OPTIMAL POLYNOMIAL RECURRENCE

NEIL LYALL ÁKOS MAGYAR

ABSTRACT. Let $P \in \mathbb{Z}[n]$ with $P(0) = 0$ and $\varepsilon > 0$. We show, using Fourier analytic techniques, that if $N \geq \exp\exp(C\varepsilon^{-1}\log\varepsilon^{-1})$ and $A \subseteq \{1, \dots, N\}$, then there must exist $n \in \mathbb{N}$ such that

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

In addition to this we also show, using the same Fourier analytic methods, that if $A \subseteq \mathbb{N}$, then the set of *ε-optimal return times*

$$R(A, P, \varepsilon) = \left\{n \in \mathbb{N} : \delta(A \cap (A + P(n))) > \delta(A)^2 - \varepsilon\right\}$$

is syndetic for every $\varepsilon > 0$. Moreover, we show that $R(A, P, \varepsilon)$ is *dense* in every sufficiently long interval, in particular we show that there exists an $L = L(\varepsilon, P, A)$ such that

$$|R(A, P, \varepsilon) \cap I| \geq c(\varepsilon, P)|I|$$

for all intervals I of natural numbers with $|I| \geq L$ and $c(\varepsilon, P) = \exp\exp(-C\,\varepsilon^{-1}\log\varepsilon^{-1})$.

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 [16] that $\mu(A \cap T^{-n}A) > 0$ for some natural number n, and hence for infinitely many.

Poincaré's result was subsequently sharpened by Khintchine [8], who showed that sets of positive measure not only return to intersect themselves infinitely often, but in fact return "frequently" with "large" intersection. In order to be more precisely we recall that a set $R \subseteq \mathbb{N}$ is said to be *syndetic* if it has bounded gaps: there exists $L \in \mathbb{N}$ such that every interval of length greater than L intersects R non-trivially. A precise formulation of Khintchine's result is that for every $\varepsilon > 0$, the set

$$\left\{n \in \mathbb{N} : \mu(A \cap T^{-n}A) > \mu(A)^2 - \varepsilon\right\} \tag{1}$$

is syndetic. 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.

In [6], Furstenberg showed that the iterates of the transformation arising in Khintchine's result can in fact be restricted to polynomial iterates, more precisely he established the following result.

Theorem A (Furstenberg [6], see also [15] or [2])**.** *Let $(X, \mathcal{M}, \mu, T)$ be an invertible measure preserving system, $A \in \mathcal{M}$ and $P \in \mathbb{Z}[n]$ with $P(0) = 0$. For every $\varepsilon > 0$, the set*

$$\{n \in \mathbb{N} : \mu(A \cap T^{-P(n)}A) > \mu(A)^2 - \varepsilon\} \tag{2}$$

is syndetic.

It follows from Furstenberg's correspondence principle (in fact, two different variants thereof) that Theorem A has the following two combinatorial consequences, in infinite and finite setting respectfully, for subsets of the natural numbers.

Corollary B (Furstenberg [6], see Frantzikinakis and Kra [4])**.** *Let $A \subseteq \mathbb{N}$ and $P \in \mathbb{Z}[n]$ with $P(0) = 0$. For every $\varepsilon > 0$, the set*

$$\left\{n \in \mathbb{N} : \delta(A \cap (A + P(n))) > \delta(A)^2 - \varepsilon\right\} \tag{3}$$

is syndetic, where $\delta(A) = \limsup_{N\to\infty} |A \cap [1, N]|/N$ denotes the upper density of any given set $A \subseteq \mathbb{N}$.

2000 *Mathematics Subject Classification.* 11B30.

Both authors were partially supported by NSF grants.

We remark that Corollary B (and Theorem 2 below) also holds if one replaces the upper density δ with the upper Banach density δ^* defined for $A \subseteq \mathbb{N}$ by $\delta^*(A) = \lim_{N\to\infty} \sup_{x\in\mathbb{N}} |A \cap (x + [1, N])|/N$.

Corollary C (Furstenberg [6], see Frantzikinakis and Kra [4])**.** *Let $P \in \mathbb{Z}[n]$ with $P(0) = 0$. For every $\varepsilon > 0$ there exists $N_1 = N_1(\varepsilon, P)$ such that if $N \geq N_1$ and $A \subseteq [1, N]$, then there exists $n \in \mathbb{N}$ such that*

$$\frac{|A \cap (A + P(n))|}{N} > \left(\frac{|A|}{N}\right)^2 - \varepsilon. \tag{4}$$

We note that Furstenberg's correspondence principle unfortunately gives no effective quantitative bounds in the finite setting of Corollary C (other than the special case when the polynomial is linear).

In this article we will be concerned with establishing, using Fourier analytic techniques, quantitative refinements of both Corollaries B and C. In addition to this we also provide what is, to the best of our knowledge, the first incidence of Fourier analysis being used to establish the syndeticity of a class of return times (results such as Corollary B).

1.2. **Further remarks.** Recently, far reaching generalizations of Furstenberg's results (Theorem A and its corollaries) have been obtained in the settings of multiple recurrence by Frantzikinakis and Kra [4]. Such results appear, however, to be inaccessible via current Fourier analytic machinary. A study of the intermediate phenomenon of simultaneous single polynomial recurrence was initiated by the authors in [13] (see also [10] and [11]) where we in particular establish, using Fourier analytic techniques and a standard decomposition theorem, Corollary C with weak tower-type bounds for $N_1(\varepsilon, P)$. The general method employed in [13] broadly follows a strategy outlined by Green and Tao in [7], we note however that in addition to the weak quantitative bounds that one obtains using this approach, these methods appears to be insufficient to establish syndeticity of optimal return times, namely results such as Corollary B.

1.3. **Statement of Main Results.** As stated above, the purpose of this article is to establish, using Fourier analytic methods, the following quantitative versions of Corollaries C and B.

1.3.1. *Quantitative refinement of Corollary C.*

Theorem 1. *Let $A \subseteq [1, N]$, $P \in \mathbb{Z}[n]$ with $P(0) = 0$ and $\varepsilon > 0$, then*

$$\left|\left\{n \in [0, L] \,:\, \frac{|A \cap (A + P(n))|}{N} > \left(\frac{|A|}{N}\right)^2 - \varepsilon\right\}\right| \geq c(\varepsilon, P)L \tag{5}$$

for all $1 \leq L \leq N^{1/k}$ where $k = \deg(P)$ and $c(\varepsilon, P) = 1/\exp\exp(C\varepsilon^{-1}\log\varepsilon^{-1})$.

In order to obtain a non-trivial conclusion from Theorem 1 we must have $L \geq c(\varepsilon, P)^{-1}$ and consequently also $N \geq c(\varepsilon, P)^{-k}$. In particular, this implies Corollary C with $N_1(\varepsilon, P) = \exp\exp(C\varepsilon^{-1}\log\varepsilon^{-1})$. We further note that Theorem 1 also gives more qualitative information as Corollary C only implies that the left side of (5) is bounded below by $c(\varepsilon, P, A)L$ with a constant that additionally depends on the set A.

1.3.2. *Quantitative refinement of Corollary B.*

Theorem 2. *Let $A \subseteq \mathbb{N}$, $P \in \mathbb{Z}[n]$ with $P(0) = 0$ and $\varepsilon > 0$, then there exists $L = L(\varepsilon, P, A)$ such that*

$$\left|\left\{n \in I \,:\, \delta(A \cap (A + P(n))) > \delta(A)^2 - \varepsilon\right\}\right| \geq c(\varepsilon, P)|I| \tag{6}$$

for all intervals I of natural numbers with $|I| \geq L$ and $c(\varepsilon, P) = 1/\exp\exp(C\,\varepsilon^{-1}\log\varepsilon^{-1})$.

Again, apart from the quantitative bounds, ergodic methods (Corollary B) imply only that the left side of (6) is bounded below by $c(\varepsilon, P, A)|I|$ with a constant again depending on the set A. We do note however that the parameter L in Theorem 2 does necessarily depend on the actual set A in question and not just on its density and include a proof of this fact in Section 7.

1.3.3. *Higher dimensional results.* The strategy we will employ to prove Theorems 1 and 2 is to lift the problem in such a way that we may then apply the following analogous higher dimensional results.

Theorem 3. *Let $B \subseteq [1,M]^k$, $\gamma(n) = (n, n^2, \dots, n^k)$ and $\varepsilon > 0$, then*

$$\left| \left\{ n \in [0,K] \,:\, \frac{|B \cap (B+\gamma(n))|}{M^k} > \left(\frac{|B|}{M^k}\right)^2 - \varepsilon \right\} \right| \geq c(\varepsilon,k)K \tag{7}$$

for all $1 \leq K \leq M^{1/k}$ with $c(\varepsilon,k) = 1/\exp\exp(C\varepsilon^{-1}\log\varepsilon^{-1})$.

Theorem 4. *Let $B \subseteq \mathbb{N}^k$, $\gamma(n) = (n, n^2, \dots, n^k)$ and $\varepsilon > 0$, then there exists $K = K(\varepsilon,k,B)$ such that*

$$\left| \left\{ n \in I \,:\, \delta(B \cap (B+\gamma(n))) > \delta(B)^2 - \varepsilon \right\} \right| \geq c(\varepsilon,k)|I| \tag{8}$$

for all intervals I of natural numbers with $|I| \geq K$ and $c(\varepsilon,k) = 1/\exp\exp(C\varepsilon^{-1}\log\varepsilon^{-1})$.

Recall that for any given set $B \subseteq \mathbb{N}^k$, it upper density $\delta(B)$ is defined to be

$$\delta(B) = \limsup_{M\to\infty} |B \cap [1,M]^k|/M^k.$$

1.4. **Outline of the paper.** The bulk of the present paper is concerned with establishing Theorems 3 and 4, from which Theorems 1 and 2 follow in an essentially straightforward manner. This approach was also used by the authors in [10] (see also [11]), however in this setting, due to the optimal nature of the results we are trying to establish, the analogous arguments are somewhat more delicate. These deductions are presented in Sections 6.1 and 6.2 respectively.

We will deduce Theorems 3 and 4 from a key dichotomy proposition, namely Proposition 2. This dichotomy proposition roughly says that either the conclusion of Theorem 3 holds for a particular interval I, or the L^2 mass of the Fourier transform of the set A is concentrated on a very thin interval depending on I. The proofs of Theorems 3 and 4 are presented in Sections 3.1 and 3.2, respectively. The arguments in these sections are close in spirit, and were influenced greatly by those of Bourgain [3] in his study of distance sets of measurable subsets of Euclidean spaces, which were first adapted to the discrete settings by the second author in [14].

In Section 4 we introduce a smooth functional variant of Proposition 2 (a technique also employed by Bourgain), the proof of which we present in Section 5 based on a certain decomposition of the indicator function of the set A. Such decompositions can in fact be obtained under very general settings via the so-called "arithmetic regularity lemma" of Green and Tao [7], and indeed have been used to proving optimal versions of Szemerédi's theorem on k-term arithmetic progressions for $k = 3$ and $k = 4$. We remark again that, while our situation is essentially simpler, results obtained this way usually give tower type bounds and do not show that the optimal return times are syndetic.

Finally, in Section 7, we include a short proof of the fact that the parameter L in Theorem 2 necessarily depends on the actual set A in question and not just its density.

We choose to begin our presentation below by recalling some basic properties of the Fourier transform on $\mathbb{Z}^k$. By observing how these properties can then be used to count differences in $B \subseteq [1,M]^k$ of the form $\gamma(t)$ will lead us naturally to the analysis of certain variants of standard Weyl sums.

Remark on Notation. We will use the letters C and c denote appropriately large or small constants, which can change from line to line.

2. Preliminaries

2.1. **Fourier analysis on $\mathbb{Z}^k$.** If $f : \mathbb{Z}^k \to \mathbb{C}$ is a function for which $\sum_{m\in\mathbb{Z}^k} |f(m)| < \infty$ we will say that $f \in L^1(\mathbb{Z}^k)$ and define

$$\|f\|_1 = \sum_{m\in\mathbb{Z}^k} |f(m)|.$$

For $f \in L^1$ we define its *Fourier transform* $\widehat{f} : \mathbb{T}^k \to \mathbb{C}$ by

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

Note that the summability assumption on f ensures that the series defining $\widehat{f}$ converges uniformly to a continuous function on $\mathbb{T}^k$ (which we will identify with the unit cube $[0,1)^k$ in $\mathbb{R}^k$) and that the Fourier inversion formula and Plancherel's identity, namely

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

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

$$\int_{\mathbb{T}^k} e^{2\pi i m\cdot\alpha} d\alpha = \begin{cases} 1 & \text{if } \ m = 0 \\ 0 & \text{if } \ m \neq 0 \end{cases} .$$

Defining the convolution of f and g to be

$$f * g(m) = \sum_{\ell\in\mathbb{Z}^k} f(m-\ell) g(\ell)$$

it follows that if $f, g \in L^1$ then $f * g \in L^1$ with

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1 \qquad \text{and} \qquad \widehat{f * g} = \widehat{f}\,\widehat{g}.$$

Finally, we remark that it follows from the Poisson Summation Formula that if $\varphi \in \mathcal{S}(\mathbb{R}^k)$, then

$$\widehat{\varphi}(\alpha) = \sum_{\ell\in\mathbb{Z}^k} \widetilde{\varphi}(\alpha - \ell) \tag{9}$$

where

$$\widetilde{\varphi}(\xi) = \int_{\mathbb{R}^k} \varphi(x) e^{-2\pi i x\cdot\xi}\, dx \tag{10}$$

denotes the Fourier transform (on $\mathbb{R}^k$) of φ.

2.2. **Counting differences of the form $\gamma(n) = (n, n^2, \ldots, n^k)$.** Let $B \subseteq [1, M]^k$ and $\delta = |B|/M^k$.

Let $1 \leq \mu \leq \lambda$ be integers with $\lambda^k \leq M/2^k$. It is easy to verify, using the properties of the Fourier transform discussed above, that the average number of pairs of elements in B whose difference is equal to $\gamma(n)$ with $n \in (\lambda, \lambda+\mu] \cap \mathbb{Z}$ can be expressed as follows:

$$\frac{1}{\mu} \sum_{n=\lambda+1}^{\lambda+\mu} |B \cap (B + \gamma(n))| = \frac{1}{\mu} \sum_{n=\lambda+1}^{\lambda+\mu} \sum_{m\in\mathbb{Z}^k} 1_B(m) 1_B(m - \gamma(n)) = \int_{\mathbb{T}^k} |\widehat{1_B}(\alpha)|^2 S_{\lambda,\mu}(\alpha)\, d\alpha \tag{11}$$

where

$$S_{\lambda,\mu}(\alpha) = \frac{1}{\mu} \sum_{n=\lambda+1}^{\lambda+\mu} e^{2\pi i \alpha\cdot\gamma(n)}. \tag{12}$$

It is easy to see that

$$S_{\lambda,\mu}(\alpha) = \frac{\lambda+\mu}{\mu} S_{\lambda+\mu}(\alpha) - \frac{\lambda}{\mu} S_\lambda(\alpha) \tag{13}$$

where

$$S_\mu(\alpha) = \frac{1}{\mu} \sum_{n=1}^{\mu} e^{2\pi i \alpha\cdot\gamma(n)} \tag{14}$$

denotes a classical (normalized) Weyl sum. Unfortunately, the rather simplistic relationship indicated in (13) will only be useful to us in the case where $\mu = \lambda$. When $\mu < \lambda$ we will make use of the following alternative:

$$S_{\lambda,\mu}(\alpha) = e^{2\pi i \alpha\cdot\gamma(\lambda)} S_\mu(T_\lambda \alpha) \tag{15}$$

where T_λ is a $k \times k$ matrix whose entries are given by

$$(T_\lambda)_{ij} = \begin{cases} \binom{j}{i} \lambda^{j-i} & j \geq i \\ 0 & j < i \end{cases} . \tag{16}$$

2.3. **Standard Weyl sum estimates.** It is clear that whenever $|\alpha_j| \ll \mu^{-j}$ there can be no cancellation in the Weyl sum (14), in fact it is easy to verify that the same is also true whenever each α_j is *close* to a rational with *small* denominator (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. For a proof of this result see either [10] or [11].

Lemma 1. *Let $\eta > 0$ and $\mu \geq \eta^{-C}$ (with C sufficiently large depending on k). If for some $1 \leq j \leq k$ we have*

$$\left|\alpha_j - \frac{a}{q}\right| > \frac{1}{\eta^k \mu^j} \tag{17}$$

for all $a \in \mathbb{Z}$ and $1 \leq q \leq \eta^{-k}$, then

$$|S_\mu(\alpha)| \leq C_1 \eta. \tag{18}$$

Remark. It follows immediately from Lemma 1 and (15) that

$$|S_{\lambda,\mu}(\alpha)| \leq C_1 \eta \tag{19}$$

for any $\eta > 0$, where $1 \leq \mu \leq \lambda$ are integers with $\mu \geq \eta^{-C}$, provided that for some $1 \leq j \leq k$

$$\left|(T_\lambda \alpha)_j - \frac{a}{q}\right| > \frac{1}{\eta^k \mu^j} \tag{20}$$

for all $a \in \mathbb{Z}$ and $1 \leq q \leq \eta^{-k}$. In fact, it is easy to see that one can conclude from Lemma 1 that estimate (19) also holds (under the same hypotheses as above with C_1 replaced with say $2C_1$) for the "perturbed" Weyl sums

$$\frac{1}{\mu} \sum_{n \in (\lambda, \lambda+\mu] \cap \mathbb{Z}} e^{2\pi i \alpha \cdot \gamma(n)}$$

where $1 \leq \mu \leq \lambda$ are now no longer assumed to take on integer values, provided C is chosen sufficiently large.

We note that from Lemma 1, relationship (15) and the Plancherel identity, we may conclude that

$$\int_{\mathbb{T}^k} |\widehat{1_B}(\alpha)|^2 S_{\lambda,\mu}(\alpha)\, d\alpha = \int_{T_\lambda^{-1} \mathfrak{M}_{\eta,\mu}} |\widehat{1_B}(\alpha)|^2 S_{\lambda,\mu}(\alpha)\, d\alpha + O(\eta M^k)$$

where

$$\mathfrak{M}_{\eta,\mu} = \bigcup_{q=1}^{\eta^{-k}} \left\{ \alpha \in \mathbb{T}^k \,:\, \left|\alpha_j - \frac{a_j}{q}\right| \leq \frac{1}{\eta^k \mu^j} \ (1 \leq j \leq k) \text{ for some } a \in \mathbb{Z}^k \right\}.$$

While in the case $\mu = \lambda$ it follows from (13) that

$$|S_{\lambda,\lambda}(\alpha)| \leq 3C_1 \eta$$

whenever $\alpha \notin \mathfrak{M}_{\eta,\lambda}$ and as a consequence of this we can in fact make the rather more favorable conclusion that

$$\int_{\mathbb{T}^k} |\widehat{1_B}(\alpha)|^2 S_{\lambda,\lambda}(\alpha)\, d\alpha = \int_{\mathfrak{M}_{\eta,\lambda}} |\widehat{1_B}(\alpha)|^2 S_{\lambda,\lambda}(\alpha)\, d\alpha + O(\eta M^k).$$

In order to carry out our Fourier analytic arguments it will be convenient to consider the (nonisotropic) lattice

$$\left\{ \left(\frac{a_1}{q_\eta}, \frac{a_2}{q_\eta^2}, \ldots, \frac{a_k}{q_\eta^k} \right) \in \mathbb{T}^k \,:\, (a_1, \ldots, a_k) \in \mathbb{Z}^k \right\}$$

of rational points where

$$q_\eta = \operatorname{lcm}\{1 \leq q \leq \eta^{-k}\} \tag{21}$$

as opposed to the much smaller, but alas more wildly distributed, set of rational points that appear as the centers of the major boxes in $\mathbb{T}^k$ that constitute $\mathfrak{M}_{\eta,\mu}$. Note that it follows from elementary considerations involving the prime numbers that

$$q_\eta \leq \exp(C\eta^{-k}) \tag{22}$$

and this accounts for one of the exponentials in the bound in Theorems 3 and 4 (as well as 1 and 2).

3. Reduction to dichotomy propositions

We now separately present the statement of two key propositions (although as we shall see the first of which follows immediately from the second) and demonstrate how they can be used to prove Theorem 3 and Theorem 4 respectively.

For $L > 1$ and $q \in \mathbb{N}$ we define

$$M_{q,L} = \left\{ \alpha \in \mathbb{T}^k \,:\, \left| \alpha_j - \frac{a_j}{q^j} \right| \leq \frac{1}{L^j} \ (1 \leq j \leq k) \text{ for some } a \in \mathbb{Z}^k \right\}. \tag{23}$$

Let $0 < \eta < 1$ and $1 \leq \mu \leq \lambda$. We define

$$\Omega_{\eta,\lambda,\mu} = \left\{ \alpha \in \mathbb{T}^k \,:\, \alpha \in M_{q_\eta, \eta^k \mu} \setminus M_{q_\eta, \eta^{-k}\lambda} \right\} \tag{24}$$

where $q_\eta = \operatorname{lcm}\{1 \leq q \leq \eta^{-k}\}$ as before.

3.1. **Proof of Theorem 3.** Although this result can in fact be deduced from the second dichotomy proposition (Proposition 2 below), we feel that the reduction of Theorem 3 to the (simpler) Proposition 1 is not only more direct and straightforward (by virtue of the fact that we can take $\mu = \lambda$), but that our decision to include it will also serve to illuminate the deduction of Theorem 4 from Proposition 2.

Proposition 1. *Let $B \subseteq [1, M]^k$ and $\varepsilon > 0$. Let $\eta_\varepsilon = \exp(-C\varepsilon^{-1} \log \varepsilon^{-1})$ and $q_\varepsilon = q_{\eta_\varepsilon}$.*

If λ is an integer that satisfies $\lambda \geq \eta_\varepsilon^{-k} q_\varepsilon$ and $M \geq C(\eta_\varepsilon^{-k}\lambda)^k$ then either

$$\left| \left\{ n \in (\lambda, 2\lambda] \cap \mathbb{Z} \,:\, \frac{|B \cap (B + \gamma(n))|}{M^k} > \left(\frac{|B|}{M^k} \right)^2 - \varepsilon \right\} \right| \geq \exp(-C\eta_\varepsilon^{-k})\lambda \tag{25}$$

or

$$\int_\Omega |\widehat{1_B}(\alpha)|^2 \, d\alpha \geq \varepsilon M^k / 10 \tag{26}$$

where $\Omega = \Omega_{\eta_\varepsilon,\lambda,\lambda}$.

Proposition 1 (and Proposition 2 below) both express, in our setting, the basic dichotomy that either B behaves as though it were a random set, or has arithmetic structure as the Fourier transform $\widehat{1_B}$ is concentrated (on small annuli) around a fixed (nonisotropic) lattice of rational points.

Proof of Theorem 3. Let $\varepsilon > 0$, $\eta_\varepsilon = \exp(-C\varepsilon^{-1} \log \varepsilon^{-1})$ and $q_\varepsilon = q_{\eta_\varepsilon}$. Suppose K and M are integers that satisfy

$$\exp(C\eta_\varepsilon^{-k}) \leq K \leq M^{1/k}$$

and $\{\lambda_j\}_{j=1}^J$ is a sequence of integers with $J = \lceil 10/\varepsilon \rceil$ with the property that $\lambda_1 \geq \eta_\varepsilon^{-k} q_\varepsilon$, $\lambda_J = c\eta_\varepsilon^k K$ and

$$\eta_\varepsilon^{-2k} \lambda_j \leq \lambda_{j+1} \leq C\eta_\varepsilon^{-2k} \lambda_j \tag{27}$$

for $1 \leq j \leq J$. That such a lacunary sequence of length $J > 10/\varepsilon$ exists is ensured by, by choosing C large enough, that $K \gg q_\varepsilon$. It is also easy to see that (27) ensures that the sets $\Omega_j = \Omega_{\eta_\varepsilon,\lambda_j,\lambda_j}$ are disjoint.

Suppose, contrary to Theorem 3, that there does exists a set $B \subseteq [1, M]^k$ such that

$$\left| \left\{ n \in [0, K] \,:\, \frac{|B \cap (B + \gamma(n))|}{M^k} > \left(\frac{|B|}{M^k} \right)^2 - \varepsilon \right\} \right| < \exp(-C\eta_\varepsilon^{-k})K \tag{28}$$

for all $C > 0$. Since $\lambda_1 \geq (C\eta_\varepsilon^{-2k})^{1-J}$, it follows that $K/\lambda_j \leq (C\eta_\varepsilon)^{-2kJ} \ll \exp(-C\eta_\varepsilon^{-k})$ for all $1 \leq j \leq J$, and hence that

$$\left| \left\{ n \in (\lambda_j, 2\lambda_j] \cap \mathbb{Z} \,:\, \frac{|B \cap (B + \gamma(n))|}{M^k} > \left(\frac{|B|}{M^k} \right)^2 - \varepsilon \right\} \right| < \exp(-C\eta_\varepsilon^{-k})\lambda_j \tag{29}$$

for all $C > 0$ and all $1 \leq j \leq J$.

Proposition 1 allows us to conclude from this that

$$\sum_{j=1}^J \int_{\Omega_j} |\widehat{1_B}(\alpha)|^2 \, d\alpha \geq J\varepsilon M^k / 10 > M^k. \tag{30}$$

On the other hand it follows from the disjointness property of the sets Ω_j (which we guaranteed by our initial choice of sequence $\{\lambda_j\}$) and the Plancherel identity that

$$\sum_{j=1}^{J}\int_{\Omega_j}|\widehat{1_B}(\alpha)|^2\,d\alpha\leq\int_{\mathbb{T}^k}|\widehat{1_B}(\alpha)|^2\,d\alpha\leq|B|\leq M^k \tag{31}$$

giving us our desired contradiction. □

3.2. **Proof of Theorem 4.** We now present the statement of our second (stronger) dichotomy proposition.

Proposition 2. *Let $B\subseteq[1,M]^k$ and $\varepsilon>0$. Let $\eta_\varepsilon=\exp(-C\varepsilon^{-1}\log\varepsilon^{-1})$ and $q_\varepsilon=q_{\eta_\varepsilon}$.*

If $1\leq\mu\leq\lambda$ are any given pair of integers that satisfy $\mu\geq\eta_\varepsilon^{-k}q_\varepsilon$ and $M\geq C(\eta_\varepsilon^{-k}\lambda)^k$ then either

$$\left|\left\{n\in(\lambda,\lambda+\mu]\cap\mathbb{Z}\,:\,\frac{|B\cap(B+\gamma(n))|}{M^k}>\left(\frac{|B|}{M^k}\right)^2-\varepsilon\right\}\right|\geq\exp(-C\eta_\varepsilon^{-k})\mu \tag{32}$$

or

$$\int_{T_\lambda^{-1}\Omega}|\widehat{1_B}(\alpha)|^2\,d\alpha\geq\varepsilon M^k/10 \tag{33}$$

where $\Omega=\Omega_{\eta_\varepsilon,\lambda,\mu}$.

Key to deducing Theorem 4 from Proposition 2 is the following combinatorial result on the annuli $\Omega_{\eta_\varepsilon,\lambda,\mu}$.

Lemma 2 (Overlapping Lemma)**.** *Let $\eta>0$. Suppose $\{\mu_j\}_{j\in\mathbb{N}}$ and $\{\lambda_j\}_{j\in\mathbb{N}}$ be sequences such that $\mu_1\geq\eta^{-k}q_\eta$ and*

$$\mu_j\leq\lambda_j\leq\frac{1}{3}\,\eta^{2k}\mu_{j+1} \tag{34}$$

for all $j\in\mathbb{N}$, then it follows that

$$\alpha\in T_{\lambda_j}^{-1}\Omega_j$$

for at most k different values of j, where $\Omega_j=\Omega_{\eta,\lambda_j,\mu_j}$.

The proof of this result is given in the Section 3.3 below.

Proof of Theorem 4. Let $\varepsilon>0$, $\eta_\varepsilon=\exp(-C\varepsilon^{-1}\log\varepsilon^{-1})$ and $q_\varepsilon=q_{\eta_\varepsilon}$. Suppose, contrary to Theorem 4, that there exists a set $B\subseteq\mathbb{N}^k$ with $\delta=\delta(B)>\varepsilon^{1/2}$ such that for all $K>0$, there exists an interval of natural numbers I with $|I|\geq K$ such that

$$\left|\left\{n\in I\,:\,\delta(B\cap(B+\gamma(n)))>\delta(B)^2-\varepsilon\right\}\right|<\exp(-C\eta_\varepsilon^{-k})|I| \tag{35}$$

for all $C>0$. In this case there necessarily exists a sequence of intervals of natural numbers $I_j=(\lambda_j,\lambda_j+\mu_j]$ with $\mu_1\geq4q_\eta$ and $\mu_j\nearrow\infty$ for which

$$\left|\left\{n\in I_j\,:\,\delta(B\cap(B+\gamma(n)))>\delta(B)^2-\varepsilon\right\}\right|<\exp(-C\eta_\varepsilon^{-k})|I_j| \tag{36}$$

for all $C>0$ and all $j\in\mathbb{N}$.

Since inequality (36) must then also hold for the *right-half* intervals $I_j'=(\lambda_j',\lambda_j'+\mu_j']$, where $\mu_j'=\mu_j/2$ and $\lambda_j'=\lambda_j+\mu_j'$, we see that we can further assume that $\lambda_j\to\infty$. By passing to a subsequence, one may without loss in generality assume that $\mu_1\geq\eta^{-k}q_\varepsilon$ and

$$\mu_j\leq\lambda_j\leq\frac{1}{3}\,\eta^{2k}\mu_{j+1}$$

for all $j\in\mathbb{N}$. We now fix an integer $J>40k/\varepsilon$.

It follows from the definition of upper density that there must exist $M\in\mathbb{N}$ such that

$$\left|B\cap[1,M]^k\right|\geq(\delta-\varepsilon/2)M^k$$

while

$$\left|B\cap(B+\gamma(n))\cap[1,M]^k\right|\leq(\delta^2-\varepsilon)M^k$$

for all $n \in \bigcup_{j=1}^{J} I_j$ for which $\delta(B \cap (B+\gamma(n))) \leq \delta(B)^2 - \varepsilon$. Letting $B' = B \cap [1,M]^k$ it follows that

$$\left| \left\{ n \in I_j \,:\, \frac{|B' \cap (B'+\gamma(n))|}{M^k} > \left(\frac{B'}{M^k}\right)^2 - \frac{\varepsilon}{4} \right\} \right| < \exp(-C\eta_\varepsilon^{-k})|I_j|$$

for all $C > 0$ and all $1 \leq j \leq J$.

Proposition 2 allows us to conclude from this that

$$\sum_{j=1}^{J} \int_{T_{\lambda_j}^{-1}\Omega_j} |\widehat{1_B}(\alpha)|^2 \, d\alpha \geq J\varepsilon M^k/40 > kM^k. \tag{37}$$

On the other hand it follows from Lemma 2 (with $\Omega_j = \Omega_{\eta_\varepsilon,\lambda_j,\mu_j}$) and the Plancherel identity that

$$\sum_{j=1}^{J} \int_{T_{\lambda_j}^{-1}\Omega_j} |\widehat{1_B}(\alpha)|^2 \, d\alpha \leq k \int_{\mathbb{T}^k} |\widehat{1_B}(\alpha)|^2 \, d\alpha \leq k|B| \leq kM^k \tag{38}$$

giving us our desired contradiction. □

3.3. **Proof of the Overlapping Lemma.** First we establish the following.

Lemma 3. *Suppose that $0 < \eta < 1/4k^2$ and $0 < \mu \leq \lambda$. If $\alpha \in T_\lambda^{-1}\Omega$, where $\Omega = \Omega_{\eta,\lambda,\mu}$, then there exist $1 \leq i \leq k$ and $a \in \mathbb{Z}$ such that*

$$\frac{1}{2}\left(\frac{\eta^k}{\lambda}\right)^i \leq \left|\alpha_i - \frac{a}{q_\eta^i}\right| \leq \frac{3}{2}\left(\frac{1}{\eta^k\mu}\right)^i. \tag{39}$$

Proof of Lemma 3. Suppose that $T_\lambda\alpha \in \Omega$, then for some $1 \leq j \leq k$ we have

$$\left|(T_\lambda\alpha)_j - \frac{a_j}{q_\eta^j}\right| \geq \left(\frac{\eta^k}{\lambda}\right)^j \tag{40}$$

for all $a \in \mathbb{Z}^k$, while for all $1 \leq j \leq k$ we have

$$\left|(T_\lambda\alpha)_j - \frac{a'_j}{q_\eta^j}\right| \leq \left(\frac{1}{\eta^k\mu}\right)^j \tag{41}$$

for some $a' \in \mathbb{Z}^k$. If we denote by i the largest integer from $\{1,\dots,k\}$ for which

$$\left|(T_\lambda\alpha)_i - \frac{a}{q_\eta^{i-1}}\right| \geq \left(\frac{\eta^k}{\lambda}\right)^i$$

for all $a \in \mathbb{Z}$, then we must have

$$\left|(T_\lambda\alpha)_{i+1} - \frac{a''_i}{q_\eta^i}\right| < \left(\frac{\eta^k}{\lambda}\right)^{i+1}$$

for some $a'' \in \mathbb{Z}^k$. In light of the fact that for each $1 \leq j \leq k$ one can write

$$(T_\lambda\alpha)_j = \alpha_j + \binom{j+1}{j}\lambda\alpha_{j+1} + \cdots + \binom{k}{j}\lambda^{k-j}\alpha_k = \alpha_j + c_j\lambda(T_\lambda\alpha)_{j+1}$$

where $0 < c_j \leq j(k-j) < k^2$, it therefore follows that

$$\left|(T_\lambda\alpha - \alpha)_i - \frac{a'''_i}{q_\eta^i}\right| < 2c_j\eta^k\left(\frac{\eta^k}{\lambda}\right)^i \leq \frac{1}{2}\left(\frac{\eta^k}{\lambda}\right)^i < \frac{1}{2}\left(\frac{1}{\eta^k\mu}\right)^i \tag{42}$$

for some $a''' \in \mathbb{Z}^k$.

We note that it follows immediately from (42) and (41) that

$$\left|\alpha_i - \frac{a_i}{q_\eta^i}\right| \leq \frac{3}{2}\left(\frac{1}{\eta^k\mu}\right)^i.$$

for some $a \in \mathbb{Z}^k$, while from (42) and (40) it follows that for all $a \in \mathbb{Z}^k$ we have

$$\left|\alpha_i - \frac{a_i}{q_\eta^i}\right| \geq \frac{1}{2}\left(\frac{\eta^k}{\lambda}\right)^i. \qquad \square$$

Proof of Lemma 2. If $T_{\lambda_j}\alpha \in \Omega_j$, then Lemma 3 guarantees the existence of an integer $1 \leq i_j \leq k$ such that

$$\frac{1}{2}\left(\frac{\eta^k}{\lambda_j}\right)^{i_j} \leq \left|\alpha_{i_j} - \frac{a}{q}\right| \leq \frac{3}{2}\left(\frac{1}{\eta^k \mu_j}\right)^{i_j} \tag{43}$$

for some $a \in \mathbb{Z}$. Suppose there exists $\alpha \in \mathbb{T}^k$ and distinct integers $j_1, \ldots, j_{k+1}$ for which

$$\alpha \in T_{\lambda_{j_1}}^{-1}\Omega_{j_1} \cap \cdots \cap T_{\lambda_{j_{k+1}}}^{-1}\Omega_{j_{k+1}}.$$

It follows from the pigeonhole principle that there must exists integers $j, j' \in \{j_1, \ldots, j_{k+1}\}$, with $j < j'$, for which $i_j = i_{j'}$. Inequality (43) and the fact that $\mu_1 \geq \eta^{-k} q_\eta$ the forces the situation that

$$\eta^{2k}\mu_{j'} < 3\lambda_j$$

which contradicts (34). $\square$

4. Formulation of smooth variants of Propositions 1 and 2

We now formulate *smooth* functional variants of Proposition 1 and 2 that are better suited to our Fourier analytic approach.

4.1. **Counting function.** For $g, h : [1, M]^k \to [0, 1]$ and $q, \lambda, \mu \in \mathbb{N}$ we define

$$\Lambda_{q,\mu}(g, h) = \frac{q}{\mu} \sum_{\substack{n \in (\lambda, \lambda+\mu] \\ q|n}} \sum_{m \in \mathbb{Z}^k} g(m)h(m - \gamma(n)). \tag{44}$$

With $g = h = 1_B$ this essentially gives a normalized count for the number of pairs of elements in B whose difference is equal to $\gamma(n)$ with $n \in (\lambda, \lambda + \mu] \cap \mathbb{Z}$ and $q|n$.

Note that it is natural to consider only those $n \in \mathbb{N}$ that are divisible by some (large) natural number q. Indeed, as a consequence of the fact that our set B could fall entirely into a subset of $\mathbb{Z}^k$ of the form $x + d\,\mathbb{Z} \times \mathbb{Z}^{k-1}$ with $1 \leq d \leq \varepsilon^{-1/2}$, it follows that if there were to exist $n \in \mathbb{N}$ such that $B \cap (B + \gamma(n)) \neq \emptyset$ for an arbitrary set B, then these n would necessarily have to be divisible by all $1 \leq d \leq \varepsilon^{-1/2}$ and hence by the least common multiple of all $1 \leq d \leq \varepsilon^{-1/2}$, a quantity of size $\exp(C\varepsilon^{-1/2})$.

As before this can be expressed this count on the transform side as

$$\Lambda_{q,\mu}(g, h) = \int_{\mathbb{T}^k} \widehat{g}(\alpha)\overline{\widehat{h}(\alpha)} S_{\lambda,\mu,q}(\alpha)\, d\alpha \tag{45}$$

where

$$S_{\lambda,\mu,q}(\alpha) = \frac{q}{\mu} \sum_{\substack{n \in (\lambda, \lambda+\mu] \\ q|n}} e^{2\pi i\, \alpha \cdot \gamma(n)}. \tag{46}$$

Remark. If the integers λ and μ are both divisible by q, then one can easily verify that

$$S_{\lambda,\mu,q}(\alpha) = S_{\lambda/q,\mu/q}(q \circ \alpha) \tag{47}$$

where

$$q \circ \alpha = (q\alpha_1, \ldots, q^k \alpha_k) \tag{48}$$

and as such we can deduce estimates for these new exponential sums, via relations (13) and (15), from those that are stated in Lemma 1. See in particular Lemma 6 below.

4.2. **Smooth variants of our dichotomy propositions.** Let $\varphi : \mathbb{R}^k \to (0,\infty)$ be a Schwartz function satisfying

$$\widetilde{\varphi}(0) = 1 \geq \widetilde{\varphi}(\xi) \geq 0 \qquad \text{and} \qquad \widetilde{\varphi}(\xi) = 0 \ \text{ for } \ |\xi| > 1$$

where $\widetilde{\varphi}$ denotes the Fourier transform (on $\mathbb{R}^k$) of φ, see (10).

For a given $q \in \mathbb{N}$ and $L > 1$ we define

$$\varphi_{q,L}(x) = \begin{cases} \left(\frac{q}{L}\right)^{k(k-1)/2} \varphi\left(\frac{q\ell_1}{L}, \dots, \frac{q^k \ell_k}{L^k}\right) & \text{if } \ x = (q\ell_1, \dots, q^k\ell_k) \ \text{ for some } \ell \in \mathbb{Z}^k \\ 0 & \text{otherwise} \end{cases} \tag{49}$$

It follows from the Poisson summation formula that the Fourier transform (on $\mathbb{Z}^k$) of $\varphi_{q,L}$ takes the form

$$\widehat{\varphi}_{q,L}(\alpha) = \sum_{\ell \in \mathbb{Z}^k} \widetilde{\varphi}\left(L\left(\alpha_1 - \frac{\ell_1}{q}\right), \dots, L^k\left(\alpha_k - \frac{\ell_k}{q^k}\right)\right). \tag{50}$$

We now define

$$\psi_{q,L}(x) = \varphi_{q,L}(T_\lambda^{*^{-1}} x)$$

where T_λ^* denotes the adjoint of T_λ, and note that

$$\widehat{\psi}_{q,L}(\alpha) = \widehat{\varphi}_{q,L}(T_\lambda \alpha).$$

Note that $\widehat{\varphi}_{q,L}$ is supported on $M_{q,L}$ (and hence $\widehat{\psi}_{q,L}$ is supported on $T_\lambda^{-1} M_{q,L}$), where $M_{q,L}$ are the major boxes defined by (23), and that we may choose our cutoff function φ so that

$$\widehat{\varphi}_{q_\varepsilon, \eta_\varepsilon^k \mu} - \widehat{\varphi}_{q_\varepsilon, \varepsilon\eta_\varepsilon^{-k}\lambda} \tag{51}$$

will be essentially supported on $\Omega_{\eta_\varepsilon,\lambda,\mu}$ in the sense that

$$\left|\widehat{\varphi}_{q_\varepsilon, \eta_\varepsilon^k \mu}(\alpha) - \widehat{\varphi}_{q_\varepsilon, \varepsilon\eta_\varepsilon^{-k}\lambda}(\alpha)\right| \leq \varepsilon/10 \tag{52}$$

whenever $\alpha \notin \Omega_{\eta_\varepsilon,\lambda,\mu}$.

The smooth variant of Proposition 2 is then the following:

Proposition 3 (Smooth variant of Proposition 2)**.** *Let $f : [1,M]^k \to [0,1]$ and set $\delta = M^{-k}\sum_{m\in\mathbb{Z}^k} f(m)$.*

Let $0 < \varepsilon \leq \delta^2$ and $1 \leq \mu \leq \lambda$ be any given pair of integers that satisfy $\mu \geq \eta_\varepsilon^{-k} q_\varepsilon$ and $M \geq C(\eta_\varepsilon^{-k}\lambda)^k$ where $q_\varepsilon = q_{\eta_\varepsilon}$ with $\eta_\varepsilon = \exp(-C\varepsilon^{-1}\log\varepsilon^{-1})$. Then there exists $0 < \eta \ll \varepsilon$ satisfying $\eta_\varepsilon \leq \varepsilon\eta$, such that either

$$\Lambda_{q,\mu}(f,f) > (\delta^2 - \varepsilon)M^k \tag{53}$$

or

$$\int_{\mathbb{T}^k} |\widehat{f}(\alpha)|^2 \left|\widehat{\psi}_{q,L_2}(\alpha) - \widehat{\psi}_{q,L_1}(\alpha)\right| d\alpha \geq \varepsilon M^k/5 \tag{54}$$

where $L_1 = \eta^{-k}\lambda$, $L_2 = \eta^k\mu$, and $q = q_\eta$.

Remark. We have chosen to not explicitly state the analogous smooth variant of Proposition 1, since this would be simply Proposition 3 with μ set equal to λ and ψ replaced with φ.

We finish this section by explicitly showing that Proposition 3 does indeed imply Proposition 2, the same argument of course also establishes that Proposition 1 would follow from its (unstated) analogous smooth variant.

Proof that Proposition 3 implies Proposition 2. Let $f = 1_B$ and $q = q_\eta$, noting that and $q \leq q_\varepsilon$.

It is easy to see that if $\Lambda_{q,\mu}(f,f) > (\delta^2 - \varepsilon)M^k$, then

$$\left|\left\{n \in (\lambda, \lambda+\mu] \cap \mathbb{Z} \,:\, |B \cap (B + \gamma(n))| > (\delta^2 - 2\varepsilon)M^k\right\}\right| \geq \frac{c\varepsilon}{q}\mu \geq \frac{c\varepsilon}{q_\varepsilon}\mu \tag{55}$$

which immediately gives (32), with 2ε in place of ε, since $q_\varepsilon \leq \exp(C\eta_\varepsilon^{-k})$.

While from the fact that $q | q_\varepsilon$ it follows that

$$\operatorname{supp}\big(\widehat{\psi}_{q,L_2} - \widehat{\psi}_{q,L_1}\big) \subseteq \operatorname{supp}\big(\widehat{\psi}_{q_\varepsilon,\eta_\varepsilon^k\mu} - \widehat{\psi}_{q_\varepsilon,\varepsilon\eta_\varepsilon^{-k}\lambda}\big)$$

and hence from the remarks preceding Proposition 3 (in particular (52)) that (54) implies (33). □

5. Proof of Proposition 3

We now present the proof of Proposition 3, finally completing the proofs of Theorems 3 and 4. As opposed to the usual Fourier proofs of Sárközy's theorem, which are based on density increment arguments, here we use an energy increment argument, (in fact a regularity lemma type decomposition) to obtain optimal recurrence.

Remark. We have already noted that in order to establish Theorem 3 we need only prove Proposition 3 with $\mu = \lambda$ and ψ replaced with φ. Making these substitutions in the proof below will indeed give a proof of the (unstated) smooth variant of Proposition 1 (one must also, in the proof of Lemma 6, (naturally) replace T_λ with the identity matrix and increase the size of some constants threefold).

5.1. **Decomposition.** Let $f : [1, M]^k \to [0, 1]$ and $\delta = M^{-k} \sum_{m\in[1,M]^k} f(m)$.

We make the decomposition

$$f = f_1 + f_2 + f_3 \tag{56}$$

where

$$f_1 = f * \psi_{q,L_1} \quad \text{and} \quad f_2 = f - f * \psi_{q,L_2} \tag{57}$$

which of course forces

$$f_3 = f * (\psi_{q,L_2} - \psi_{q,L_1}). \tag{58}$$

One should think of $f_1(m)$ (respectively $f*\psi_{q,L_2}(m)$) as being essentially the average value of the function f over arithmetic grids of the form $\{q_\eta \circ \ell \, : \, \ell \in \mathbb{Z}^k\}$ of (total) size $L_1 \times L_1^2 \times \cdots \times L_1^k$ (respectively $L_2 \times L_2^2 \times \cdots \times L_2^k$) centered at m.

5.2. **Proof of Proposition 3.** Note that

$$\Lambda_{q,\mu}(f, f) = \Lambda_{q,\mu}(f_1, f_1) + \underbrace{\Lambda_{q,\mu}(f_2, f_1) + \Lambda_{q,\mu}(f, f_2)}_{(\star)} + \underbrace{\Lambda_{q,\mu}(f_3, f_1) + \Lambda_{q,\mu}(f, f_3)}_{(\star\star)} \tag{59}$$

where both terms in $(\star)$ involve a f_2 and both terms in $(\star\star)$ involve a f_3.

The proof of Proposition 3 will follow as an almost immediate consequence of the following two lemmas.

Lemma 4 (Main term)**.** *Let $\varepsilon > 0$. If $0 < \eta \ll \varepsilon$, then*

$$\Lambda_{q,\mu}(f_1, f_1) \geq (\delta^2 - \varepsilon/2)\, M^k \tag{60}$$

Lemma 5 (Error term)**.** *Let $\varepsilon > 0$, then there exists $\eta > 0$ satisfying $\exp(-C\varepsilon^{-1} \log \varepsilon^{-1}) \leq \eta \ll \varepsilon$, such that*

$$\int_{\mathbb{T}^k} |\widehat{f}(\alpha)|^2 |(1 - \widehat{\psi}_{q,L_2}(\alpha)) S_{\lambda,\mu,q}(\alpha)|\, d\alpha \leq (\varepsilon/20) M^k \tag{61}$$

and hence

$$|\Lambda_{q,\mu}(f_2, f_1) + \Lambda_{q,\mu}(f, f_2)| \leq (\varepsilon/10) M^k. \tag{62}$$

Proof of Proposition 3. If $\Lambda_{q,\mu}(f, f) \leq (\delta^2 - \varepsilon)M^k$, then it follows from Lemma 4 that

$$|\Lambda_{q,\mu}(f, f) - \Lambda_{q,\mu}(f_1, f_1)| \geq (\varepsilon/2)M^k.$$

Since

$$|\Lambda_{q,\mu}(f_3, f_1) + \Lambda_{q,\mu}(f, f_3)| \geq |\Lambda_{q,\mu}(f, f) - \Lambda_{q,\mu}(f_1, f_1)| - |\Lambda_{q,\mu}(f_2, f_1) + \Lambda_{q,\mu}(f, f_2)|$$

it consequently follows from Lemma 5 that

$$|\Lambda_{q,\mu}(f_3, f_1) + \Lambda_{q,\mu}(f, f_3)| \geq (2\varepsilon/5)M^k.$$

The proposition then follows from the observation that

$$\max\{|\Lambda_{q,\mu}(f_3, f_1)|, |\Lambda_{q,\mu}(f, f_3)|\} \leq \int_{\mathbb{T}^k} |\widehat{f}(\alpha)|^2 \big|\widehat{\psi}_{q,L_2}(\alpha) - \widehat{\psi}_{q,L_1}(\alpha)\big|\, d\alpha. \tag{63}$$

which follows from standard properties of convolutions under the action of the Fourier transform, identity (45), and trivial bounds for the exponential sum $S_{\lambda,\mu,q}$. □

5.3. **Proof of Lemma 4.** Let $q = q_\eta$. If $q|n$ and $\lambda < n \leq \lambda + \mu$ (and hence $n \leq 2\eta^k L_1$), then it is straightforward to see that φ can be chosen so that f_1 is essentially invariant under translation by $\gamma(n)$ in the the sense that

(64)
$$|f_1(m) - f_1(m - \gamma(n))| = \left(\frac{q}{L_1}\right)^{k(k-1)/2} \sum_{\ell \in \mathbb{Z}^k} \left|\varphi\left(\frac{q\ell_1 - n}{L_1}, \ldots, \frac{q^k\ell_k - n^k}{L_1^k}\right) - \varphi\left(\frac{q\ell_1}{L_1}, \ldots, \frac{q^k\ell_k}{L_1^k}\right)\right| \leq c\eta^k$$

for some constant $c > 0$.

Therefore, provided η is chosen so that $c\eta^k \leq \varepsilon/4$, we have

$$\Lambda_{q,\mu}(f_1) \geq \sum_{m \in \mathbb{Z}^k} f_1(m)^2 - \frac{\varepsilon}{4} \sum_{m \in \mathbb{Z}^k} f_1(m). \tag{65}$$

Since ψ_{q,L_1} is L^1-normalized it follows that

$$\sum_{m \in \mathbb{Z}^k} f_1(m) = \sum_{m,\ell \in \mathbb{Z}^k} f(m - \ell)\psi_{q,L_1}(\ell) = \sum_{m \in \mathbb{Z}^k} f(m) = \delta M^k. \tag{66}$$

Using Cauchy-Schwarz, one obtains

$$\sum_{m \in \mathbb{Z}^k} f_1(m)^2 \geq \sum_{-\sigma M \leq m_j \leq M + \sigma M} f_1(m)^2 \geq \frac{1}{(1+2\sigma)^k M^k} \left(\sum_{-\sigma M \leq m_j \leq M + \sigma M} f_1(m)\right)^2 \tag{67}$$

Since f is supported on $[1, M]^k$ (and ψ_{q,L_1} is L^1-normalized) it follows that

$$\sum_{-\sigma M \leq m_j \leq M + \sigma M} f_1(m) \geq \sum_{m \in \mathbb{Z}^k} f(m) \left(1 - \sum_{|\ell_j| \geq \sigma M} \psi_{q,L_1}(\ell)\right) \geq \delta M^k (1 - \sigma) \tag{68}$$

as φ can be chosen so that $\sum_{|\ell_j| \geq \sigma M} \psi_{q,L_1}(\ell) \leq \sigma$ whenever $M \gg L_1$.

Note that

$$\frac{(1-\sigma)^2}{(1+2\sigma)^k} \geq (1-\sigma)^2(1 - 2\sigma k) \geq 1 - 4k\sigma$$

provided $2\sigma k < 1$. Hence taking $\sigma = \varepsilon/16k$ completes the proof. □

5.4. **Proof of Lemma 5.** It is in establishing Lemma 5 that we finally exploit the arithmetic properties of the curve $\gamma(n)$. In particular, we will make use of the following "minor arc" estimates for the exponential sums $S_{\lambda,\mu,q}$.

Lemma 6 (Corollary of Lemma 1)**.** *Let $\eta_0 > 0$. If $0 < \eta' < \eta < 1$ with $\eta' = \eta_0\eta$, then*

$$\|(1 - \widehat{\psi}_{q',L_2'})S_{\lambda,\mu,q}\|_\infty \leq 2C_1\eta_0 \tag{69}$$

where $q' = q_{\eta'}$ and $L_2' = \eta'^k\mu$.

Proof. Let $\alpha \in \mathbb{T}^k$ be fixed. If there exists $a \in \mathbb{Z}^k$ such that

$$\left|(T_\lambda\alpha)_i - \frac{a}{q'^i}\right| \leq \frac{\eta_0}{(\eta'^k\mu)^i},$$

for all $1 \leq i \leq k$, then (as in (52) above) we note that φ can be chosen such that

$$|1 - \widehat{\psi}_{q',L_2'}(\alpha)| \leq \eta_0. \tag{70}$$

While if for some $1 \leq i \leq k$ we have

$$\left|(T_\lambda\alpha)_i - \frac{a}{q'^i}\right| > \frac{\eta_0}{(\eta'^k\mu)^i}$$

for all $a \in \mathbb{Z}^k$, then

$$\left|q^i(T_\lambda\alpha)_i - \frac{a}{q_0^i}\right| > \frac{q^i}{(\eta_0^k\mu)^i}$$

for all $a \in \mathbb{Z}^k$, since $qq_0|q'$ where $q_0 = q_{\eta_0}$.

Since

$$q \circ (T_\lambda \alpha) = T_{\lambda/q}(q \circ \alpha)$$

it follows from (47) that

$$S_{\lambda,\mu,q}(\alpha) = \frac{1}{\mu'} \sum_{s \in (\lambda' + q^{-1}, \lambda' + \mu'] \cap \mathbb{Z}} e^{2\pi i \gamma(s) \cdot (q \circ \alpha)}$$

where $\lambda' = \lambda/q$ and $\mu' = \mu/q$ and the remark proceeding Lemma 1 that

$$|S_{\lambda,\mu,q}(\alpha)| \leq 2C_1\eta_0. \tag{71}$$

□

Proof of Lemma 5. We first construct the number $\eta > 0$. Choosing a lacunary sequence $\{\eta_j\}$ for which

$$\eta_1 \ll \varepsilon \quad \text{and} \quad \eta_{j+1} \leq (\varepsilon/80C_1)\eta_j \tag{72}$$

for each $j \geq 1$ it is easy to see that

$$\sup_{\alpha \in \mathbb{T}^k} \sum_{j=1}^{\infty} \big| \widehat{\psi}_{q_{j+1}, \eta_{j+1}^k \mu}(\alpha) - \widehat{\psi}_{q_j, \eta_j^k \mu}(\alpha) \big| \leq C_2$$

where $q_j = q_{\eta_j}$, as the terms have essentially disjoint supports. It follows that there exist $1 \leq j \leq 40C_2/\varepsilon$ such that

$$\int_{\mathbb{T}^k} |\widehat{f}(\alpha)|^2 |\widehat{\psi}_{q_{j+1}, \eta_{j+1}^k \mu}(\alpha) - \widehat{\psi}_{q_j, \eta_j^k \mu}(\alpha)| \, d\alpha \leq (\varepsilon/40) M^k. \tag{73}$$

We set $\eta = \eta_j$ and $\eta' = \eta_{j+1}$ for this value of j and note that η satisfies the inequality

$$\exp(-C\varepsilon^{-1} \log \varepsilon^{-1}) \leq \eta \ll \varepsilon.$$

Lemma 5 now follows immediately from Lemma 6 and (73), since

$$|(1 - \widehat{\psi}_{q,L_2})(\alpha) S_{\lambda,\mu,q}(\alpha)| \leq |(1 - \widehat{\psi}_{q',L_2'})(\alpha) S_{\lambda,\mu,q}(\alpha)| + |(\widehat{\psi}_{q,L_2}(\alpha) - \widehat{\psi}_{q',L_2'}(\alpha)) S_{\lambda,\mu,q}(\alpha)| \tag{74}$$

and $\eta'/\eta \leq \varepsilon/80C_1$.

Indeed, arguing as in the proof of Proposition 3 above, we obtain

$$\begin{aligned} \max\{|\Lambda_{q,\mu}(f_2, f_1)|, |\Lambda_{q,\mu}(f, f_2)|\} &\leq \int_{\mathbb{T}^k} |\widehat{f}(\alpha)|^2 \, |(1 - \widehat{\psi}_{q,L_2}(\alpha)) \, S_{\lambda,\mu,q}(\alpha)| \, d\alpha \\ &\leq \|(1 - \widehat{\psi}_{q',L_2}) S_{\lambda,\mu,q}\|_\infty M^k + (\varepsilon/40) M^k \end{aligned}$$

where the last inequality follows from Plancherel and the fact that $\|f\|_2^2 \leq \|f\|_1 \leq M^k$. □

6. The proofs of Theorems 1 and 2

In both of the proofs below we fix a polynomial $P(n)$ with integer coefficients, namely

$$P(n) = c_1 n + \cdots + c_k n^k$$

and let $\mathcal{P} : \mathbb{Z}^k \to \mathbb{Z}$ denote the mapping given by

$$\mathcal{P}(b) = c_1 b_1 + \cdots + c_k b_k.$$

Furthermore, given any set $A \subseteq \mathbb{Z}$ we define

$$A_j = \{a \in A \, : \, a \equiv j \mod m\} = A \cap (m\mathbb{Z} + j)$$

for each $0 \leq j \leq m - 1$ where

$$m = \gcd(c_1, \ldots, c_k).$$

6.1. **Deduction of Thereom 1 from Theorem 3.** Let $\varepsilon > 0$ and $A \subseteq [1, N]$ with $\delta = |A|/N$ satisfying $0 < \varepsilon \leq \delta^2$. We suppose that

$$|A \cap (A + P(n))| \leq (\delta^2 - \varepsilon)N$$

for some $n \in \mathbb{N}$. Without loss in generality we will make the convenient additional assumption that $m|N$.

It is easy to see that there necessarily exists $0 \leq j \leq m-1$ such that

$$|A_j \cap (A_j + P(n))| \leq (\delta_j^2 - \varepsilon)N/m$$

with $\delta_j = m|A_j|/N$. If we now let

$$B' = \{b \in \mathbb{Z}^k \,:\, \mathcal{P}(b) \in A_j - j\} \cap Q$$

where

$$Q = \mathcal{P}^{-1}(m\mathbb{Z} \cap [1, N]) \cap [-N', N']^k$$

and N' is some suitably large multiple of N (depending only on the coefficients of P) then it follows that

$$\delta_j = |B'|/|Q|$$

and

$$\frac{|B' \cap (B' + \gamma(n))|}{|Q|} = \frac{|A_j \cap (A_j + P(n))|}{N/m}.$$

We now set $M = \eta N/m$ for some suitably small $\eta > 0$,

$$X = \{x \in (M\mathbb{Z})^k \,:\, x + [1, M]^k \subseteq Q\}$$

and

$$Q' = \bigcup_{x \in X} (x + [1, M]^k),$$

noting that we can clearly choose $\eta \ll \varepsilon$ to ensure that

$$\frac{|Q \setminus Q'|}{|Q|} \leq \frac{\varepsilon}{10} \quad \Longleftrightarrow \quad \frac{|Q|}{|Q'|} \leq 1 + \frac{\varepsilon}{9}.$$

Thus, if we set $B'' = B' \cap Q'$ and $\beta = |B''|/|Q'|$, it follows that

$$\beta \geq \delta_j - \varepsilon/10$$

and

$$\frac{|B'' \cap (B'' + \gamma(n))|}{|Q'|} \leq \frac{|B' \cap (B' + \gamma(n))|}{|Q|} \frac{|Q|}{|Q'|} \leq \beta^2 - \varepsilon/2.$$

It follows that there must exist $x \in X$ such that if we set

$$B = B'' \cap (x + [1, M]^k)$$

then

$$\frac{|B \cap (B + \gamma(n))|}{M^k} \leq \left(\frac{|B|}{M^k}\right)^2 - \varepsilon/2.$$

In summary we have shown that for any given set $A \subseteq [1, N]$ and $\varepsilon > 0$ there exists a set $B \subseteq [1, M]^k$ with $M \ll \varepsilon N/m$ such that

$$\left\{n \in \mathbb{N} \,:\, \frac{|B \cap (B + \gamma(n))|}{M^k} > \left(\frac{|B|}{M^k}\right)^2 - \varepsilon/2\right\} \subseteq \left\{n \in \mathbb{N} \,:\, \frac{|A \cap (A + P(n))|}{N} > \left(\frac{|A|}{N}\right)^2 - \varepsilon\right\}$$

and hence Theorem 1 now follows from Theorem 3. □

6.2. **Deduction of Theorem 2 from Theorem 4.** Let $\varepsilon > 0$ and $A \subseteq \mathbb{N}$ with $\delta = \delta(A)$ satisfying $0 < \varepsilon \leq \delta^2$. We suppose that

$$\delta(A \cap (A + P(n))) \leq \delta^2 - \varepsilon$$

for some $n \in \mathbb{N}$. It follows from the definition of upper density that there exists a sequence of intervals $\{I_i\}$ with $|I_i| = N_i$, where $\{N_i\} \subseteq m\mathbb{N}$ and $N_i \nearrow \infty$, such that

$$\frac{|(A \cap I_i) \cap ((A \cap I_i) + P(n))|}{N_i} \leq \delta^2 - \varepsilon/2$$

while

$$\frac{|A \cap I_i|}{N_i} \geq \delta - \varepsilon/10.$$

If we define $\delta_i = |A \cap I_i|/N_i$, it therefore follows that

$$\frac{|(A \cap I_i) \cap ((A \cap I_i) + P(n))|}{N_i} \leq \delta_i^2 - \varepsilon/5.$$

Note that $(A \cap I_i)_j = A_j \cap I_i$. If we set $\delta_{ij} = m|A_j \cap I_i|/N_i$, then

$$\delta_i = \frac{1}{m} \sum_{j=0}^{m-1} \delta_{ij}$$

and as a consequence of the Cauchy-Schwarz inequality we have

$$\delta_i^2 \leq \frac{1}{m} \sum_{j=0}^{m-1} \delta_{ij}^2.$$

It therefore follows immediately from the fact that

$$\sum_{j=0}^{m-1} |(A_j \cap I_i) \cap ((A_j \cap I_i) + P(n))| = |(A \cap I_i) \cap ((A \cap I_i) + P(n))| \leq \frac{1}{m} \sum_{j=0}^{m-1} (\delta_{ij}^2 - \varepsilon/5) N_i$$

that for each i there must exist $0 \leq j \leq m-1$ for which

$$|(A_j \cap I_i) \cap ((A_j \cap I_i) + P(n))| \leq (\delta_{ij}^2 - \varepsilon/5) N_i/m.$$

We will assume (by refining our collection $\{I_i\}$ if necessary) that the same j is selected for each i. Since $(A \cap I_i)_j = A_j \cap I_i$ it follows, from the definition of upper density, that

$$\delta(A_j) \geq \delta_{ij}/m$$

and

$$\delta(A_j \cap (A_j + P(n))) \leq (\delta_{ij}^2 - \varepsilon/5)/m.$$

If we now define

$$B = \{b \in \mathbb{Z}^k \, : \, \mathcal{P}(b) \in A_j - j\}$$

it follows immediately that

$$\delta(B) = m\delta(A_j)$$

$$\delta(B \cap (B + \gamma(n))) = m\delta(A_j \cap (A_j + P(n)))$$

and consequently

$$\delta(B \cap (B + \gamma(n))) \leq \delta(B)^2 - \varepsilon/5.$$

In summary we have shown that for any given set $A \subseteq \mathbb{N}$ with $\delta(A) > 0$ and $\varepsilon > 0$ there exists a set $B \subseteq \mathbb{Z}^k$ with $\delta(B) > 0$ such that

$$\left\{n \in \mathbb{N} \, : \, \delta(B \cap (B + \gamma(n))) > \delta(B)^2 - \varepsilon/5\right\} \subseteq \left\{n \in \mathbb{N} \, : \, \delta(A \cap (A + P(n))) > \delta(A)^2 - \varepsilon\right\}$$

and hence Theorem 2 follows immediately from Theorem 4. □

7. The parameter L in Theorem 2 necessarily depends on the set A

In this final section we construct an example to show that the parameter L in Theorem 2 necessarily depends on the actual set A and not just on its density.

Proposition 4. *Let $P \in \mathbb{Z}[n]$ with $P(0) = 0$ and $L \in \mathbb{N}$, then there exist $A \subseteq \mathbb{N}$ with $\delta(A) = 1/3$ and an unbounded increasing sequence $\{\lambda_j\}$ with the property that $A \cap (A + P(n)) = \emptyset$ whenever $n \in \bigcup_{j=0}^{\infty} [\lambda_j, \lambda_j + L]$.*

Proof. With out loss in generality we assume that the leading coefficient of P. Set $M = P(aL)$ with $a \in \mathbb{N}$ chosen so that P is increasing and $2P(aL) \geq P((a+1)L)$. We definine $A \subseteq \mathbb{N}$ such that $A = A + 3M$ and $A \cap [1, 3M] = [M + 1, 2M]$.

Since $P(n) = P(m) \pmod{3M}$ whenever $n = m \pmod{3M}$, it is easy to see that if $\lambda_j = j3M + aL$, then the fact that $A \cap (A + P(n)) = \emptyset$ whenever $n \in [\lambda_j, \lambda_j + L]$ for some j, follows from the fact that this holds for $j = 0$ (as can be easily verified by the reader). □

Acknowledgements

The authors would like to thank the anonymous referee for their careful reading of the article and their subsequent useful remarks and comments.

References

[1] V. Bergelson, *The multifarious Poincar recurrence theorem*, Descriptive set theory and dynamical systems (Marseille-Luminy, 1996), 31–57, London Math. Soc. Lecture Note Ser., 277, Cambridge Univ. Press, Cambridge, 2000.
[2] V. Bergelson, *Combinatorial and Diophantine applications of ergodic theory*, Handbook of dynamical systems. Vol. 1B, 745–869, Elsevier B. V., Amsterdam, 2006.
[3] J. Bourgain, *A Szemerdi type theorem for sets of positive density in R^k*, Israel J. Math. 54 (1986), no. 3, 307–316.
[4] N. Frantzikinakis and B. Kra, *Ergodic averages for independent polynomials and applications*, J. London Math. Soc. (2) 74 (2006), no. 1, 131–142.
[5] N. Frantzikinakis and R. McCutcheon, *Ergodic Theory: Recurrence*, Encyclopedia of Complexity and System Science, Springer, (2009), Part 5, 3083-3095
[6] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. d'Analyse Math, 71 (1977), 204-256.
[7] B. Green and T. Tao, *An arithmetic regularity lemma, associated counting lemma, and applications*, An irregular mind: Szemeredi is 70, Bolyai Society Math. Studies 21 (2010)
[8] A. Y. Khintchine, *Eine Verschärfung des Poincaréscen "Wiederkehrsatzes"*, Composio Math. 1 (1934), 177-179.
[9] B. Kra, *Ergodic methods in additive combinatorics*, Additive combinatorics, 103–144, CRM Proc. Lecture Notes, 43, Amer. Math. Soc., Providence, RI, 2007.
[10] N. Lyall and Á. Magyar, *Polynomial configurations in difference sets*, J. Num. Theory, v. 129/2, pp. 439-450, 2009.
[11] N. Lyall and Á. Magyar, *Polynomial configurations in difference sets (Revised version)*, `arxiv.org/abs/0903.4504`.
[12] N. Lyall and Á. Magyar, *An optimal version of Sárközy's theorem*, `arxiv.org/abs/1010.3451`
[13] N. Lyall and Á. Magyar, *Simultaneous polynomial recurrence*, Bull. Lond. Math. Soc. 43 (2011), no. 4, 765785
[14] Á. Magyar, *On distance sets of large sets of integer points*, Israel J. Math. 164 (2008), 251–263.
[15] R. McCutcheon, *Elemental methods in ergodic Ramsey theory*, Lecture Notes in Mathematics, 1722. Springer-Verlag, Berlin, 1999.
[16] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste. I*, Gathiers-Villars, Paris 1892; II, 1893; III, 1899.

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`