POLYNOMIAL LARGENESS OF SUMSETS AND TOTALLY ERGODIC SETS

ALEXANDER FISH

ABSTRACT. We prove that a sumset of a TE subset of $\mathbb{N}$ (these sets can be viewed as "aperiodic" sets) with a set of positive upper density intersects any polynomial sequence. For WM sets (subclass of TE sets) we prove that the intersection has lower Banach density one. In addition we obtain a generalization of the latter result to the case of several polynomials.

1. INTRODUCTION

A random set in $\mathbb{N}$ satisfies the property that its sumset with a set of positive density intersects every polynomial sequence. We call a set $A \subset \mathbb{N}$ **p-good** if for every $B \subset \mathbb{N}$ of positive upper density and every $p(n) \in \mathbb{Z}[n]$ with a positive leading coefficient we have

$$(A + B) \cap \{p(n) | n \in \mathbb{N}\} \neq \emptyset.$$

A random set is p-good. The paper provides explicit constructions for p-good sets.

A p-good set cannot be periodic. One proposes a dynamical approach for constructing (aperiodic) p-good sets.

In ergodic theory there are many different notions for aperiodicity (randomness) of a measure preserving system, i.e. a quadruple $(X, \mathbb{B}_X, \mu, T)$, where X is a compact metric space, $\mathbb{B}_X$ is Borel σ-algebra on X, $T : X \to X$ is a continuous map and μ is a Borel probability measure on X which is preserved under the action of T.

We will always assume that a system $(X, \mathbb{B}_X, \mu, T)$ is **totally ergodic**[1], i.e. the systems $((X, \mathbb{B}_X, \mu, T^n))_{n \in \mathbb{N}}$ are **ergodic**. There are many equivalent definitions for ergodicity of a system. For our purposes the most convenient definition is that the conclusion of pointwise ergodic theorem holds true:
For any $f \in L^1_\mu(X)$ for almost every $x \in X$ with respect to μ we have

$$\frac{1}{N} \sum_{n=1}^{N} f(T^n x) \to \int f d\mu.$$

Let $f \in L^1_\mu(X)$. Denote by $\mathcal{A}_f$ the algebra of functions generated by f and all of its translates by T. By ergodic theorem there exists a set of full measure $X_f \subset X$

1991 *Mathematics Subject Classification.* Primary: 11B13, 37A30.
Key words and phrases. Sumsets, Total ergodicity, Weak mixing.

[1]Measure preserving systems on cyclic groups which are obviously exhibit a periodicity are not totally ergodic.

such that for every $x_0 \in X_f$, any $k \in \mathbb{N}$ and any function $g \in \mathcal{A}_f$ we have

$$\frac{1}{N}\sum_{n=1}^{N} g(T^{kn}x_0) \to \int g d\mu.$$

We will call the set X_f the set of **f-generic points**.

The space of continuous functions on X is separable, therefore by ergodic theorem there exists a set of full measure $X' \subset X$, such that for every $x \in X'$, every $f \in C(X)$ and every $k \in \mathbb{N}$ we have

$$\frac{1}{N}\sum_{n=1}^{N} f(T^{kn}x) \to \int f d\mu.$$

The set X' we call the set of **generic points** in X.

For a convenience we introduce the set $\mathbb{N}_0 = \{0, 1, , 2 \ldots\}$.

A bounded sequence $(\xi(n))_{n\in\mathbb{N}_0}$ will be called **totally ergodic** if there exists a totally ergodic system $(X, \mathbb{B}_X, \mu, T)$, a function $f \in L^\infty(X)$ and an f-generic point $x_0 \in X$ such that

$$\xi(n) = f(T^n x_0),\ \forall n \in \mathbb{N}_0.$$

We associate $\{0, 1\}$-valued sequences with subsets of $\mathbb{N}_0$ in a natural way. A set $S \subset \mathbb{N}_0$ corresponds to the sequence $1_S \in \{0, 1\}^{\mathbb{N}_0}$. We say that $S \subset \mathbb{N}_0$ is **TE set**[2] if 1_S is a totally ergodic sequence and the density of S:

$$d(S) = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} 1_S(n)$$

is positive. Notice that the density of a totally ergodic set always exists by the genericity assumption.

It was shown in [3] that any rotation by $\alpha \notin \mathbb{Q}$ on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and any interval $[a, b] \in \mathbb{T}$ generate the TE set

$$R_{\alpha,[a,b]} = \{n \in \mathbb{N}_0 \,|\, n\alpha \bmod 1 \in [a, b]\}.$$

In other words, $R_{\alpha,[a,b]}$ is the set of return times for a uniquely ergodic rotation on the compact abelian group $\mathbb{T}$ into the interval $[a, b]$. Similarly, for any homomorphism τ from $\mathbb{Z}$ to a compact abelian metrizable connected group K with $\overline{\tau(\mathbb{Z})} = K$ and any Jordan measurable set $J \subset K$ of positive Haar measure (Jordan measurability means that the boundary of J has zero Haar measure) the set

$$R_J = \tau^{-1}(J) \cap \mathbb{N}_0$$

is a TE set.

In the paper we use different notions of density for subsets of $\mathbb{N}$. For $S \subset \mathbb{N}$, the **upper density** $\overline{d}(S)$ of S is defined by

$$\overline{d}(S) = \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} 1_S(n).$$

[2]If $S \subset \mathbb{N}$ then we regard 1_S as a sequence in $\{0, 1\}^{\mathbb{N}_0}$.

The **lower density** $\underline{d}(S)$ of S is defined by

$$\underline{d}(S) = \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} 1_S(n).$$

We say $S \subset \mathbb{N}$ has density and denote it by $d(S)$ if $\overline{d}(S) = \underline{d}(S)$.
The **upper Banach density** $d^*(S)$ of S is defined by

$$d^*(S) = \limsup_{M-N\to\infty} \frac{1}{M-N} \sum_{n=N}^{M-1} 1_S(n).$$

The **lower Banach density** $d_*(S)$ of S is defined by

$$d_*(S) = \liminf_{M-N\to\infty} \frac{1}{M-N} \sum_{n=N}^{M-1} 1_S(n).$$

Note that the positivity of the lower Banach density of a set is equivalent to having bounded gaps.

The main result of the paper is that any TE set is p-good.

Theorem 1. *Let $A \subset \mathbb{N}$ be a TE set. Then for any $B \subset \mathbb{N}$ of positive upper density and any non-constant polynomial $p(n) \in \mathbb{Z}[n]$ with a positive leading coefficient we have $(A+B) \cap \{p(n) \mid n \in \mathbb{N}\} \neq \emptyset$. Moreover, if the lower density of B is positive then the set $R_p = \{n \in \mathbb{N} \mid p(n) \in A+B\}$ has bounded gaps.*

If the system $(X, \mathbb{B}_X, \mu, T)$ which was involved in the definition of a TE set is **weak-mixing**, i.e. the system $(X \times X, \mathbb{B}_X \times \mathbb{B}_X, \mu \times \mu, T \times T)$ is ergodic, then one can prove stronger results.

We introduce the notion of a WM set. A sequence $(\xi(n))_{n\in\mathbb{N}_0}$ is **weakly mixing** if there exists a weak-mixing system $(X, \mathbb{B}_X, \mu, T)$, a function $f \in L^\infty_\mu(X)$ and an f-generic point $x_0 \in X$ such that

$$\xi(n) = f(T^n x_0), \quad ,\forall n \in \mathbb{N}_0.$$

Similarly to the definition of a TE set, a set $S \subset \mathbb{N}$ is a **WM set** if 1_S is a weakly mixing sequence and the density of S is positive.
A weak-mixing system is totally ergodic, thus any WM set is a TE set.

We propose here a simple dynamical construction of WM sets. Take the shift space (Ω, σ), where $\Omega = \{0,1\}^{\mathbb{N}_0}$ is endowed with Tychonoff topology and σ is the shift to left. Take any Borel probability measures μ on Ω which preserved under the shift σ and which generate a weak-mixing system $(\Omega, \mathbb{B}_\Omega, \mu, T)$.Take any cylinder set $A \subset \Omega$ with $\mu(A) > 0$. Notice that any cylinder is a clopen set, i.e. the indicator function of A, $\chi_A \in C(\Omega)$. Then any generic point $\omega \in \Omega$ generates a WM set

$$S_{\omega,A} = \{n \in \mathbb{N} \mid \sigma^n \omega \in A\}$$

with $d_{S_{\omega,A}} = \mu(A)$.

If A in Theorem 1 is a WM set, then we can prove that the set R_p is of lower Banach density 1.

Theorem 2. *Let $A \subset \mathbb{N}$ be a WM set, let $B \subset \mathbb{N}$ of positive upper density and let $p(n) \in \mathbb{Z}[n]$ with a positive leading coefficient. Then the set $R_p = \{n \in \mathbb{N} \mid p(n) \in A+B\}$ is of lower Banach density 1.*

We use the notion of essentially distinct polynomials introduced by Bergelson in [1].

The polynomials $\{p_1, \dots, p_n\}$ are called **essentially distinct** if for every $1 \leq i < j \leq n$ we have $p_i - p_j$ is a non-constant polynomial.

All polynomials $p(n)$ that we consider are with integer coefficients and satisfy $p(n) \to \infty$ as $n \to \infty$. The following theorem is a generalization of Theorem 2.

Theorem 3. *Let $A \subset \mathbb{N}$ be a WM set, let $p_1(n), \dots, p_k(n) \in Z[n]$ be essentially distinct polynomials of the same degree, let $B \subset \mathbb{N}$ of positive upper density. Then the set*

$$R_{p_1,\dots,p_k} = \{n \in \mathbb{N} \,|\, \exists b \in B : \; p_1(n), p_2(n), \dots, p_k(n) \in A + b\}$$

has lower Banach density 1.

Notice that any element $n \in R_{p_1,\dots,p_k}$ corresponds to a solution of the equation:

$$\left\{ \begin{array}{l} x + y_1 = p_1(n) \\ x + y_2 = p_2(n) \\ \dots \\ x + y_k = p_k(n) \end{array} \right. \tag{1.1}$$

where $x \in B, y_1, \dots, y_k \in A$.

If among $p_1(n), \dots, p_k(n)$ there are two polynomials with degrees which differ by at least two, then there exists a WM set A such that the set

$$R_{p_1,\dots,p_k} = \{n \in \mathbb{N} \,|\; p_1(n), p_2(n), \dots, p_k(n) \in A + A\}$$

is empty. To prove the last claim we take an arbitrary WM set A. Notice that by the definition of a WM set[3] for any set of density zero $N \subset \mathbb{N}$ the set $A \setminus N$ is again a WM set. In particular, we can exclude from A all solutions of the system (1.1) by removing a set of density zero. If $\deg p_1 \leq \deg p_2 - 2$ then replace A by

$$A' = A \setminus \left(\bigcup_{n \in \mathbb{N}} [p_2(n) - p_1(n), p_2(n)] \right)$$

which is again a WM set. Within A' the system (1.1) is unsolvable.

Acknowledgment: I would like to thank Hillel Furstenberg who proposed me the problem. Without his encouragement this work would never be done. Also I would like to thank Vitaly Bergelson for fruitful discussions.

2. Orthogonality of polynomial shifts along totally ergodic sequences

Throughout the paper we use the notation $L^2(N)$ to denote the space of real-valued functions on the finite set $\{1, 2. \dots, N\}$ endowed with the scalar product:

$$\langle u, v \rangle_N = \frac{1}{N} \sum_{n=1}^{N} u(n)v(n).$$

The main tool for the proof of Theorem 1 is the almost orthogonality of polynomial shifts along a totally ergodic sequence.

[3]The same is true for a TE set.

Proposition 1. *Let $(\xi(n))_{n\in\mathbb{N}_0}$ be a totally ergodic sequence of zero mean. Let $p(n) \in \mathbb{Z}[n]$ be a non-constant polynomial with a positive leading coefficient. For every $\varepsilon > 0$ there exists $J(\varepsilon)$ such that for every $J \geq J(\varepsilon)$ there exists $N(J)$ such that for every $N \geq N(J)$ we have*[4]

$$\left\| \frac{1}{J} \sum_{j=1}^{J} \xi(p(N+j) - n) \right\|_{p(N)} < \varepsilon.$$

First we will establish an auxiliary statement which is also an almost orthogonality of other polynomial shifts. For a non-constant polynomial $q[n] \in \mathbb{Z}[n]$ with a positive leading coefficient which has a smaller degree than $p(n)$ and any $j \in \mathbb{N}$ we define the vector $v_j^q \in L^2(p(N))$ by

$$v_j^q(n) = \xi(n + q(N+j)),\ 1 \leq n \leq p(N).$$

Lemma 1. *Let $\varepsilon > 0$. There exists $J(\varepsilon)$ such that for every $J \geq J(\varepsilon)$ there exists $N(J)$ such that for every $N \geq N(J)$ we have*

$$\left\| \frac{1}{J} \sum_{j=1}^{J} v_j^q \right\|_{p(N)} < \varepsilon.$$

Proof. The proof is by induction on $\deg q(n)$.

Case $\deg q(n) = 1$**:** Assume $q(x) = ax + b$. Then

$$\left\| \frac{1}{J} \sum_{j=1}^{J} v_j^q \right\|_{p(N)}^2 = \frac{1}{p(N)} \sum_{n=1}^{p(N)} \left(\frac{1}{J} \sum_{j=1}^{J} \xi(n + aj) \right)^2 + \delta_{N,J},$$

where $\delta_{N,J} \to 0$ as $N \to \infty$. By total ergodicity of the sequence $(\xi(n))_{n\in\mathbb{N}}$ there exists a totally ergodic system $(X, \mathbb{B}_X, \mu, T)$, a function $f \in L^\infty(X)$ and an f-generic point $x_0 \in X$ such that

$$\xi(n) = f(T^n x_0),\ \text{ for all } n \in \mathbb{N}_0.$$

Therefore

$$\frac{1}{p(N)} \sum_{n=1}^{p(N)} \left(\frac{1}{J} \sum_{j=1}^{J} \xi(n + aj) \right)^2 = \frac{1}{p(N)} \sum_{n=1}^{p(N)} T^n \left(\frac{1}{J} \sum_{j=1}^{J} T^{aj} f \right)^2 (x_0) \tag{2.1}$$

The function $g_J(x) = \left(\frac{1}{J}\sum_{j=1}^{J} T^{aj} f(x)\right)^2$ is in $\mathcal{A}_f$, therefore by f-genericity of the point x_0 we get

$$\frac{1}{p(N)} \sum_{n=1}^{p(N)} T^n g_J(x_0) \to \int g_J d\mu,\ \text{ as } N \to \infty.$$

[4]In the case when a sequence depends on many parameters, like in this case j, N, n the L^2-norm is taken with respect to n.

We claim that g_J converges in $L^1(X)$ to zero as $J \to \infty$. By f-genericity of x_0 we have

$$\frac{1}{J}\sum_{j=1}^{J} T^{aj} f(x_0) \to \int f d\mu, \text{ as } J \to \infty.$$

By f-genericity of x_0 the function f has zero integral.

$$\int f d\mu = \lim_{J\to\infty} \frac{1}{J}\sum_{j=1}^{J} f(T^j x_0) = \lim_{J\to\infty} \frac{1}{J}\sum_{j=1}^{J} \xi(j) = 0.$$

By L^2-ergodic theorem we have

$$\frac{1}{J}\sum_{j=1}^{J} T^{aj} f(x) \to \int f d\mu, \text{ as } J \to \infty$$

where the convergence is in $L^2(X)$. By Hölder inequality the latter implies that

$$\int g_J d\mu \to 0, \text{ as } J \to \infty.$$

By equation (2.1) the latter implies the statement of the lemma. [5]

Case $\deg q(x) = n > 1$**:**
Let $\varepsilon > 0$. The vectors v_j^q are uniformly bounded by $\|\xi\|_\infty$. Without loss of generality assume that $\|\xi\|_\infty \leq 1$. Let $I = I(\varepsilon)$ be as in the finitary version of van der Corput lemma (Lemma 3 in the appendix). It is enough to show that there exists $J(I)$ such that for every $J \geq J(I)$ there exists $N(J)$ such that for every $N \geq N(J)$ and every $i : 1 \leq i \leq I$ we have

$$\left| \frac{1}{J}\sum_{j=1}^{J} \langle v_j^q, v_{j+i}^q \rangle_{p(N)} \right| < \frac{\varepsilon}{2}. \tag{2.2}$$

An easy calculation shows that

$$\frac{1}{J}\sum_{j=1}^{J} \langle v_j^q, v_{j+i}^q \rangle_{p(N)} = \frac{1}{J}\sum_{j=1}^{J} \frac{1}{p(N)} \sum_{n=1}^{p(N)} \xi(n + q(N+j))\xi(n + q(N+j+i))$$

$$= \frac{1}{p(N)} \sum_{n=1}^{p(N)} \xi(n) \frac{1}{J}\sum_{j=1}^{J} \xi(n + q(N+j+i) - q(N+j)) + \delta_{N,J,i},$$

where $\delta_{N,J,i} \to 0$ as $N \to \infty$.
Denote by $w_{i,j}^q(n) = \xi(n + q(N+j+i) - q(N+j))$, $r(x) = q(x+i) - q(x)$. Note that $\deg r(x) = \deg q(x) - 1$ and $r(x) \to \infty$ as $x \to \infty$. By induction's hypothesis

[5] Notice that we also proved that for any non-constant polynomial $p(n) \in \mathbb{Z}[n]$ with a positive leading coefficient and any $a \in \mathbb{N}$, for every $\varepsilon > 0$ there exists $J(\varepsilon)$ such that for any $J \geq J(\varepsilon)$ there exists $N(J)$ such that for every $N \geq N(J)$ we have

$$\frac{1}{p(N)} \sum_{n=1}^{p(N)} \left(\frac{1}{J}\sum_{j=1}^{J} \xi(n + aj) \right)^2 < \varepsilon.$$

there exists $J(i)$ such that for every $J \geq J(i)$ there exists $N(J,i)$ such that for every $N \geq N(J,i)$ we have

$$\left\| \frac{1}{J} \sum_{j=1}^{J} w_{i,j}^{q} \right\|_{p(N)} < \frac{\varepsilon}{4}$$

The latter implies that there exists $J(I)$ such that for every $J \geq J(I)$ there exists $N(J)$ such that for every $N \geq N(J)$ we have for every $i \in \{1,2,\ldots,I\}$ the following

$$\left\| \frac{1}{J} \sum_{j=1}^{J} w_{i,j}^{q} \right\|_{p(N)} < \frac{\varepsilon}{4}.$$

Cauchy-Schwartz inequality implies

$$\left| \frac{1}{J} \sum_{j=1}^{J} \langle v_j^q, v_{j+i}^q \rangle_{p(N)} \right| \leq \|\xi\|_{p(N)} \left\| \frac{1}{J} \sum_{j=1}^{J} w_{i,j}^{q} \right\|_{p(N)} + |\delta_{N,J,i}| = \frac{\varepsilon}{4} + |\delta_{N,J,i}|$$

for any $i \in \{1,2,,\ldots,I\}, J \geq J(I)$ and every $N \geq N(J)$. Taking into account that $\delta_{N,J,i} \to 0$ as $N \to \infty$ implies that the inequality (2.2) is fulfilled for all $i \in \{1,2,\ldots,I\}$, any $J \geq J(I)$ and any $N \geq N(J)$. □

Proof of Proposition 1. Denote by $u_j(n) = \xi(p(N+j)-n)$.
Case $\deg p(x) = 1$**:** Assume $p(x) = ax+b$. Then

$$\left\| \frac{1}{J} \sum_{j=1}^{J} u_j \right\|_{p(N)}^2 = \frac{1}{aN+b} \sum_{n=1}^{aN+b} \left(\frac{1}{J} \sum_{j=1}^{J} \xi(n+aj) \right)^2 + \delta_{N,J},$$

where $\delta_{N,J} \to 0$ as $N \to \infty$. By the remark in the proof of Lemma 1 the case $\deg p(x) = 1$ follows immediately.

Case $\deg p(x) > 1$**:** We use van der Corput lemma (Lemma 3). Without loss of generality we assume that $\|\xi\|_\infty \leq 1$. Let $\varepsilon > 0$. Let $I = I(\varepsilon)$ be as in van der Corput lemma. One knows

$$\frac{1}{J} \sum_{j=1}^{J} \langle u_j, u_{j+i} \rangle_{p(N)} = \langle \xi(n), \frac{1}{J} \sum_{j=1}^{J} \xi(n + p(N+j+i) - p(N+j)) \rangle_{p(N)} + \delta_{N,J,i},$$

where $\delta_{N,J,i} \to 0$ as $N \to \infty$. By Lemma 1 there exists $J(i)$ such that for any $J \geq J(i)$ there exists $N(J,i)$ such that for every $N \geq N(J,i)$ we have

$$\left| \langle \xi(n), \frac{1}{J} \sum_{j=1}^{J} \xi(n + q(N+j)) \rangle_{p(N)} \right| \leq \frac{\varepsilon}{2}.$$

The latter implies that there exists $J(I)$ such that for any $J \geq J(I)$ there exists $N(J)$ such that for every $N \geq N(J)$ and every $i \in \{1, 2, \dots, I\}$ we have

$$\left| \frac{1}{J} \sum_{j=1}^{J} \langle u_j, u_{j+i} \rangle_{p(N)} \right| < \varepsilon.$$

Van der Corput lemma implies the statement of the Proposition. □

3. Orthogonality of polynomial shifts along weakly mixing sequences

We start with a statement which is analogous to Proposition 1. The only difference that we assume that the sequence $(\xi(n))$ is weakly mixing rather than totally ergodic. As a consequence we get a stronger conclusion than in Proposition 1.

Proposition 2. *Let $(\xi(n))_{n \in \mathbb{N}_0}$ be a weakly mixing sequence of zero mean, $p_1, \dots, p_k \in \mathbb{Z}[n]$ be essentially distinct polynomials of the same degree $d \geq 1$, with positive leading coefficients such that $p_1(n) - p_i(n) \to \infty, \forall 1 < i \leq k$ as $n \to \infty$. For every $\varepsilon > 0$ there exists $J(\varepsilon)$ such that for any $J \geq J(\varepsilon)$ there exists $N(J)$ such that for every $N \geq N(J)$ and any $\{0,1\}$-valued sequence $(a_n)_{n \in \mathbb{N}}$ we have*[6]

$$\left\| \frac{1}{J} \sum_{j=1}^{J} a_{N+j} \xi(p_1(N+j) - n) \xi(p_2(N+j) - n) \dots \xi(p_k(N+j) - n) \right\|_{p_1(N)} < \varepsilon.$$

For the proof of Proposition 2 we will prove first the following claim.

Lemma 2. *Let $(\xi(n))_{n \in \mathbb{N}_0}$ be a weakly mixing sequence of zero mean, $p_1, \dots, p_k \in \mathbb{Z}[n]$ be essentially distinct polynomials with positive leading coefficients, and $q(n) \in \mathbb{Z}[n]$ be such that for every $i : 1 \leq i \leq k$ we have $\frac{q(n)}{p_i(n)} \to k_i \in (1, +\infty]$ as $n \to \infty$*[7]*. For every $\varepsilon > 0$ there exists $J(\varepsilon)$ such that for every $J \geq J(\varepsilon)$ there exists $N(J)$ such that for every $N \geq N(J)$ and any $\{0,1\}$-valued sequence $(a_n)_{n \in \mathbb{N}}$ we have*

$$\left\| \frac{1}{J} \sum_{j=1}^{J} a_{N+j} \xi(n - p_1(N+j)) \xi(n - p_2(N+j)) \dots \xi(n - p_k(N+j)) \right\|_{q(N)} < \varepsilon.$$

Proof. We prove this statement by using an analog of Bergelson's PET induction, see [1]. Let $F = \{p_1, \dots, p_k\}$ be a finite set of polynomials and assume that the largest of the degrees of p_i equals d. For every $i : 1 \leq i \leq d$ we denote by n_i the number of different groups of polynomials of degree i, where two polynomials p_{j_1}, p_{j_2} of degree i are in the same group if and only if they have the same leading coefficient. We will say that $(n_1, \dots, n_d)$ is the **characteristic vector** of F.
We prove a more general statement than the statement of the lemma.
Let $\mathcal{F}(n_1, \dots, n_d)$ be the family of all finite sets of essentially distinct polynomials having characteristic vector $(n_1, \dots, n_d)$. Consider the following two statements:

[6] We assume that for all negative n we have $\xi(n) = 0$.
[7] We will say that the polynomial q **grows faster to infinity** than the family $\{p_1, \dots, p_k\}$.

$L(k;n_1,\dots,n_d)$: *For every* $\{g_1,\dots,g_{n_1},q_1,\dots,q_l\}\in\mathcal{F}(n_1,\dots,n_d)$, *where* $g_1,\dots,g_{n_1}$ *are linear polynomials,* $q(n)\in\mathbb{Z}[n]$ *which grows faster to infinity than the family* $\{g_1,\dots,g_{n_1},q_1,\dots,q_l\}$, *every* $(\boldsymbol{c}_i)_{i=1}^{n_1}\in(\mathbb{Z}\setminus\{0\})^k$ *and every* $\varepsilon,\delta>0$ *there exists* $H(\delta,\varepsilon,(\boldsymbol{c}_i))\in\mathbb{N}$ *such that for every* $H\ge H(\delta,\varepsilon,(\boldsymbol{c}_i))$ *there exists* $J(H)$ *such that for every* $J\ge J(H)$ *there exists* $N(J)$ *such that for every* $N\ge N(J)$ *for a set of* $\boldsymbol{h}\in\{1,2,\dots,H\}^k$ *of density at least* $1-\delta$ *for every* $\{0,1\}$*-valued sequence* $(a_n)_{n\in\mathbb{N}}$ *we have*[8]

$$\left\|\frac{1}{J}\sum_{j=1}^{J}a_{N+j}\prod_{i=1}^{n_1}\prod_{\epsilon\in\{0,1\}^k}\xi(n-g_i(N+j)-(\boldsymbol{c}_i\epsilon)\cdot\boldsymbol{h})\prod_{i=1}^{l}\xi(n-q_i(N+j))\right\|_{q(N)}<\varepsilon,$$

where $\boldsymbol{c}_i\epsilon=(c_1^i\epsilon_1,\dots,c_k^i\epsilon_k)$ *for* $\boldsymbol{c}_i=(c_1^i,\dots,c_k^i)$, $\epsilon=(\epsilon_1,\dots,\epsilon_k)$.

$L(k;\overline{n_1,\dots,n_i},n_{i+1},\dots,n_d)$: $L(k;n_1,\dots,n_d)$ *is valid for any* $n_1,\dots,n_i$.

Lemma 2 is the statement $L(0;\overline{n_1,\dots,n_d})$. In order to prove the latter it is enough to establish $L(k;1)$, $\forall k\in\mathbb{N}_0$, and to prove the following implications:

$$S.1_d:\ L(k+1;n_1,n_2,\dots,n_d)\Rightarrow L(k;n_1+1,n_2,\dots,n_d);$$
$$k,n_1,\dots,n_{d-1}\ge 0,n_d\ge 1,d\ge 1$$
$$S.2_{d,i}:\ L(0;\overline{n_1,\dots,n_{i-1}},n_i,\dots,n_d)\Rightarrow L(k;\underbrace{0,\dots,0}_{i-1\,zeros},n_i+1,n_{i+1},\dots,n_d);$$
$$k,n_1,\dots,n_{d-1}\ge 0,n_d\ge 1,d\ge i>1$$
$$S.3_d:\ L(k;\overline{n_1,\dots,n_d})\Rightarrow L(k;\underbrace{0,\dots,0}_{d\,zeros},1),\ \ k\ge 0\,,\,d\ge 1$$

Proof of $S.2_{d,i}$**.** Suppose that F is a finite set of essentially distinct polynomials and assume that the characteristic vector of F equals $(\underbrace{0,\dots,0}_{i-1 zeros},n_i+1,n_{i+1},\dots,n_d)$.

Fix any of the n_i+1 groups of polynomials of degree i and denote its polynomials by $g_1,\dots,g_m$. Denote the remaining polynomials in F by $q_1,\dots,q_l$. Notice that there are no linear polynomials among the polynomials of F. Let $\boldsymbol{c}_1\in(\mathbb{Z}\setminus\{0\})^k$. To establish $L(k;\underbrace{0,\dots,0}_{i-1\,zeros},n_i+1,n_{i+1},\dots,n_d)$ we have to prove that for every $\varepsilon,\delta>0$ there exists $H(\varepsilon,\delta,\boldsymbol{c}_1)$ such that for every $H\ge H(\varepsilon,\delta,\boldsymbol{c}_1)$ there exists $J(H)$ such that for any $J\ge J(H)$ there exists $N(J)$ such that for every $N\ge N(J)$ for a set of $\boldsymbol{h}\in\{1,\dots,H\}^k$ of density which is at least $1-\delta$ and for any $\{0,1\}$-valued sequence $(a_n)_{n\in\mathbb{N}}$ we have

$$\left\|\frac{1}{J}\sum_{j=1}^{J}a_{N+j}\prod_{\epsilon\in\{0,1\}^k}\xi(n-(\boldsymbol{c}_1\epsilon)\cdot\boldsymbol{h})\prod_{c=1}^{m}\xi(n-g_c(N+j))\prod_{e=1}^{l}\xi(n-q_e(N+j))\right\|_{q(N)}<\varepsilon.$$

Let $(a_n)_{n\in\mathbb{N}}$ be a $\{0,1\}$-valued sequence and $\boldsymbol{h}\in\{1,2,\dots,H\}^k$. Denote by

$$u_j(n)=a_{N+j}\prod_{c=1}^{m}\xi(n-g_c(N+j))\prod_{e=1}^{l}\xi(n-q_e(N+j)),$$

[8]In the case $n_1=0$ and $k>0$ we require that the similar inequality holds true for $\boldsymbol{c}_1\in(\mathbb{Z}\setminus\{0\})^k$.

$$w(n) = \prod_{\epsilon\in\{0,1\}^k} \xi(n - (\boldsymbol{c}_1\epsilon)\cdot \boldsymbol{h}),$$
$$v_j(n) = w(n)u_j(n).$$

Let $\varepsilon > 0$. Without loss of generality we can assume that $\|\xi\|_\infty \leq 1$. This implies that $\|w\|_\infty \leq 1$ and therefore to prove that $\|\frac{1}{J}\sum_{j=1}^J v_j\|_{q(N)} < \varepsilon$ it is sufficient to show that $\|\frac{1}{J}\sum_{j=1}^J u_j\|_{q(N)} < \varepsilon$.
Let $h \geq 1$. A simple routine calculation gives that

$$\frac{1}{J}\sum_{j=1}^{J}\langle u_j, u_{j+h}\rangle_{q(N)} = \langle \xi(n), \frac{1}{J}\sum_{j=1}^{J} b_{N+j}\prod_{c=1}^{2m+2l-1}\xi(n - r_c(N+j))\rangle_{q(N)} + \delta_{N,J},$$

where $b_{N+j} = a_{N+j}a_{N+j+h}$, $\delta_{N,J} \to 0$ as $N \to \infty$ and

$$\begin{cases} r_t(n) = g_{t+1}(n) - g_1(n)\,,\ t:\ 1 \leq t \leq m-1 \\ r_t(n) = q_{t-(m-1)}(n) - g_1(n)\,,\ t:\ m \leq t \leq m+l-1 \\ r_t(n) = g_{t-(m+l-1)}(n+h) - g_1(n)\,,\ t:\ m+l \leq t \leq 2m+l-1 \\ r_t(n) = q_{t-(2m+l-1)}(n+h) - g_1(n)\,,\ t:\ 2m+l \leq t \leq 2m+2l-1. \end{cases}$$

For all but a finite number of h's the polynomials $(r_t(n))_{t=1}^{2m+2l-1}$ are essentially distinct. We notice that if we take two polynomials r_t's from the same group (there are 4 groups), then their difference is a non-constant because the initial polynomials are essentially distinct. If we take two polynomials from different groups then three cases are possible. In the first case the difference of these polynomials is $g_t(n+h) - g_t(n)$ or $q_t(n+h) - q_t(n)$ for some t. The assumption $i > 1$ implies that $\deg(q_t), \deg(g_t) > 1$ and from this it follows that $g_t(n+h) - g_t(n)$ and $q_t(n+h) - q_t(n)$ are non-constant polynomials. In the second case we get for some $t_1 \neq t_2$: $g_{t_1}(n+h) - g_{t_2}(n)$ or $q_{t_1}(n+h) - q_{t_2}(n)$. Here we note that the map $h \mapsto p(n+h)$ is an injective map from $\mathbb{N}$ to the set of essentially distinct polynomials, if $\deg(p) > 1$. Thus, for all but a finite number of h's we get again a non-constant difference. In the third case we get for some t_1, t_2: $g_{t_1}(n+h) - q_{t_2}(n)$ or $q_{t_1}(n+h) - g_{t_2}(n)$. The resulting polynomial has the same degree as q_t.
The characteristic vector of the set of polynomials $\{r_1, \ldots, r_{2m+2l-1}\}$ has the form $(c_1, \ldots, c_{i-1}, n_i, n_{i+1}, \ldots, n_d)$. The polynomials from the second and the fourth group have the same degree as q_t and the same leading coefficient as q_t if $\deg(q_t) > \deg(g_1)$ and the leading coefficient will be the difference of leading coefficients of q_t and g_1 if $\deg(q_t) = \deg(g_1)$. The polynomials from the first and the third group will be of degree smaller than $\deg(g_1)$.
$L(0; \overline{n_1, \ldots, n_{i-1}}, n_i, \ldots, n_d)$[9] and Cauchy-Schwartz inequality imply that for all but a finite number of h's there exists $J(\varepsilon, h)$ such that for every $J \geq J(\varepsilon, h)$ there exists $N(J)$ such that for every $N \geq N(J)$ and any $\{0,1\}$-valued sequence $(a_n)_{n\in\mathbb{N}}$ we have[10]

$$\left|\frac{1}{J}\sum_{j=1}^{J}\langle u_j, u_{j+h}\rangle_{q(N)}\right| < \frac{\varepsilon}{2}.$$

Van der Corput lemma implies that there exists $J(\varepsilon)$ such that for every $J \geq J(\varepsilon)$ there exists $N(J)$ such that for every $N \geq N(J)$ and any $\{0,1\}$-valued sequence

[9] Notice that for all t the polynomial $q(n)$ grows faster to infinity than $r_t(n)$.
[10] The sequence $(a_n)_{n\in\mathbb{N}}$ is involved in the definition of u_j's.

$(a_n)_{n\in\mathbb{N}}$ we have

$$\left\|\frac{1}{J}\sum_{j=1}^{J} u_j\right\|_{q(N)} < \varepsilon.$$

Thus we have shown the validity of $L(k;\underbrace{0,\dots,0}_{i-1 zeros},n_i+1,n_{i+1},\dots,n_d)$.

Proof of $S.1_d$**.** Let F be a family of essentially distinct polynomials having the characteristic vector $(n_1+1,n_2,\dots,n_d)$. Denote the linear polynomials from F by[11] $g_1(n)=c_1n+d_1,\dots,g_{n_1+1}=c_{n_1+1}n+d_{n_1+1}$. The remaining polynomials in F we denote by $q_1,\dots,q_l$. Let $(a_n)_{n\in\mathbb{N}}$ be a $\{0,1\}$-valued sequence, $\boldsymbol{h}\in\{1,2,\dots,H\}^k$ and $\boldsymbol{c}_i\in(\mathbb{Z}\setminus\{0\})^k$ for $1\le i\le n_1+1$. Denote by $u_j(n)$ the following vectors:

$$u_j(n)=a_{N+j}\prod_{i=1}^{n_1+1}\prod_{\epsilon\in\{0,1\}^k}\xi(n-g_i(N+j)-(\boldsymbol{c}_i\epsilon)\cdot\boldsymbol{h})\prod_{i=1}^{l}\xi(n-q_i(N+j)),$$

$$n=1,\dots,q(N).$$

Denote by $b_{N+j}=a_{N+j}a_{N+j+h}$, $r_i(n)=(c_{i+1}-c_1)n+(d_{i+1}-d_1)\,,\ i:1\le i\le n_1$, $s_i(n)=q_i(n)-g_1(n)\,,\ t_i(n)=q_i(n+h)-g_1(n)\,,\ i:1\le i\le l$. Then we have

$$\frac{1}{J}\sum_{j=1}^{J}\langle u_j,u_{j+h}\rangle_{q(N)}=\delta_{N,J}+$$

$$\langle\prod_{\epsilon\in\{0,1\}^k}\psi_1(n-(\boldsymbol{c}_i\epsilon)\cdot\boldsymbol{h}),\frac{1}{J}\sum_{j=1}^{J}b_{N+j}\prod_{i=1}^{n_1}\prod_{\epsilon\in\{0,1\}^k}\psi_2^i(n-r_i(N+j)-(\boldsymbol{c}_i\epsilon)\cdot\boldsymbol{h})\prod_{i=1}^{l}\psi_3^i(n)\rangle_{q(N)},$$

where

$$\psi_1(n)=\xi(n)\xi(n-c_1h),$$

$$\psi_2^i(n)=\xi(n)\xi(n-c_{i+1}h),$$

$$\psi_3^i(n)=\xi(n-s_i(N+j))\xi(n-t_i(N+j)).$$

Notice that $\delta_{N,J}\to 0$ as $N\to\infty$. For every $i\,:\,1\le i\le l$ the polynomials s_i,t_i are in the same group (they have the same degree and the same leading coefficient), therefore the characteristic vector of the family $\{s_1,t_1,\dots,s_l,t_l\}$ is the same as of the family $\{s_1,s_2,\dots,s_l\}$ and the latter family has the same characteristic vector as the family $\{q_1,q_2,\dots,q_l\}$. Thus the characteristic vector of the family $\{s_1,t_1,\dots,s_l,t_l\}$ is equal to $(0,n_2,n_3,\dots,n_d)$. $L(k+1;n_1,\dots,n_d)$, Cauchy-Schwartz inequality and van der Corput lemma imply the validity of $L(k;n_1+1,n_2,\dots,n_d)$.

Proof of $S.3_d$**:** It is similar to the proof of $S.2_{d,i}$.

Proof of $L(k;1)\,,\ \forall k\in\mathbb{N}_0$**:**
Let $g_1(n)=c_1n+d_1$ with $c_1>0$, $\boldsymbol{c}_1=(c_1^1,\dots,c_k^1)\in(\mathbb{Z}\setminus\{0\})^k$ and $q(n)\in\mathbb{Z}[n]$ with $q(n)-g_1(n)\to\infty$ as $n\to\infty$. We need to prove the following statement.

For every $\varepsilon,\delta>0$ there exists $H(\delta,\varepsilon,\boldsymbol{c}_1)$ such that for every $H\ge H(\delta,\varepsilon,\boldsymbol{c}_1)$ there exists $J(H)$ such that for every $J\ge J(H)$ there exists $N(J)$ such that for every

[11] In any group of degree one there is only one polynomial.

$N \geq N(J)$ for a set of $(h_1,\dots,h_k) \in \{1,\dots,H\}^k$ of density which is at least $1-\delta$ for any $\{0,1\}$-valued sequence $(a_n)_{n\in\mathbb{N}}$ we have

$$\left\| \frac{1}{J}\sum_{j=1}^{J} a_{N+j} \prod_{\epsilon\in\{0,1\}^k} \xi(n - g_1(N+j) - \epsilon_1 c_1^1 h_1 - \ldots - \epsilon_k c_k^1 h_k) \right\|_{q(N)} < \varepsilon.$$

By total ergodicity of the sequence $(\xi(n)_{n\in\mathbb{N}_0})$ there exist a totally ergodic system $(X, \mathbb{B}_X, \mu, T)$, a function $f \in L^\infty(X)$ and an f-generic point $x_0 \in X$ such that

$$\xi(n) = f(T^n x_0),\ \forall n \in \mathbb{N}_0.$$

Let $(b_j)_{j\in\mathbb{N}}$ be a $\{0,1\}$-valued sequence. Then by f-genericity of ξ we have

$$\frac{q(N)}{q(N)-g_1(N)} \left\| \frac{1}{J}\sum_{j=1}^{J} b_j \prod_{\epsilon\in\{0,1\}^k} \xi(n - g_1(N+j) - \epsilon_1 c_1^1 h_1 - \ldots - \epsilon_k c_k^1 h_k) \right\|_{q(N)}^2 \to$$

(3.1)

$$\int_X \left(\frac{1}{J}\sum_{j=1}^{J} b_{J+1-j} T^{c_1 j} \left(\prod_{\epsilon\in\{0,1\}^k} T^{\epsilon_1 c_1^1 h_1 + \ldots + \epsilon_k c_k^1 h_k} f(x) \right) \right)^2 d\mu(x) \text{ as } N \to \infty.$$

Denote by $g_{h_1,\dots,h_k}$ the following function on X:

$$g_{h_1,\dots,h_k}(x) = \prod_{\epsilon\in\{0,1\}^k} T^{\epsilon_1 h_1 + \ldots + \epsilon_k h_k} f(x).$$

The following statement is a corollary of Theorem 13.1 of Host and Kra in [4].

For every $\varepsilon, \delta > 0$ there exists $H(\delta,\varepsilon) \in \mathbb{N}$ such that for every $H \geq H(\delta,\varepsilon)$ for a set of $(h_1,\dots,h_k) \in \{1,\dots,H\}^k$ which has density at least $1-\delta$ we have[12]

$$\left| \int_X g_{h_1,\dots,h_k}(x) d\mu(x) \right| < \varepsilon.$$

Let $\varepsilon, \delta > 0$. By the foregoing statement there exists $H(\delta,\varepsilon)$ such that for every $H \geq H(\delta,\varepsilon)$ the set of those $(h_1,\dots,h_k) \in \{1,\dots,H\}^k$ such that

$$\left| \int_X g_{h_1,\dots,h_k}(x) d\mu(x) \right| < \sqrt{\frac{\varepsilon}{8}}$$

has density at least $1-\delta$.
For any fixed $\boldsymbol{h} = (h_1,\dots,h_k)$ Lemma 4 implies that there exists $J(\varepsilon,\boldsymbol{h})$ such that for every $J \geq J(\varepsilon,\boldsymbol{h})$ and any $\{0,1\}$-valued sequence $(e_n)_{n\in\mathbb{N}}$ we have

$$\left\| \frac{1}{J}\sum_{j=1}^{J} e_j T^{c_1 j} \left(g_{h_1,\dots,h_k}(x) - \int_X g_{h_1,\dots,h_k}(x) d\mu(x) \right) \right\|_{L^2(X)} < \sqrt{\frac{\varepsilon}{8}}.$$

[12]The mean zero of ξ is equivalent to $\int f d\mu = 0$.

Therefore, by merging last two statements we conclude that there exists $H(\delta,\varepsilon)$ such that for every $H \geq H(\delta,\varepsilon)$ there exists $J(H)$ such that for every $J \geq J(H)$ and for a set of $(h_1,\ldots,h_k) \in \{1,\ldots,H\}^k$ which has density at least $1-\delta$ we have

$$\left\| \frac{1}{J}\sum_{j=1}^{J} e_j T^{c_1 j} g_{h_1,\ldots,h_k}(x) \right\|_{L^2(X)} < \sqrt{\frac{\varepsilon}{2}}$$

for any $\{0,1\}$-valued sequence $(e_j)_{j\in\mathbb{N}}$.

By making δ smaller we conclude that the same statement holds true when we replace $g_{h_1,\ldots,h_k}(x)$ by the function

$$\prod_{\epsilon\in\{0,1\}^k} T^{\epsilon_1 c_1^1 h_1+\ldots+\epsilon_k c_k^1 h_k} f(x).$$

By (3.1), the fact that $\lim_{N\to\infty}\frac{q(N)}{q(N)-g_1(N)} > 0$ and the last statement we get that there exists $N(J)$ such that for every $N \geq N(J)$, for a set of $\boldsymbol{h} \in \{1,\ldots,H\}^k$ of density $1-\delta$ and every $\{0,1\}$-valued sequence $(b_j)_{1\leq j\leq J}$ we have

$$\left\| \frac{1}{J}\sum_{j=1}^{J} b_j \prod_{\epsilon\in\{0,1\}^k} \xi(n-g_1(N+j)-\epsilon_1 c_1^1 h_1-\ldots-\epsilon_k c_k^1 h_k) \right\|_{q(N)} < \varepsilon.$$

The latter statement implies the validity of $L(k;1)$.

□

Proof of Proposition 2. For a family of polynomials $F=\{p_1,\ldots,p_k\}$ with a maximal degree d denote by n_d the number of different leading coefficients of polynomials of degree d.

As in the proof of Lemma 2 we fix one of the groups of polynomials of degree d (all polynomials in the same group have the same leading coefficient). Assume that the group $\{g_1,\ldots,g_m\}$ has the maximal leading coefficient among all polynomials $p_1,\ldots,p_k$. The rest of the polynomials we denote by $q_1,\ldots,q_l$. Without loss of generality assume that $p_1=g_1,\ldots,p_m=g_m$. For any integer j denote by u_j the vector

$$u_j(n) = a_{N+j}\xi(p_1(N+j)-n)\xi(p_2(N+j)-n)\ldots\xi(p_k(N+j)-n),\ 1\leq n\leq p_1(N).$$

Denote by $r_i(n)=p_1(n)-q_i(n)$; $s_i(n)=p_1(n)-q_i(n+h)$, $i: 1\leq i\leq l$ and $t_i(n)=p_1(n)-p_i(n)$; $f_i(n)=p_1(n)-p_i(n+h)$, $i: 1\leq i\leq m$. Also denote by

$$\psi_i(x,y)=\xi(x-r_i(y))\xi(x-s_i(y)),\ \text{for } 1\leq i\leq l.$$

For any $h\geq 1$ we have

$$\frac{1}{J}\sum_{j=1}^{J}\langle u_j,u_{j+h}\rangle_{p_1(N)} = \delta_{J,N}+$$

$$\langle \xi(n), \frac{1}{J}\sum_{j=1}^{J} b_{N+j}\prod_{i=1}^{m-1}\xi(n-t_{i+1}(N+j))\prod_{i=1}^{l}\psi_i(n,N+j)\prod_{i=1}^{m}\xi(n-f_i(N+j))\rangle_{p_1(N)},$$

where $b_n=a_n a_{n+h}$ and $\delta_{J,N}\to 0$ as $N\to 0$.

For all but a finite number of h's the polynomials in the family

$$\tilde{F}=\{r_1,\ldots,r_l,s_1,\ldots,s_l,t_2,\ldots,t_m,f_1,\ldots,f_m\}$$

are essentially distinct and p_1 grows faster to infinity than any polynomial in $\tilde{F}$. For all but a finite number of h's, by Lemma 2 for any $\varepsilon > 0$ there exists $J(\varepsilon)$ such that for any $J \geq J(\varepsilon)$ there exists $N(J)$ such that for every $N \geq N(J)$ we have

$$\left\| \frac{1}{J} \sum_{j=1}^{J} b_{N+j} \prod_{i=1}^{m-1} \xi(n - t_{i+1}(N+j)) \prod_{i=1}^{l} \psi_i(n, N+j) \prod_{i=1}^{m} \xi(n - f_i(N+j)) \right\|_{p_1(N)} < \varepsilon.$$

Cauchy-Schwartz inequality and van der Corput's lemma imply the validity of the statement of the lemma.

□

4. Proof of Theorem 1

We remind the statement.

Theorem 1. *Let $A \subset \mathbb{N}$ be a TE set. Then for any $B \subset \mathbb{N}$ of positive upper density and any non-constant polynomial $p(n) \in \mathbb{Z}[n]$ with a positive leading coefficient we have $(A+B) \cap \{p(n) \,|\, n \in \mathbb{N}\} \neq \emptyset$. Moreover, if the lower density of B is positive then the set $R_p = \{n \in \mathbb{N} \,|\, p(n) \in A+B\}$ has bounded gaps.*

Proof. Let $B \subset \mathbb{N}$ be a set of positive upper density, $A \subset \mathbb{N}$ be a TE set and $p(n) \in \mathbb{Z}[n]$ a non-constant polynomial with a positive leading coefficient. Denote by $(\xi(n))_{n \in \mathbb{N}_0}$ the sequence[13]

$$\xi(n) = 1_A(n) - d(A).$$

Denote by $c = \overline{d}(B) > 0, u_j(n) = \xi(p(N+j) - n)$; $1 \leq n \leq p(N)$, $1 \leq j \leq J$.
If $(A+B) \cap \{p(n) | n \in \mathbb{N}\} = \emptyset$ then for any $b \in B$ and for all N, j we have $p(N+j) - b \notin A$. Thus

$$\langle 1_B, \frac{1}{J} \sum_{j=1}^{J} u_j \rangle_{p(N)} = \frac{1}{p(N)} \sum_{n=1}^{p(N)} 1_B(n) \frac{1}{J} \sum_{j=1}^{J} \xi(p(N+j) - n) =$$

$$-d(A) \frac{|B \cap \{1, 2, \ldots, p(N)\}|}{p(N)}.$$

Therefore for infinitely many N's we have[14]

$$\left| \langle 1_B, \frac{1}{J} \sum_{j=1}^{J} u_j \rangle_{p(N)} \right| \geq \frac{d(A)c}{2}.$$

Cauchy-Schwartz inequality together with Proposition 1 imply a contradiction.

If we assume that the lower density of B is positive, then for N big enough and for any J we have

$$\left| \langle 1_B, \frac{1}{J} \sum_{j=1}^{J} u_j \rangle_{p(N)} \right| \geq \frac{d(A)\underline{d}(B)}{2}.$$

---

[13]We assume that $\xi(0) = 0$.

[14]$\overline{d}(B) > 0$ implies that there exists a subsequence $(N_k)_{k \in \mathbb{N}}$ such that for every k we have $\frac{|B \cap \{1,2,\ldots,p(N_k)\}|}{p(N_k)} > \frac{c}{2}$. The latter uses that $\frac{p(N+1)}{p(N)} \to 1$ as $N \to \infty$.

Again by Cauchy-Schwartz inequality and Proposition 1 we get that there exist J big enough and $N(J)$ such that for every $N \geq N(J)$ we have $(A+B) \cap \{p(N+1), \dots, p(N+J)\} \neq \emptyset$. Thus the set

$$R_p = \{n \in \mathbb{N} \,|\, p(n) \in A+B\}$$

has bounded gaps. □

5. Proof of Theorem 3

We remind the statement.

Theorem 3. *Let $A \subset \mathbb{N}$ be a WM set, let $p_1(n), \dots, p_k(n) \in Z[n]$ be essentially distinct polynomials of the same degree, let $B \subset \mathbb{N}$ of positive upper density. Then the set*

$$R_{p_1,\dots,p_k} = \{n \in \mathbb{N} \,|\, \exists b \in B : \, p_1(n), p_2(n), \dots, p_k(n) \in A+b\}$$

has lower Banach density 1.

Proof. Let A be a WM set and let $p_1, \dots, p_k \in \mathbb{Z}[n]$ be essentially distinct polynomials of the same degree $d \geq 1$ with positive leading coefficients. Assume that for sufficiently large n's we have $p_1(n) > p_i(n)\,, \forall i : \, 2 \leq i \leq k$. We notice that $n \in R_{p_1,\dots,p_k}$ if and only if there exists $(x, y_1, \dots, y_k) \in B \times A^k$ such that the system

$$\begin{cases} x + y_1 = p_1(n) \\ x + y_2 = p_2(n) \\ \dots \\ x + y_k = p_k(n) \end{cases} \tag{5.1}$$

holds. Let F be the set of all n's for which the statement of the theorem fails.

$$F = \{n \in \mathbb{N} \,|\, \text{for any } (x, y_1, \dots, y_k) \in B \times A^k \text{ the system (5.1) fails to hold}\}.$$

We prove that $d^*(F) = 0$. Denote by $(a_n)_{n \in \mathbb{N}}$ the indicator sequence of F, i.e., $a_n = 1_F(n)$. Let ξ be the sequence

$$\xi(n) = 1_A(n) - d(A), \text{ for all } n \in \mathbb{N}.$$

Denote by $B_{N,J}$ the following expression

$$B_{N,J} = \langle 1_B(n), \frac{1}{J} \sum_{j=1}^{J} a_{N+j} 1_A(p_1(N+j) - n) \dots 1_A(p_k(N+j) - n) \rangle_{p_1(N)}.$$

Suppose that $d^*(F) > 0$. Then for every J there exist intervals $(I_\ell^J)_{\ell \in \mathbb{N}}$ such that $I_\ell^J = \{N_\ell^J + 1, \dots, N_\ell^J + J\}$ and $N_\ell^J \to \infty$ as $\ell \to \infty$. Also we demand from $(I_\ell^J)_{\ell \in \mathbb{N}}$ that $\frac{|F \cap I_\ell^J|}{J} > \frac{d^*(F)}{2}$ for J big enough and every ℓ. Denote by

$$c = \min_{2 \leq i \leq k} \frac{c_i}{c_1},$$

where c_i is a leading coefficient of polynomial p_i. Proposition 2 implies the validity of the following statement.

Claim 1: *For any $\varepsilon > 0$ there exists $J(\varepsilon)$ such that for every $J \geq J(\varepsilon)$ there exists $\ell(J)$ such that for every $\ell \geq \ell(J)$ we have*

$$\left| B_{N_\ell^J, J} \right| \geq c(1-\varepsilon) \overline{d}(B) d^k(A) \frac{d^*(F)}{3}.$$

For i, $0 \le i \le k-1$ denote by $\psi_j(x,y)$ and by $\phi_j(x,y)$ the following expressions:

$$\psi_j(x,y) = \prod_{m=1}^{i} 1_A(p_m(x+j)-y),\ \phi_j(x,y) = \prod_{m=i+1}^{k} \xi(p_m(x+j)-y).$$

By Proposition 2 and an induction on i, $0 \le i \le k-1$ the following statement holds true.

Claim 2: *For any $\varepsilon > 0$ there exists $J(\varepsilon)$ such that for any $J \ge J(\varepsilon)$ there exists $\ell(J)$ such that for every $\ell \ge \ell(J)$ and any $\{0,1\}$-valued seuquence $(b_n)_{n\in\mathbb{N}}$ we have*[15]

$$\left| \langle 1_B(n), \frac{1}{J}\sum_{j=1}^{J} b_{N_\ell^J+j}\psi_j(N_\ell^J,n)\phi_j(N_\ell^J,n)\rangle_{p_1(N_\ell^J)} \right| < \varepsilon.$$

Claim 2 for $i = k-1$ and an induction on k imply the validity of Claim 1.

By the definition of F it follows that for every J and N the expression $B_{N,J} = 0$. The latter contradicts Claim 1. The latter implies that, indeed, $d^*(F) = 0$. □

6. Appendix

Lemma 3. *(van der Corput) Let $\varepsilon > 0$ and $(u_j)_{j\in\mathbb{N}}$ be a bounded sequence of vectors in Hilbert space. There exists $I(\varepsilon)$ such that for every $I \ge I(\varepsilon)$ there exists $J(I)$, such that for any $J \ge J(I)$ for which we have*

$$\left| \frac{1}{J}\sum_{j=1}^{J} \langle u_j, u_{j+i}\rangle \right| < \frac{\varepsilon}{2},$$

for set of i's in the interval $\{1,\dots,I\}$ of density $1-\frac{\varepsilon}{3}$ the following holds

$$\left\| \frac{1}{J}\sum_{j=1}^{J} u_j \right\| < \varepsilon.$$

This is a finitary modification of Bergelson's lemma in [1]. Its proof may be found in [2], Lemma 5.1.

The following lemma is a simple fact that for a weakly mixing system we have a convergence in L^2-norm even of weighted ergodic averages. The precise statement is the follwoing.

Lemma 4. *Let $(X,\mathbb{B},\mu,T)$ be a weakly mixing system and $f \in L^2(X)$ with $\int_X f d\mu = 0$. Let $\varepsilon > 0$. There exists $J(\varepsilon)$ such that for any $J > J(\varepsilon)$ and any $\{0,1\}$-valued sequence $(b_n)_{n\in\mathbb{N}}$ we have*

$$\left\| \frac{1}{J}\sum_{j=1}^{J} b_j T^j f \right\|_{L^2(X)} < \varepsilon.$$

[15] The statement holds true for any integer k.

Proof. Weak mixing implies that for any $f \in L^2(X)$ with $\int_X f d\mu(x) = 0$ we have

$$\frac{1}{N}\sum_{n=1}^{N} |\langle T^n f, f\rangle| \to 0.$$

Denote by $c_n = c_{(-n)} = |\langle T^n f, f\rangle|$. Then we have

$$\frac{1}{N}\sum_{n=1}^{N} c_n \to 0 \text{ as } N \to \infty. \tag{6.1}$$

Let $\varepsilon > 0$. From (6.1) it follows that there exists $J(\varepsilon)$ such that for any $J > J(\varepsilon)$ we have

$$\left\| \frac{1}{J}\sum_{j=1}^{J} b_j T^j f \right\|^2 \le \frac{1}{J^2}\sum_{j,k=1}^{J} b_j b_k c_{j-k} \le \frac{1}{J^2}\sum_{j,k=1}^{J} c_{j-k} < \varepsilon.$$

□

References

[1] Bergelson, V. Weakly mixing PET. Ergodic Theory Dynam. Systems 7 (1987), no. 3, 337–349.

[2] Fish, A. Solvability of linear equations within weak mixing sets, to appear in Israel Journal of Mathematics.

[3] Fish, A. Ramsey Theory of subsets of ℕ, Ph.D. thesis, Hebrew University of Jerusalem, 2007.

[4] Host, B.; Kra, B. Nonconventional ergodic averages and nilmanifolds. Ann. of Math. (2) 161 (2005), no. 1, 397–488.

Department of Mathematics, University of Wisconsin, Madison 480 Lincoln Drive Madison, WI 53706

E-mail address: afish@math.wisc.edu