



Complexities of finite families of polynomials, Weyl systems, and constructions in combinatorial number theory

Vitaly Bergelson, Alexander Leibman, Emmanuel Lesigne

► To cite this version:

Vitaly Bergelson, Alexander Leibman, Emmanuel Lesigne. Complexities of finite families of polynomials, Weyl systems, and constructions in combinatorial number theory. *Journal d'analyse mathématique*, 2007, 103, pp.47-92. hal-00017730v2

HAL Id: hal-00017730

<https://hal.science/hal-00017730v2>

Submitted on 24 Oct 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Complexities of finite families of polynomials, Weyl systems, and constructions in combinatorial number theory

V. Bergelson*, A. Leibman*, and E. Lesigne

July 12, 2007

Abstract

We introduce two notions of complexity of a system of polynomials $p_1, \dots, p_r \in \mathbb{Z}[n]$ and apply them to characterize the limits of the expressions of the form $\mu(A_0 \cap T^{-p_1(n)} A_1 \cap \dots \cap T^{-p_r(n)} A_r)$ where T is a skew-product transformation of a torus $\mathbb{T}^d$ and $A_i \subseteq \mathbb{T}^d$ are measurable sets. The dynamical results obtained allow us to construct subsets of integers with specific combinatorial properties related to the polynomial Szemerédi theorem.

Table of Contents

0. Introduction	1
A. Configurations in sets of positive density	1
B. Dynamically defined sets of integers	3
C. Formulation of main results	6
1. Orbit of the diagonal	9
2. Span and Rank of a system of polynomials	10
3. Weyl dynamical systems	12
4. Weyl and Vandermonde complexities	16
5. Properties of Vandermonde and Weyl complexities and examples	19
6. Characteristic factors and large intersections	24
7. Combinatorics	33

0. Introduction

A. Configurations in sets of positive density

The Szemerédi theorem on arithmetic progressions ([Sz]) states that if a set $E \subseteq \mathbb{Z}$ has positive upper Banach density, then for any $r \in \mathbb{N}$, there exist integers a and $n \neq 0$ such that $\{a, a + n, a + 2n, \dots, a + rn\} \subset E$. (We recall that the upper Banach density $d^*(E)$ of a set E of integers is given by $d^*(E) = \limsup_{N-M \rightarrow \infty} \frac{|E \cap [M, N]|}{N-M+1}$. When the limit $\lim_{N-M \rightarrow \infty} \frac{|E \cap [M, N]|}{N-M+1}$ exists we will denote it by $d_u(E)$.) Szemerédi's theorem has

*Bergelson and Leibman were supported by NSF grants DMS-0345350 and DMS-0600042.

been extended to multidimensional “polynomial” progressions ([BL]). In particular, it was proven in [BL] that if $p_1, \dots, p_r$ are integral polynomials (that is, polynomials with rational coefficients taking integer values on $\mathbb{Z}$) with $p_i(0) = 0$ for $i = 1, \dots, r$, then any set $E \subseteq \mathbb{Z}$ with $d^*(E) > 0$ contains configurations of the form $C(a, n) = \{a, a + p_1(n), a + p_2(n), \dots, a + p_r(n)\}$ with $n \in \mathbb{Z} \setminus \{0\}$. Moreover, in the linear case as well as the polynomial case, the set $\{n \in \mathbb{Z} : C(a, n) \subset E \text{ for some } a \in \mathbb{Z}\}$ is *syndetic*, i.e., this set meets any sufficiently long interval in $\mathbb{Z}$. (See [F1] and [BM].)

The homogeneity condition, $p_i(0) = 0$, in the formulation is quite essential. Thus for $p(n) = 2n + 1$, a simple non-homogeneous polynomial, the 2-element configuration $\{a, a + p(n)\}$ does not appear in $E = 2\mathbb{Z}$ or $E = 2\mathbb{Z} + 1$. (By a change of variables $q_i(n) = p_i(n - n_0)$ the homogeneity condition can be replaced by the condition that the polynomials vanish simultaneously for some $n_0 \in \mathbb{Z}$.)

It is significant that the foregoing counterexample to existence of a non-homogeneous configuration is rooted in a “non-randomness” displayed by the set E . We will be giving a precise definition of “randomness” of a set of integers, and for sets E with this property and $d^*(E) > 0$, *every* non-trivial polynomial progression type will occur in E . (A polynomial progression type is trivial if for some $i \neq j$, $p_i - p_j = \text{const.}$) For sets with only the condition $d^*(E) > 0$, the existence of a specific non-homogeneous configuration depends on the configuration in question. For example, as we will see, one can produce a set E with $d_u(E) > 0$ so that *all* linear 2-term configuration occur (and moreover, do so in abundance) but some linear 3-term configurations do not occur.

The length of the configuration plays a role here and the foregoing example is generalized in the following.

Proposition 0.1. *Let $r \in \mathbb{N}$, let $b_1, b_2, \dots, b_r$ be distinct integers, and let $m \in \mathbb{N}$ and $l \in \mathbb{Z}$, $l \not\equiv 0 \pmod{m}$. There exists a set E with $d_u(E) > 0$ such that*

(i) *For any integers $b'_1, b'_2, \dots, b'_{r-1}$, $m' \in \mathbb{N}$ and $l' \in \mathbb{Z}$ the set*

$$\mathcal{N}' = \{n \in \mathbb{Z} : \{a, a + b'_1(m'n + l'), a + b'_2(m'n + l'), \dots, a + b'_{r-1}(m'n + l')\} \subset E \text{ for some } a \in \mathbb{Z}\}$$

is syndetic, but

(ii) *the set*

$$\mathcal{N} = \{n \in \mathbb{Z} : \{a, a + b_1(mn + l), a + b_2(mn + l), \dots, a + b_r(mn + l)\} \subset E \text{ for some } a \in \mathbb{Z}\}$$

is empty.

One can ask quite generally if for two kinds of configurations there exists a set of positive upper density containing configurations of one kind but not of the other kind. We will give a partial answer to this based on an appropriate notion of the “complexity” of a polynomial configuration, for which, given polynomial systems $P = \{p_1, \dots, p_r\}$ and $Q = \{q_1, \dots, q_s\}$ of different complexity, there exists a set E with $d_u(E) > 0$ containing one configuration but not the other. To each system of homogeneous (vanishing at 0) polynomials, $P = \{p_1, \dots, p_r\}$ we assign a complexity index and determine a particular set

E_P of positive density for which this complexity index will play a role. Namely, there will be a non-homogeneous version of a P -progression, specifically, $C(a, n) = \{a, a + p_1(mn + l), a + p_2(mn + l), \dots, a + p_r(mn + l)\}$ for some $m, n \neq 0$ so that no $C(a, n)$ appears in E_P with $n \neq 0$, while at the same time, for every polynomial system $Q = \{q_1, \dots, q_s\}$ of complexity less than P , the analogous configurations $C'(a, n) = \{a, a + q_1(mn + l), a + q_2(mn + l), \dots, a + q_s(mn + l)\}$ appear in E_P for a syndetic set of $n \in \mathbb{Z}$ for every choice of m and l . The foregoing can be strengthened and we can speak of arbitrary “composite Q -configurations” $\{a, a + q_1(h(n)), a + q_2(h(n)), \dots, a + q_s(h(n))\}$ for any non-constant integral polynomial h , and these too will be found in E_P , when the complexity of Q is less than that of P , for a syndetic set of n .

Remark. As a matter of fact, we will find it convenient to introduce not one but two complexity indices, $W(P)$ and $V(P)$ (Weyl complexity and Vandermonde complexity). The reason for doing so and the pertinent details will be clarified in the ensuing discussion.

B. Dynamically defined sets of integers

As in the ergodic theoretical proof of Szemerédi’s theorem (as well as its extension to polynomial progressions), the analogy between sets of integers with positive upper Banach density and subsets of positive measure in the state space of a dynamical system plays a role in the present discussion, and the sets E_P referred to in the foregoing section will be defined using certain dynamical systems called *Weyl systems*. We will not be using ergodic theoretic techniques explicitly, but some important ergodic theoretical ideas underly the construction that we will try to explain in this section.

We recall that an ergodic system $(X, \mathcal{B}, \mu, T)$ consists of a probability space $(X, \mathcal{B}, \mu)$ and an (invertible) measure preserving transformation $T: X \rightarrow X$ having no nontrivial $A \in \mathcal{B}$ ($0 < \mu(A) < 1$) invariant. Ergodic systems form a plentiful source of sets of integers of positive density. By the ergodic theorem, if $A \in \mathcal{B}$ has $\mu(A) > 0$, then for almost every $x \in X$ the visiting time set $E(A, x) = \{n : T^n x \in A\}$ has density equal to $\mu(A)$, and particularly $d^*(E(A, x)) > 0$. A simple but crucial observation is that outside of a (fixed) set of measure 0 of points $x \in X$, a configuration $C(a, n) = \{a, a + p_1(n), a + p_2(n), \dots, a + p_r(n)\}$ will occur in $E(A, x)$ if $\{p_1(n), p_2(n), \dots, p_r(n)\}_{n \in \mathbb{Z}}$ is a “recurrence pattern” for A , i.e., if for some n

$$\mu(A \cap T^{-p_1(n)} A \cap \dots \cap T^{-p_r(n)} A) > 0. \quad (0.1)$$

The initial a is determined by the condition $T^a x \in A \cap T^{-p_1(n)} A \cap \dots \cap T^{-p_r(n)} A$ and thus (for a.e. x) forms a set of positive density. A converse of the foregoing is also valid: if $A \cap T^{-p_1(n)} A \cap \dots \cap T^{-p_r(n)} A = \emptyset$ (rather than just a set of measure 0) then clearly $C(a, n)$ cannot occur in any $E(A, x)$. This makes it plausible that sets of the form $E(A, x)$ will play an important role for our purposes.

The notion of recurrence patterns for ergodic systems has been studied extensively, and it is known ([B]) that for the class of weakly mixing (WM) systems, these include all polynomial configurations (homogeneous or not) provided the p_i have positive degree and do not differ by constants. We can now give a precise definition to the notion of randomness of a set E alluded to in §A. We call E a WM set if $E = E(A, x)$ for an invertible WM system $(X, \mathcal{B}, \mu, T)$ where x is “generic” for $(X, \mathcal{B}_A, \mu, T)$; that is, $\frac{1}{N} \sum_{n=1}^N 1_B(T^n x) \rightarrow \mu(B)$ for

every set $B \in \mathcal{B}_A$, the algebra of sets spanned by $\{T^n A\}_{n \in \mathbb{Z}}$.

For general ergodic systems it has been shown ([BL]) that *any* system of *homogeneous* polynomials forms a recurrence pattern, and this underlines the fact that homogeneous progressions occur for any E with $d^*(E) > 0$. (Here E need not have the form $E(A, x)$ exactly as above, but this will hold for an extended notion of genericity.)

In this paper we confine ourselves to the Weyl systems. A *Weyl system* is an ergodic system $(X, \mathcal{B}, \mu, T)$ where X is a compact commutative Lie group (which is either a torus or the product of a torus and of a finite commutative group) and T is a unipotent affine transformation thereof. To motivate the special role of Weyl systems in our discussion, it will be instructive to make reference to the notion of characteristic factors for non-conventional ergodic averages ([HK1] and [Z]). Recall that a factor $(Y, \mathcal{D}, \nu, S)$ of an ergodic system $(X, \mathcal{B}, \mu, T)$ is an ergodic system with a measurable, measure preserving map $\pi: X \rightarrow Y$ satisfying $\pi \circ T = S \circ \pi$. One says that $(Y, \mathcal{D}, \nu, S)$ is a *characteristic factor* of $(X, \mathcal{B}, \mu, T)$ for the pattern $\{p_1(n), \dots, p_r(n)\}_{n \in \mathbb{Z}}$ if for any set of functions $f_1, \dots, f_r \in L^\infty(X)$ and their “projections” $\bar{f}_1, \dots, \bar{f}_r \in L^\infty(Y)$,

$$\lim_{N-M \rightarrow \infty} \left\| \frac{1}{N-M} \sum_{n=M+1}^N f_1(T^{p_1(n)}x) \dots f_r(T^{p_r(n)}x) - \frac{1}{N-M} \sum_{n=M+1}^N \bar{f}_1(T^{p_1(n)}x) \dots \bar{f}_r(T^{p_r(n)}x) \right\|_{L^2(X)} = 0.$$

The consequence of this relevant for our purposes is that for a system $(X, \mathcal{B}, \mu, T)$ and a characteristic factor $(Y, \mathcal{D}, \nu, S)$

$$\lim_{N-M \rightarrow \infty} \frac{1}{N-M} \sum_{n=M+1}^N \mu(A \cap T^{-p_1(n)}A \cap \dots \cap T^{-p_r(n)}A) > 0 \quad (0.2)$$

for every $A \in \mathcal{B}$ with $\mu(A) > 0$ if and only if

$$\lim_{N-M \rightarrow \infty} \frac{1}{N-M} \sum_{n=M+1}^N \mu(B \cap S^{-p_1(n)}B \cap \dots \cap S^{-p_r(n)}B) > 0 \quad (0.3)$$

for every $B \in \mathcal{D}$ with $\nu(B) > 0$.

A second fact relevant to us is the fundamental result in [HK1] and [Z] asserting that, for any linear pattern, every dynamical system has a characteristic factor which is a nilsystem or an inverse limit of such (a pro-nilsystem). A nilsystem is a system $(Y, \mathcal{D}, \nu, S)$ where Y is a homogeneous space G/Γ of a nilpotent Lie group G and Γ is a co-compact subgroup of G , and there is an element $g \in G$ so that $S(x\Gamma) = gx\Gamma$ for each $x \in G$. Let $G_2 = [G, G]$ and $G_l = [G_{l-1}, G]$ for $l \geq 2$; we define the “level” of a nilpotent group G as the first l such that $G_{l+1} = \{1\}$. In [HK1] and [Z], for every system X and every $k \in \mathbb{N}$ a natural factor with the structure of a level k pro-nilsystem is constructed; we will call it *the k th HKZ factor of X* . It is proven in [HK2] and [L1] that for any system X and any

“polynomial pattern” $P = \{p_1(n), \dots, p_r(n)\}$ a certain HKZ factor of X is characteristic. The minimal k such that for every system X the $(k-1)$ st HKZ factor of X is characteristic with respect to P is called *the complexity* of $P = \{p_1, \dots, p_r\}$ and is denoted by $C(P)$. The *Weyl complexity* of P , $W(P)$, is the complexity of P measured with respect to Weyl systems only; it is an open question whether $W(P)$ may differ from $C(P)$.

Suppose now that $E(A, x)$ is a dynamically defined set corresponding to a system $(X, \mathcal{B}, \mu, T)$, and that $A \cap T^{-p_1(n)} A \cap \dots \cap T^{-p_r(n)} A = \emptyset$ for all n . Then (0.2) does not hold and if $(Y, \mathcal{D}, \nu, S)$ is a nilsystem which is a characteristic factor for $\{p_1, \dots, p_r\}$, we will have for some $B \in \mathcal{D}$ with $\nu(B) > 0$, $\frac{1}{N-M} \sum_{n=M+1}^N \mu(B \cap S^{-p_1(n)} B \dots S^{-p_r(n)} B) \xrightarrow{N-M \rightarrow \infty} 0$.

This implies that for most $y \in Y$, $E(B, y)$ will be a set of positive density in which configurations $\{a, a + p_1(n), \dots, a + p_r(n)\}$ are rare, and it makes it plausible that for some y they will be absent entirely from $E(B, y)$. If this heuristic principle is valid then we can restrict to sets defined by nilsystems to obtain sets of positive density where certain polynomial progressions are absent. We remark now that Weyl systems can be viewed as a special class of nilsystems (see Section 3 below) and it will be shown in the sequel that these special systems will be adequate for our purposes. In the examples constructed from Weyl systems we will find a dichotomy: either a polynomial progression does not occur at all (nontrivially, i.e. not in the form $(a, a, \dots, a)$), or it occurs in “abundance”. In the latter case the corresponding configurations $C(a, n)$ occur in E with “positive density” of (a, n) . This will accord with the dynamical property inherent in (0.2).

In a way, Weyl systems form a rather natural class of ergodic systems for the task at hand. Indeed, it follows from [B] that dynamically generated WM sets (namely, the visiting time sets coming from weakly mixing systems) contain all possible polynomial configurations of the form $\{a, a + p_1(n), \dots, a + p_k(n)\}$ provided the (not necessarily homogeneous) polynomials p_i are not constant and do not differ by constants. Now, weak mixing is characterized by the absence of a Kronecker factor, namely the factor that is formed by a rotation on a compact monothetic commutative groups. It follows that dynamically generated sets of positive density with non-trivial combinatorial properties of the sort we are interested in should come from ergodic systems having nontrivial Kronecker factor. Since the Kronecker systems allow one to construct sets with rather restricted combinatorial properties (see the remark after Theorem 0.6 below), one is led to dealing with extensions of Kronecker systems, and Weyl systems form the simplest class of extensions which is sufficient for achieving the stated goals.

The actual calculation of $W(P)$ for a system P of polynomials will be carried out algebraically and we will find that a related calculation leads to another measure of complexity which we call *Vandermonde complexity* and denote by $V(P)$ (see Section 4). Our analysis will show that this parameter also plays a role in measuring the “size” of $\{(a, n) : C(a, n) \subset E\}$. The minimal complexity $V = W = 1$ is achieved by linearly independent systems of polynomials. Roughly speaking, the more linear relations are between polynomials $p_1, \dots, p_r$ and their powers $p_1^k, \dots, p_r^k$, $k \in \mathbb{N}$, the higher is the complexity of the system $\{p_1, \dots, p_r\}$, and the more difficult it will be to have associated configurations in a given set of integers.

Here are some values of complexities of polynomial systems:

1. $V(p_1, \dots, p_r) = W(p_1, \dots, p_r) = 1$ if the polynomials p_i are linearly independent;

2. $V(b_1n, b_2n, \dots, b_rn) = W(b_1n, b_2n, \dots, b_rn) = r$, and $V(b_1p(n), b_2p(n), \dots, b_rp(n)) = W(b_1p(n), b_2p(n), \dots, b_rp(n)) = r$ for any nonconstant polynomial p and any nonzero distinct integers $b_1, \dots, b_r$;
 3. $V(n, 2n, n^2) = 2$ and $W(n, 2n, n^2) = 3$;
 4. $V(n, 2n, n^3) = W(n, 2n, n^3) = 2$.
- (See Section 5 for details.)

C. Formulation of main results

Let $(X, \mathcal{B}, \mu, T)$ be a Weyl system. To make the discussion in this introduction more concrete, let us restrict ourselves to a *standard* Weyl system (X, T) : X is the d -dimensional torus $\mathbb{T}^d$ and $T(x_1, x_2, \dots, x_d) = (x_1 + \alpha, x_2 + x_1, \dots, x_d + x_{d-1})$, $(x_1, \dots, x_d) \in X$, where $\alpha \in \mathbb{T}$ is irrational. For $k = 1, \dots, d+1$, let $L_k = \{(0, \dots, 0, x_k, \dots, x_d) \in X\}$, and let X_{k-1} be the factor-torus X/L_k (with the canonical projection $X \rightarrow X_{k-1}$). One can show (see Section 3) that the tori $X_1, X_2, \dots$ are the HKZ factors of (X, T) . Let $P = \{p_1, \dots, p_r\}$ be a system of (distinct) integral polynomials with zero constant term; then $W(P) = k$ iff X_{k-1} is a characteristic factor for P , and we have:

Theorem 0.2. (See Proposition 6.2 below.) *Let $k = W(P)$ and let $A_0, \dots, A_r$ be measurable subsets of X independent of X_{k-1} . Then*

$$\lim_{N-M \rightarrow \infty} \frac{1}{N-M} \sum_{n=M+1}^N \mu(A_0 \cap T^{-p_1(n)} A_1 \cap \dots \cap T^{-p_r(n)} A_r) = \prod_{i=0}^r \mu(A_i).$$

One can also show (see Lemma 6.3) that the $(k-1)$ st factor X_{k-1} is “optimal” in this theorem and cannot be replaced by X_{k-2} .

The next, k -dimensional factor-torus $X_k = X/L_{k+1}$ of X is characteristic for P in a stronger sense:

Theorem 0.3. (See Proposition 6.5 below.) *Let $k = W(P)$ and let $A_0, \dots, A_r$ be measurable subsets of X independent of X_k . Then*

$$\lim_{n \rightarrow \infty} \mu(A_0 \cap T^{-p_1(n)} A_1 \cap \dots \cap T^{-p_r(n)} A_r) = \prod_{i=0}^r \mu(A_i).$$

Lemma 6.7 says that, again, X_k cannot be replaced by X_{k-1} in this theorem.

Turning to the Vandermonde complexity, and interpreting X as the direct sum $X_{k-1} \oplus L_k$ for any k , we have the following:

Theorem 0.4. (See Proposition 6.9 below.) *Let $k = V(P)$ and let $A_i = X_{k-1} \times I_i$, $i = 0, \dots, r$, where $I_0, \dots, I_r$ are subsets of L_k of positive measure. Then*

$$\lim_{N-M \rightarrow \infty} \frac{1}{N-M} \sum_{n=M+1}^N \mu(A_0 \cap T^{-p_1(n)} A_1 \cap \dots \cap T^{-p_r(n)} A_r) > 0.$$

Again, X_{k-2} does not work for this theorem, at least for a nonconnected Weyl system; this follows from Proposition 7.5.

The reader will find an example illustrating the above results at the beginning of Section 6.

The Weyl and Vandermonde complexities induce a hierarchy on the set of all systems of polynomials so that, applying the dynamical method to a suitable Weyl system, one can construct a set of integers which contains many configurations corresponding to systems of smaller complexities and no configuration of a certain form corresponding to a system of larger complexity. In order to give a more precise formulation of our results let us first introduce some notions describing the “largeness” and the “regularity” of occurrences of polynomial configurations in a set of integers.

For a set of integers E , we will say that E has *uniform density* α and write $UD(E) = \alpha$ if the limit $\lim_{N \rightarrow \infty} \frac{|E \cap \Phi_N|}{|\Phi_N|}$ exists and equals α for every Følner sequence $\{\Phi_N\}$ in $\mathbb{Z}$. (We recall that a Følner sequence in $\mathbb{Z}$ is a sequence (Φ_N) of finite sets $\Phi_N \subset \mathbb{Z}$ with the property that $\lim_{N \rightarrow \infty} \frac{|(\Phi_N + n) \Delta \Phi_N|}{|\Phi_N|} = 0$ for every $n \in \mathbb{Z}$.) For a sequence of real numbers α_n we will write $UC\text{-}\lim_n \alpha_n = \alpha$ if $\lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{n \in \Phi_N} \alpha_n = \alpha$ for any Følner sequence $\{\Phi_N\}$ in $\mathbb{Z}$. Let $E \subseteq \mathbb{Z}$ have positive uniform density. Let $P = \{p_1, \dots, p_r\}$ be a system of integral polynomials. We will say that E is *UC-positive* with respect to P (“UC” for “Uniform Cesàro”) if $UD(\{a \in \mathbb{Z} : a, a + p_1(n), \dots, a + p_r(n) \in E\})$ exists for all $n \in \mathbb{Z}$ and

$$UC\text{-}\lim_n UD(\{a \in \mathbb{Z} : a, a + p_1(n), \dots, a + p_r(n) \in E\}) > 0,$$

that is, the *UC-limit* exists and is positive. (In particular, this implies that there exists $\delta > 0$ such that the set of n for which

$$UD(\{a \in \mathbb{Z} : a, a + p_1(n), \dots, a + p_r(n) \in E\}) > \delta$$

is syndetic in $\mathbb{Z}$.)

We will say that E is *UC-balanced* with respect to P if

$$UC\text{-}\lim_n UD(\{a \in \mathbb{Z} : a, a + p_1(n), \dots, a + p_r(n) \in E\}) = UD(E)^{r+1}.$$

(In particular, this implies that for any $\varepsilon > 0$ the set of n for which

$$UD(\{a \in \mathbb{Z} : a, a + p_1(n), \dots, a + p_r(n) \in E\}) > UD(E)^{r+1} - \varepsilon$$

is syndetic in $\mathbb{Z}$.)

Finally, we will say that E is *balanced* with respect to P if

$$\lim_{n \rightarrow \infty} UD(\{a \in \mathbb{Z} : a, a + p_1(n), \dots, a + p_r(n) \in E\}) = UD(E)^{r+1}$$

(this implies that for any $\varepsilon > 0$ one has

$$UD(\{a \in \mathbb{Z} : a, a + p_1(n), \dots, a + p_r(n) \in E\}) > UD(E)^{r+1} - \varepsilon$$

for all but finitely many $n \in \mathbb{Z}$).

Our main combinatorial result is that using Weyl systems one can construct a set E of integers with strong combinatorial properties:

Theorem 0.5. (See Theorem 7.1 below.) *Let $\{p_1, \dots, p_r\}$ be a system of integral polynomials with zero constant term and let $k = V(p_1, \dots, p_r)$. There exists a set $E \subset \mathbb{Z}$ of positive uniform density such that*

- (i) *for any system $\{q_1, \dots, q_s\}$ of integral polynomials with zero constant term and with $V(q_1, \dots, q_s) < k$ and any nonconstant integral polynomial h , E is UC-positive with respect to the system $\{q_1(h(n)), \dots, q_s(h(n))\}$;*
- (ii) *for any system $\{q_1, \dots, q_s\}$ of integral polynomials with zero constant term and with $W(q_1, \dots, q_s) < k$, any nonconstant integral polynomial h and any integers $c_1, \dots, c_s$, E is UC-balanced with respect to the system $\{q_1(h(n)) + c_1, \dots, q_s(h(n)) + c_s\}$;*
- (iii) *for any system $\{q_1, \dots, q_s\}$ of integral polynomials with zero constant term and with $W(q_1, \dots, q_s) < k - 1$, any nonconstant integral polynomial h and any integers $c_1, \dots, c_s$, E is balanced with respect to the system $\{q_1(h(n)) + c_1, \dots, q_s(h(n)) + c_s\}$;*
- (iv) *there exist nonzero integers m and l such that E contains no configuration of the form $\{a, a + p_1(mn + l), \dots, a + p_r(mn + l)\}$, $a, n \in \mathbb{Z}$.*

Remark. The reader may notice that the assertion (i) of Theorem 0.5 is “weaker” than the assertions (ii) and (iii), since the “shifting” constants c_i are absent in it. It is not clear whether a “shifted” version of (i) is true; the methods employed in this paper do not allow one to get such a result. (See the remark after Proposition 6.9 below.)

The integers m and l appearing in the formulation of Theorem 0.5 are not arbitrary (for instance, l cannot be divisible by m because in this case the polynomial Szemerédi theorem guarantees the existence of corresponding configurations). In general, these m and l depend on the system $P = \{p_1, \dots, p_r\}$ (see Theorem 7.2). However, when the system P consists of linear polynomials $p_i(n) = b_i n$, $i = 1, \dots, r$, the foregoing nondivisibility restriction is the only restriction that cannot be avoided. This fact is reflected in the following theorem which, while similar in spirit to Theorem 0.5, has a different arrangement of quantifiers.

Theorem 0.6. (See Corollary 7.4 below.) *For any $r, m \geq 2$ there exists a set $E \subset \mathbb{Z}$ of positive uniform density such that*

- (i) *for any system $\{q_1, \dots, q_s\}$ of integral polynomials with zero constant term and with $V(q_1, \dots, q_s) < r$ and any nonconstant integral polynomial h , E is UC-positive with respect to the system $\{q_1(h(n)), \dots, q_s(h(n))\}$;*
- (ii) *for any system $\{q_1, \dots, q_s\}$ of integral polynomials with zero constant term and with $W(q_1, \dots, q_s) < r$, any nonconstant integral polynomial h and any integers $c_1, \dots, c_s$, E is UC-balanced with respect to the system $\{q_1(h(n)) + c_1, \dots, q_s(h(n)) + c_s\}$;*
- (iii) *for any system $\{q_1, \dots, q_s\}$ of integral polynomials with zero constant term and with $W(q_1, \dots, q_s) < r - 1$, any nonconstant integral polynomial h and any integers $c_1, \dots, c_s$, E is balanced with respect to the system $\{q_1(h(n)) + c_1, \dots, q_s(h(n)) + c_s\}$;*
- (iv) *E contains no arithmetic progressions of the form $\{a, a + (mn + l), \dots, a + r(mn + l)\}$, $a, n, l \in \mathbb{Z}$, with l not divisible by m .*

Remark. It is natural to inquire whether our examples can be constructed by utilizing Kronecker systems (rotations on compact commutative groups) which form a rather particular subclass of Weyl systems. It turns out that some special cases of the situation described in items (i) of Theorems 0.5 and 0.6 can indeed be obtained (albeit in a very

cumbersome way) with the help of Kronecker systems. On the other hand the closer scrutiny reveals that Kronecker systems can not provide the more complicated examples which appear in the formulations of these theorems.

The paper is organized as follows. In Section 1 we give a detailed description of the dynamical method that will be used throughout this paper. In Section 2 we introduce some linear algebra notation related to a system of polynomials. In Section 3 we define Weyl dynamical systems and discuss their elementary properties. In Section 4 we introduce the Weyl and the Vandermonde complexities of a system of integral polynomials; in Section 5 we describe their properties and give examples. In Section 6 we obtain measure-theoretical results similar to Theorems 0.2, 0.3, and 0.4. In Section 7 we prove (somewhat more precise versions of) Theorems 0.5 and 0.6.

Acknowledgement. We thank H. Furstenberg for constructive criticism and numerous useful suggestions.

1. Orbit of the diagonal

As we have already mentioned in the Introduction, the ergodic approach to combinatorics, which goes back to Furstenberg's seminal paper [F1], establishes a two-way connection between the family of sets of integers satisfying $d^*(S) > 0$ and the family of visiting time sets $E(A, x)$ in ergodic measure preserving systems. In one direction, this connection is manifested by Furstenberg's correspondence principle which is behind the derivation of Szemerédi's theorem from a corresponding multiple recurrence result. In the other direction, this connection provides a natural method of creating sets of integers with preassigned combinatorial properties. We will now describe this method in some detail.

Let $(X, \mathcal{B}, \mu, T)$ be an ergodic Borel measure preserving system on a compact space X , let A be a measurable set in X with $\mu(A) > 0$, let $x_0 \in X$, and let $P = \{p_1, \dots, p_r\}$ be a system of integer polynomials (or, at this stage, a system of integer valued sequences). Define $E = \{n \in \mathbb{N} : T^n x_0 \in A\}$. Then for $a, n \in \mathbb{N}$ we have

$$\begin{aligned} a \in E \cap (E - p_1(n)) \cap \dots \cap (E - p_r(n)) &\text{ iff } a, a + p_1(n), \dots, a + p_r(n) \in E \\ \text{iff } T^a x_0, T^{a+p_1(n)} x_0, \dots, T^{a+p_r(n)} x_0 \in A &\text{ iff } T^a x_0 \in A \cap T^{-p_1(n)} A \cap \dots \cap T^{-p_r(n)} A. \end{aligned}$$

Thus, for any $n \in \mathbb{N}$, the set $E_n = \{a : a, a + p_1(n), \dots, a + p_r(n) \in E\}$ is the same as $\{a : T^a x_0 \in A \cap T^{-p_1(n)} A \cap \dots \cap T^{-p_r(n)} A\}$.

Let $\Delta_{X^{r+1}} = \left\{ \begin{pmatrix} x \\ x \\ \vdots \\ x \end{pmatrix}, x \in X \right\}$ be the diagonal in X^{r+1} . Consider the “polynomial action” $g(n) \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} x_0 \\ T^{p_1(n)} x_1 \\ \vdots \\ T^{p_r(n)} x_r \end{pmatrix}$, $n \in \mathbb{N}$, on X^{r+1} corresponding to the system $\hat{P} = \{0, p_1, \dots, p_r\}$ and let $\mathcal{O}(\hat{P}, \Delta_{X^{r+1}}) = \bigcup_{n \in \mathbb{N}} g(n) \Delta_{X^{r+1}}$ be the “orbit” of $\Delta_{X^{r+1}}$ under this action. Then, for $n \in \mathbb{N}$,

$$\begin{aligned} T^a x_0, T^{a+p_1(n)} x_0, \dots, T^{a+p_r(n)} x_0 \in A &\text{ iff } x, T^{p_1(n)} x, \dots, T^{p_r(n)} x \in A \text{ for } x = T^a x_0 \\ \text{iff } g(n) \begin{pmatrix} x \\ x \\ \vdots \\ x \end{pmatrix} \in A^{r+1} &\text{ only if } g(n) \Delta_{X^{r+1}} \cap A^{r+1} \neq \emptyset \text{ only if } \mathcal{O}(\hat{P}, \Delta_{X^{r+1}}) \cap A^{r+1} \neq \emptyset. \end{aligned}$$

So, a configuration of the form $a, a + p_1(n), \dots, a + p_r(n)$ is contained in E only if $\mathcal{O}(\widehat{P}, \Delta_{X^{r+1}}) \cap A^{r+1} \neq \emptyset$.

On the other hand, let $\mu_{\Delta_{X^{r+1}}}$ be the measure on $\Delta_{X^{r+1}}$ induced by the measure μ on X , that is, $\mu_{\Delta_{X^{r+1}}}(A_0 \times A_1 \times \dots \times A_r) = \mu(A_0 \cap A_1 \cap \dots \cap A_r)$, $A_i \subseteq X$. Suppose that $\tilde{\mu}_P(A_0 \times \dots \times A_r) = \lim_{N-M \rightarrow \infty} \frac{1}{N-M} \sum_{n=M+1}^N g(n) \mu_{\Delta_{X^{r+1}}}(A_0 \times \dots \times A_r)$ exists for any measurable $A_0, \dots, A_r \subseteq X$ (it does if P is a system of polynomials – see [HK2] and [L1]); then $\tilde{\mu}_P$ is a probability measure on X^{r+1} supported by the topological closure $\overline{\mathcal{O}(\widehat{P}, \Delta_{X^{r+1}})}$ of $\mathcal{O}(\widehat{P}, \Delta_{X^{r+1}})$. Now assume that the set A^{r+1} “has substantial intersection with $\overline{\mathcal{O}(\widehat{P}, \Delta_{X^{r+1}})}$ ”, namely, suppose that $\tilde{\mu}_P(A^{r+1}) = \delta > 0$. This means that $\lim_{N-M \rightarrow \infty} \frac{1}{N-M} \sum_{n=M+1}^N \mu_{\Delta_{X^{r+1}}}(g(n)^{-1} A^{r+1}) = \delta$, or equivalently,

$$\lim_{N-M \rightarrow \infty} \frac{1}{N-M} \sum_{n=M+1}^N \mu(A \cap T^{-p_1(n)} A \cap \dots \cap T^{-p_r(n)} A) = \delta.$$

For $S \subseteq \mathbb{Z}$ let $d(S)$ denote, if it exists, the density of S , $d(S) = \lim_{N \rightarrow \infty} \frac{1}{N} |S \cap \{1, \dots, N\}|$. Choose our point x_0 so that for any set B of the form $B = T^{c_1} A \cap \dots \cap T^{c_l} A$, $c_1, \dots, c_l \in \mathbb{Z}$, one has $d(\{a : T^a x_0 \in B\}) = \mu(B)$. (This is always possible when T is ergodic by the ergodic theorem.) Then for every $n \in \mathbb{N}$, $d(E_n) = \mu(A \cap T^{-p_1(n)} A \cap \dots \cap T^{-p_r(n)} A)$ and hence $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N d(E_n) = \delta$. So, not only E contains configurations of the form $\{a, a + p_1(n), \dots, a + p_r(n)\}$, but contains many such configurations, and they occur in E quite regularly.

This suggests how one can attempt to construct a set $E \subset \mathbb{Z}$ which contains no configurations $\{a, a + p_1(n), \dots, a + p_r(n)\}$ and many configurations $\{a, a + q_1(n), \dots, a + q_s(n)\}$ for another system $Q = \{q_1, \dots, q_s\}$ of integral polynomials: it suffices to find a dynamical system (X, T) and a set $A \subset X$ such that $\tilde{\mu}_Q(A^{s+1}) > 0$ whereas $\tilde{\mu}_P(A^{r+1}) = 0$. Then one chooses a “typical” point x_0 of A and defines E to be the set of return times of x_0 to A , $E = E(A, x_0)$.

2. Span and Rank of a system of polynomials

Let us first introduce some linear algebra notation. Given a system of vectors $\begin{pmatrix} a_{1,1} \\ \vdots \\ a_{r,1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1,k} \\ \vdots \\ a_{r,k} \end{pmatrix} \in \mathbb{R}^r$, we define $\text{span} \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,k} \\ \vdots & \vdots & & \vdots \\ a_{r,1} & a_{r,2} & \dots & a_{r,k} \end{pmatrix}$ to be the subspace of $\mathbb{R}^r$ spanned by these vectors, and, as usual,

$$\text{rank} \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,k} \\ \vdots & \vdots & & \vdots \\ a_{r,1} & a_{r,2} & \dots & a_{r,k} \end{pmatrix} = \dim \text{span} \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,k} \\ \vdots & \vdots & & \vdots \\ a_{r,1} & a_{r,2} & \dots & a_{r,k} \end{pmatrix}.$$

Given r real-valued polynomials $q_1, \dots, q_r$ with zero constant term, we will denote by $\text{Span} \begin{pmatrix} q_1 \\ \vdots \\ q_r \end{pmatrix}$ the subspace of $\mathbb{R}^r$ spanned by the range of $\begin{pmatrix} q_1 \\ \vdots \\ q_r \end{pmatrix}$:

$$\text{Span} \begin{pmatrix} q_1 \\ \vdots \\ q_r \end{pmatrix} = \text{span} \left\{ \begin{pmatrix} q_1(x) \\ \vdots \\ q_r(x) \end{pmatrix}, x \in \mathbb{R} \right\}.$$

For polynomials $q_{i,j}$, $i = 1, \dots, r$, $j = 1, \dots, k$, with zero constant term, we define

$$\text{Span} \begin{pmatrix} q_{1,1} & q_{1,2} & \dots & q_{1,k} \\ \vdots & \vdots & & \vdots \\ q_{r,1} & q_{r,2} & \dots & q_{r,k} \end{pmatrix} = \text{Span} \begin{pmatrix} q_{1,1} \\ \vdots \\ q_{r,1} \end{pmatrix} + \text{Span} \begin{pmatrix} q_{1,2} \\ \vdots \\ q_{r,2} \end{pmatrix} + \dots + \text{Span} \begin{pmatrix} q_{1,k} \\ \vdots \\ q_{r,k} \end{pmatrix}$$

and

$$\text{Rank} \begin{pmatrix} q_{1,1} & q_{1,2} & \dots & q_{1,k} \\ \vdots & \vdots & & \vdots \\ q_{r,1} & q_{r,2} & \dots & q_{r,k} \end{pmatrix} = \dim \text{Span} \begin{pmatrix} q_{1,1} & q_{1,2} & \dots & q_{1,k} \\ \vdots & \vdots & & \vdots \\ q_{r,1} & q_{r,2} & \dots & q_{r,k} \end{pmatrix}.$$

Clearly, we have

Lemma 2.1. *The Span and Rank of a polynomial matrix $\begin{pmatrix} q_{1,1} & q_{1,2} & \dots & q_{1,k} \\ \vdots & \vdots & & \vdots \\ q_{r,1} & q_{r,2} & \dots & q_{r,k} \end{pmatrix}$ are invariant under column operations (that is, multiplying a column by a nonzero constant or a nonzero polynomial, and adding one column to another) on the matrix.*

The following lemmas will be utilized below for finding the Span of a polynomial matrix:

Lemma 2.2. *Let $q_{i,j}$ be polynomials with zero constant term and $\deg q_{i,j} \leq d$ for all i, j . Then*

(i) *Span $\begin{pmatrix} q_{1,1} & q_{1,2} & \dots & q_{1,k} \\ \vdots & \vdots & & \vdots \\ q_{r,1} & q_{r,2} & \dots & q_{r,k} \end{pmatrix}$ is spanned by the values of $\begin{pmatrix} q_{1,1}(x) & q_{1,2}(x) & \dots & q_{1,k}(x) \\ \vdots & \vdots & & \vdots \\ q_{r,1}(x) & q_{r,2}(x) & \dots & q_{r,k}(x) \end{pmatrix}$ at any distinct nonzero $x_1, \dots, x_d \in \mathbb{R}$:*

$$\text{Span} \begin{pmatrix} q_{1,1} & q_{1,2} & \dots & q_{1,k} \\ \vdots & \vdots & & \vdots \\ q_{r,1} & q_{r,2} & \dots & q_{r,k} \end{pmatrix} = \text{span} \begin{pmatrix} q_{1,1}(x_1) & \dots & q_{1,k}(x_1) & q_{1,1}(x_2) & \dots & q_{1,k}(x_2) & \dots & q_{1,1}(x_d) & \dots & q_{1,k}(x_d) \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ q_{r,1}(x_1) & \dots & q_{r,k}(x_1) & q_{r,1}(x_2) & \dots & q_{r,k}(x_2) & \dots & q_{r,1}(x_d) & \dots & q_{r,k}(x_d) \end{pmatrix}.$$

(ii) *Span $\begin{pmatrix} q_{1,1} & q_{1,2} & \dots & q_{1,k} \\ \vdots & \vdots & & \vdots \\ q_{r,1} & q_{r,2} & \dots & q_{r,k} \end{pmatrix}$ is spanned by the coefficients of $\begin{pmatrix} q_{1,1} & q_{1,2} & \dots & q_{1,k} \\ \vdots & \vdots & & \vdots \\ q_{r,1} & q_{r,2} & \dots & q_{r,k} \end{pmatrix}$: if $q_{i,j}(x) = c_{i,j,1}x + \dots + c_{i,j,d}x^d$, $c_{i,j,l} \in \mathbb{R}$, then*

$$\text{Span} \begin{pmatrix} q_{1,1} & q_{1,2} & \dots & q_{1,k} \\ \vdots & \vdots & & \vdots \\ q_{r,1} & q_{r,2} & \dots & q_{r,k} \end{pmatrix} = \text{span} \begin{pmatrix} c_{1,1,1} & \dots & c_{1,1,d} & c_{1,2,1} & \dots & c_{1,2,d} & \dots & c_{1,k,1} & \dots & c_{1,k,d} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ c_{r,1,1} & \dots & c_{r,1,d} & c_{r,2,1} & \dots & c_{r,2,d} & \dots & c_{r,k,1} & \dots & c_{r,k,d} \end{pmatrix}.$$

(iii) *More generally, let $h_1, \dots, h_l$ be linearly independent polynomials without constant term such that $q_{i,j}$ are their linear combinations: $q_{i,j} = b_{i,j,1}h_1 + \dots + b_{i,j,l}h_l$, $b_{i,j,t} \in \mathbb{R}$. Then*

$$\text{Span} \begin{pmatrix} q_{1,1} & q_{1,2} & \dots & q_{1,k} \\ \vdots & \vdots & & \vdots \\ q_{r,1} & q_{r,2} & \dots & q_{r,k} \end{pmatrix} = \text{span} \begin{pmatrix} b_{1,1,1} & \dots & b_{1,1,l} & b_{1,2,1} & \dots & b_{1,2,l} & \dots & b_{1,k,1} & \dots & b_{1,k,l} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ b_{r,1,1} & \dots & b_{r,1,l} & b_{r,2,1} & \dots & b_{r,2,l} & \dots & b_{r,k,1} & \dots & b_{r,k,l} \end{pmatrix}.$$

Proof. (i) follows from Lagrange's interpolation formula for polynomials of degree $\leq d$.

(ii) is a special case of (iii).

Here is the proof of (iii): for $j = 1, \dots, k$ define $B_j = \begin{pmatrix} b_{1,j,1} & \dots & b_{1,j,l} \\ \vdots & & \vdots \\ b_{r,j,1} & \dots & b_{r,j,l} \end{pmatrix}$. Then

$$\begin{aligned} \text{Span} \begin{pmatrix} q_{1,j} \\ \vdots \\ q_{r,j} \end{pmatrix} &= \text{span} \left\{ \begin{pmatrix} q_{1,j}(x) \\ \vdots \\ q_{r,j}(x) \end{pmatrix}, x \in \mathbb{R} \right\} = \text{span} \left\{ B_j \begin{pmatrix} h_1(x) \\ \vdots \\ h_l(x) \end{pmatrix}, x \in \mathbb{R} \right\} \\ &= B_j \text{span} \left\{ \begin{pmatrix} h_1(x) \\ \vdots \\ h_l(x) \end{pmatrix}, x \in \mathbb{R} \right\} = B_j \mathbb{R}^l = \text{span } B_j. \end{aligned}$$

So,

$$\text{Span} \begin{pmatrix} q_{1,1} & q_{1,2} & \cdots & q_{1,k} \\ \vdots & \vdots & & \vdots \\ q_{r,1} & q_{r,2} & \cdots & q_{r,k} \end{pmatrix} = \sum_{j=1}^k \text{Span} \begin{pmatrix} q_{1,1} \\ \vdots \\ q_{r,1} \end{pmatrix} = \sum_{j=1}^k \text{span } B_j.$$

■

As a corollary, we get:

Lemma 2.3. *Let $q_{i,j}$ be polynomials with zero constant term, let h be a nonconstant polynomial, and let $\hat{q}_{i,j}(x) = q_{i,j}(h(x)) - q_{i,j}(h(0))$. Then $\text{Span} \begin{pmatrix} \hat{q}_{1,1} & \cdots & \hat{q}_{1,k} \\ \vdots & & \vdots \\ \hat{q}_{r,1} & \cdots & \hat{q}_{r,k} \end{pmatrix} = \text{Span} \begin{pmatrix} q_{1,1} & \cdots & q_{1,k} \\ \vdots & & \vdots \\ q_{r,1} & \cdots & q_{r,k} \end{pmatrix}$.*

Proof. It is enough to check this for $k = 1$, that is, to show that, for integral polynomials q_i with zero constant term, $\text{Span} \begin{pmatrix} \hat{q}_1 \\ \vdots \\ \hat{q}_r \end{pmatrix} = \text{Span} \begin{pmatrix} q_1 \\ \vdots \\ q_r \end{pmatrix}$. It follows from Lemma 2.2(i) that

$$\text{Span} \begin{pmatrix} \hat{q}_1 \\ \vdots \\ \hat{q}_r \end{pmatrix} = \text{span} \left\{ \begin{pmatrix} q_1(x+h(0)) - q_1(h(0)) \\ \vdots \\ q_r(x+h(0)) - q_r(h(0)) \end{pmatrix}, x \in \mathbb{R} \right\}, \text{ and we have}$$

$$\begin{aligned} \text{span} \left\{ \begin{pmatrix} q_1(x+h(0)) - q_1(h(0)) \\ \vdots \\ q_r(x+h(0)) - q_r(h(0)) \end{pmatrix}, x \in \mathbb{R} \right\} &= \text{span} \left\{ \begin{pmatrix} q_1(x) - q_1(h(0)) \\ \vdots \\ q_r(x) - q_r(h(0)) \end{pmatrix}, x \in \mathbb{R} \right\} \\ &= \text{span} \left\{ \begin{pmatrix} q_1(x) - q_1(0) \\ \vdots \\ q_r(x) - q_r(0) \end{pmatrix}, x \in \mathbb{R} \right\} = \text{span} \left\{ \begin{pmatrix} q_1(x) \\ \vdots \\ q_r(x) \end{pmatrix}, x \in \mathbb{R} \right\} = \text{Span} \begin{pmatrix} q_1 \\ \vdots \\ q_r \end{pmatrix}. \end{aligned}$$

(We used the fact that for any system V of vectors and any $v \in V$, $\text{span}(V - v) = \text{span}(V - v')$ for any other $v' \in V$.) ■

3. Weyl dynamical systems

A *Weyl system* is a dynamical system (X, T) where X is a compact commutative Lie group and T is an affine unipotent transformation of X (that is, $Tx = \varphi(x) + \alpha$ where $\alpha \in X$ and φ is an automorphism of X satisfying $(\varphi - \text{Id}_X)^d = 0$ for some d). A natural example of a Weyl system is given by the system (X, T) where $X = \mathbb{T}^l \times Z$, $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and Z is a finite Abelian group, and

$$T(x_1, \dots, x_l, z) = \left(x_1 + \alpha_1, x_2 + m_{2,1}x_1 + \alpha_2, \dots, x_l + \sum_{i=1}^{l-1} m_{l,i}x_i + \alpha_l, z + w \right),$$

$\alpha_i \in \mathbb{T}$, $m_{j,i} \in \mathbb{Z}$ and $w \in Z$. When X is connected (and so, is a torus) we will say that (X, T) is a *connected Weyl system*.

Any closed subgroup M of X is, topologically, either a torus or a union of finitely many tori, and is a single torus if connected. We will call a *subtorus* of X any translate $x + M$ of a closed connected subgroup M of X .

Clearly, the product of several Weyl systems is a Weyl system.

A Weyl system possesses a sequence of *natural* factors: for $k = 1, \dots, d$ let $L_k = \ker(\varphi - \text{Id}_X)^{d-k+1}$ and $X_{k-1} = X/L_k$; then the projection maps $\pi_k: X \rightarrow X_k$, $k =$

$0, \dots, d-1$, commute with the action of T . The natural factors X_k , $k = 1, \dots, d$, are the *HKZ factors* for the system (X, T) . Indeed, a Weyl system (X, T) can be viewed as a nilsystem: if we define a group G as the extension of the torus X by the group of automorphisms of X generated by φ , then G is a nilpotent (nonconnected) Lie group, and X is a homogeneous space of G with $T = \varphi + \alpha$ being a translation on X . Under this interpretation, the sequence $G = \langle L_1, \varphi \rangle \supseteq L_2 \supseteq \dots \supseteq L_d$ is the lower central series of G . It is proved in [Z], Lemma 6.58 that the k -th natural factor $X_k = X/L_{k+1}$ is the characteristic factor of X for all systems of k linear integral polynomials, that is, the $(k+1)$ -st HKZ factor. (See [L2].)

A Weyl system (X, T) has good ergodic properties: if T is ergodic, then the orbit of every point is uniformly distributed in X ; if T is not ergodic, then the closure of the orbit of any point is a coset of a closed subgroup of X , and the orbit is uniformly distributed in this coset. Moreover, “polynomial orbits” of points, and even of subtori of X , also possess an analogous property.

Let us be more precise. Recall that for a set $E \subseteq \mathbb{Z}$ we write $UD(E) = \alpha$ if the limit $\lim_{N \rightarrow \infty} \frac{|E \cap \Phi_N|}{|\Phi_N|}$ exists and equals α for every Følner sequence $\{\Phi_N\}$ in $\mathbb{Z}$, and for a sequence $\{\beta_n\}_{n \in \mathbb{Z}}$ of real numbers we write $UC\text{-}\lim_n \beta_n = \beta$ if the limit $\lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{n \in \Phi_N} \beta_n$ exists and equals β for every Følner sequence $\{\Phi_N\}$ in $\mathbb{Z}$. For a torus (or, more generally, a compact commutative Lie group) M we will denote by μ_M the normalized Haar measure on M ; if M is a closed subgroup of X and $x \in X$, we will denote by μ_{x+M} the image of μ_M by the translation by x . We will say that a sequence $\{x_n\}_{n \in \mathbb{Z}}$ (in a, potentially, “larger” space $X \supseteq M$) is *well distributed in M* if $x_n \in M$ for all n , and for any continuous function f on M one has $UC\text{-}\lim_n \int_M f d\mu_M = \int_M f d\mu_M$. More generally, we will say that a sequence of tori $\{D_n\}_{n \in \mathbb{Z}}$ is *well distributed in M* if $D_n \subseteq M$ for all n , and for any continuous function f on M one has $UC\text{-}\lim_n \int_{D_n} f d\mu_{D_n} = \int_M f d\mu_M$.

Let $P = \{p_1, \dots, p_r\}$ be a system of integral polynomials with zero constant term. For a point $y \in X^r$ we will denote by $\mathcal{O}(P, y)$ the orbit of y under the action of P , that is, $\mathcal{O}(P, y) = \left\{ \begin{pmatrix} T^{p_1(n)} \\ \vdots \\ T^{p_r(n)} \end{pmatrix} y \right\}_{n \in \mathbb{Z}}$. For a torus $D \subseteq X^r$ we will denote by $\mathcal{O}(P, D)$ the orbit

of D under the action of P , that is, $\mathcal{O}(P, D) = \bigcup_{n \in \mathbb{Z}} \begin{pmatrix} T^{p_1(n)} \\ \vdots \\ T^{p_r(n)} \end{pmatrix} D$. $\overline{\mathcal{O}(P, D)}$ will stand for the topological closure of $\mathcal{O}(P, D)$. We will also denote by $\hat{P}$ the “extended system” $\{0, p_1, \dots, p_r\}$ and by $\mathcal{O}(\hat{P}, y)$ and $\mathcal{O}(\hat{P}, D)$ the orbits $\left\{ \begin{pmatrix} \text{Id}_X \\ T^{p_1(n)} \\ \vdots \\ T^{p_r(n)} \end{pmatrix} y \right\}_{n \in \mathbb{Z}}$ and, respectively,

$$\bigcup_{n \in \mathbb{Z}} \begin{pmatrix} \text{Id}_X \\ T^{p_1(n)} \\ \vdots \\ T^{p_r(n)} \end{pmatrix} D.$$

For a vector $u \in \mathbb{R}^l$ we will denote by $u \bmod 1$ the image of u in $\mathbb{T}^l = \mathbb{R}^l / \mathbb{Z}^l$. If N is a rational subspace of $\mathbb{R}^l$ (that is, a subspace defined by linear equations with rational coefficients), then $N \bmod 1$ is a subtorus of $\mathbb{T}^l$.

The following fundamental fact is a direct consequence of the classical Weyl’s work on uniform distribution ([We]) (modulo the easy modification that upgrades Weyl’s results

from uniform to well distribution).

Theorem 3.1. *Let $q_1, \dots, q_l$ be real-valued polynomials with zero constant term. The closure of the sequence $\left\{ \begin{pmatrix} q_1(n) \\ \vdots \\ q_l(n) \end{pmatrix} \bmod 1 \right\}_{n \in \mathbb{Z}}$ is a disjoint union of finitely many subtori of $\mathbb{T}^l$. If this sequence is dense in a single subtorus (which may coincide with $\mathbb{T}^l$), then it is well distributed in this subtorus.*

Let (X, T) be a connected Weyl system and let $P = \{p_1, \dots, p_r\}$ be a system of integral polynomials. It is easy to see that the orbit $\mathcal{O}(P, y)$ of a point $y \in X^r$ under the action of P is described by a polynomial formula, that is, in coordinates on the torus X^r , $\begin{pmatrix} T^{p_1(n)} \\ \vdots \\ T^{p_r(n)} \end{pmatrix} y = (q_1(n), \dots, q_l(n))$, $n \in \mathbb{Z}$, where q_i are polynomials and $l = r \dim X$. Hence, it follows from Theorem 3.1 that $M = \overline{\mathcal{O}(P, y)}$ is a subtorus of X^r or a disjoint union of finitely many subtori of X^r , and when M is a single subtorus, the sequence $\left\{ \begin{pmatrix} T^{p_1(n)} \\ \vdots \\ T^{p_r(n)} \end{pmatrix} y \right\}_{n \in \mathbb{Z}}$ is well distributed in M .

Assume now that (X, T) is a disconnected Weyl system having c components. Replacing T by T^c we arrive at a system which is a union of finitely many connected Weyl systems and our argument, applied to each of these systems, shows that $M = \overline{\mathcal{O}(P, y)}$ is still a finite union of subtori of X^r .

We will need the following more general fact:

Proposition 3.2. *Let (X, T) be a Weyl system and let $P = \{p_1, \dots, p_r\}$ be a system of integral polynomials. For any subtorus D of X^r the closure $M = \overline{\mathcal{O}(P, D)}$ of the orbit of D under the action of P is a subtorus of X^r or a union of finitely many subtori of X^r . If M is a single subtorus, the sequence $D_n = \left\{ \begin{pmatrix} T^{p_1(n)} \\ \vdots \\ T^{p_r(n)} \end{pmatrix} D \right\}_{n \in \mathbb{Z}}$ is well distributed in M .*

Proposition 3.2 can be deduced from the multiparameter version of Theorem 3.1 (which is also essentially contained in [We]). Indeed, choose an element $y \in D$ such that the sequence $\{my\}_{m \in \mathbb{Z}}$ is dense (and so, well distributed) in D . Then M is equal to the closure of the two-parameter sequence $y_{m,n} = \begin{pmatrix} T^{p_1(n)} \\ \vdots \\ T^{p_r(n)} \end{pmatrix} my$, $n, m \in \mathbb{Z}$, which, by the two-parameter version of Theorem 3.1, is a finite union of subtori of X^r . If M is a single torus, $\{y_{m,n}\}_{n,m \in \mathbb{Z}}$ is well distributed in M , that is, for any continuous function f on M and any Følner sequence $\{\Psi_N\}$ in $\mathbb{Z}^2$ one has $\lim_{N \rightarrow \infty} \frac{1}{|\Psi_N|} \sum_{(m,n) \in \Psi_N} f(y_{m,n}) = \int_M f d\mu_M$. On the other hand, for each fixed n the sequence $\{y_{m,n}\}_{m \in \mathbb{Z}}$ is well distributed in D_n . Consider a Følner sequence $\{\Phi_N\}$ in $\mathbb{Z}$ and a sequence $\{K_N\}$ of integers that increases fast enough. We then have

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{|\Phi_N| K_N} \sum_{\substack{n \in \Phi_N \\ 1 \leq m \leq K_N}} f(y_{m,n}) &= \lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{n \in \Phi_N} \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{m=1}^K f(y_{m,n}) \\ &= \lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{n \in \Phi_N} \int_{D_n} f d\mu_{D_n}. \end{aligned}$$

Since $\Phi_N \times \{1, \dots, K_N\}$ is a Følner sequence in $\mathbb{Z}^2$, we obtain $\lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{n \in \Phi_N} \int_{D_n} f d\mu_{D_n} = \int_M f d\mu_M$.

Here is another corollary of Theorem 3.1:

Corollary 3.3. *Let $\alpha_1, \dots, \alpha_k$ be rationally independent elements of $\mathbb{T}$ and let $q_{i,j}$, $i = 1, \dots, k$, $j = 1, \dots, l$, be integral polynomials with zero constant term. The sequence $\left\{ \begin{pmatrix} q_{1,1}(n)\alpha_1 + \dots + q_{k,1}(n)\alpha_k \\ \vdots \\ q_{1,l}(n)\alpha_1 + \dots + q_{k,l}(n)\alpha_k \end{pmatrix} \right\}_{n \in \mathbb{Z}}$ is well distributed in the subtorus $\text{Span} \begin{pmatrix} q_{1,1} & \dots & q_{k,1} \\ \vdots & & \vdots \\ q_{1,l} & \dots & q_{k,l} \end{pmatrix} \bmod 1$ of $\mathbb{T}^l$.*

Proof. Let $L = \text{Span} \begin{pmatrix} q_{1,1} & \dots & q_{k,1} \\ \vdots & & \vdots \\ q_{1,l} & \dots & q_{k,l} \end{pmatrix}$ and let $M = L \bmod 1$. (Note that since L is a “rational” subspace of $\mathbb{R}^l$ (that is, is spanned by vectors with rational coordinates), M is a subtorus.) For any i and any $n \in \mathbb{Z}$, $\begin{pmatrix} q_{i,1}(n) \\ \vdots \\ q_{i,l}(n) \end{pmatrix} \in L$, so $\begin{pmatrix} q_{1,1}(n)\alpha_1 + \dots + q_{k,1}(n)\alpha_k \\ \vdots \\ q_{1,l}(n)\alpha_1 + \dots + q_{k,l}(n)\alpha_k \end{pmatrix} \in M$ for all n .

Assume that χ is a character on $\mathbb{T}^l$ (with values in the additive torus $\mathbb{T}$), $\chi(x_1, \dots, x_l) = \sum_{j=1}^l a_j x_j$ with $a_1, \dots, a_l \in \mathbb{Z}$, such that $\chi \begin{pmatrix} q_{1,1}(n)\alpha_1 + \dots + q_{k,1}(n)\alpha_k \\ \vdots \\ q_{1,l}(n)\alpha_1 + \dots + q_{k,l}(n)\alpha_k \end{pmatrix} = 0$ for all n . Then, in $\mathbb{T}$, we have

$$\left(\sum_{j=1}^l a_j q_{1,j}(n) \right) \alpha_1 + \dots + \left(\sum_{j=1}^l a_j q_{k,j}(n) \right) \alpha_k = 0$$

for all n , and since $\alpha_1, \dots, \alpha_k$ are rationally independent in $\mathbb{T}$, we deduce that $\sum_{j=1}^l a_j q_{i,j}(n) = 0$ for all n and for all $i = 1, \dots, k$. This implies that $\sum_{j=1}^l a_j q_{i,j} = 0$ for all $i = 1, \dots, k$, and thus $\chi|_M = 0$.

We have established that the sequence $\left\{ \begin{pmatrix} q_{1,1}(n)\alpha_1 + \dots + q_{k,1}(n)\alpha_k \\ \vdots \\ q_{1,l}(n)\alpha_1 + \dots + q_{k,l}(n)\alpha_k \end{pmatrix} \right\}_{n \in \mathbb{Z}}$ is not contained in a proper subgroup of M ; by Theorem 3.1, it is well distributed in M . ■

To achieve the goals formulated in the introduction, it will be sufficient to deal with Weyl systems of a special form. A *standard Weyl system (of depth d)* is the system (X, T) with $X = \mathbb{T}^d$ and

$$T(x_1, x_2, \dots, x_d) = (x_1 + \alpha, x_2 + x_1, x_3 + x_2, \dots, x_d + x_{d-1}), \quad x = (x_1, \dots, x_d) \in X,$$

where $\alpha \in \mathbb{T}$ is irrational. By [F2] Proposition 3.11, (X, T) is ergodic. A *quasi-standard Weyl system* of depth d is a system (X, T) where $X = \mathbb{T}^d$ and

$$T(x_1, \dots, x_d) = \left(x_1 + \alpha_1, x_2 + m_{2,1}x_1 + \alpha_2, \dots, x_d + \sum_{i=1}^{d-1} m_{d,i}x_i + \alpha_d \right),$$

with $\alpha_i \in \mathbb{T}$, $m_{j,i} \in \mathbb{Z}$, α_1 is irrational and $m_{j,j-1} \neq 0$ for all $j = 2, \dots, d$. A quasi-standard Weyl system is also ergodic. For a standard or a quasi-standard Weyl system the natural factors X_k have form $X_k = X/L_{k+1}$ where $L_{k+1} = \{(0, \dots, 0, x_{k+1}, \dots, x_d), x_i \in \mathbb{T}\}$.

Lemma 3.4. *Any quasi-standard Weyl system is a factor, $\eta: \tilde{X} \rightarrow X$, of a standard Weyl system $(\tilde{X}, \tilde{T})$ of the same depth; η has finite fibers and commutes with the projections π_k , $k = 0, \dots, d-1$, onto the natural factors.*

Rather than formally proving this lemma we illustrate it on a simple example, which indicates how the general result can be proved. Consider the quasi-standard Weyl system $(\mathbb{T}^3, T)$ where

$$T(x, y, z) = (x + \alpha_1, y + 2x + \alpha_2, z + 4x + 3y + \alpha_3).$$

This system is the factor of the standard Weyl system $(\mathbb{T}^3, \tilde{T})$

$$\tilde{T}(x, y, z) = (x + \alpha_1, y + x, z + y)$$

via the factor-map

$$(x, y, z) \mapsto (x + a, 2y + x + b, 6z + 7y + x)$$

with $a, b \in \mathbb{T}$ satisfying $2a = \alpha_1 - \alpha_2$ and $3b = -\alpha_1 + 2\alpha_2 - \alpha_3$.

We would like to remark that while we find it convenient to work with quasi-standard Weyl systems, all the combinatorial results in this paper could be obtained by employing the standard Weyl systems only. We would also like to mention the following fact, which demonstrates a universal property of standard Weyl systems:

Theorem 3.5. ([FrK]) *Any ergodic connected Weyl system is a factor of a product of several standard Weyl systems.*

Disconnected Weyl systems do not provide much novelty: any disconnected ergodic Weyl system (Y, R) is a union $Y = X^{(1)} \cup \dots \cup X^{(m)}$ of $m \geq 2$ isomorphic tori; R cyclically permutes these tori and, for each $i = 1, \dots, m$, $(X^{(i)}, R^m|_{X^{(i)}})$ is a connected Weyl system.

4. Weyl and Vandermonde complexities

We fix a standard Weyl system (X, T) , where $X = \mathbb{T}^d$ and

$$T(x_1, x_2, \dots, x_d) = (x_1 + \alpha, x_2 + x_1, x_3 + x_2, \dots, x_d + x_{d-1}), \quad x = (x_1, \dots, x_d) \in X,$$

with an irrational $\alpha \in \mathbb{T}$; we will always assume that d is large enough so that its value will not affect our further computations. (Namely, d will have to be larger than the complexity of the polynomial systems under consideration.) For $k = 0, \dots, d$ we define subtori $L_k = \{(0, \dots, 0, x_k, \dots, x_d), x_i \in \mathbb{T}\} \subset \mathbb{T}^d$, $F_k = \{(0, \dots, 0, x_k, 0, \dots, 0), x_k \in \mathbb{T}\} \subset \mathbb{T}^d$, a factor torus $X_k = X/L_{k+1}$, and let $\pi_k: X \rightarrow X_k$ be the projection map. We will also denote by F_k the image of F_k under π_k , so that if we identify X_k with $\mathbb{T}^k$, we identify F_k with the subgroup $F_k = \{(0, \dots, 0, x_k), x_k \in \mathbb{T}\}$ of $\mathbb{T}^k$.

Let now $P = \{p_1, \dots, p_r\}$ be a system of integral polynomials with zero constant term. (We will always assume that the polynomials $p_1, \dots, p_r$ are all distinct.) One checks by induction that for any $n \in \mathbb{Z}$,

$$T^n(x_1, x_2, \dots, x_d) = (x_1 + n\alpha, x_2 + nx_1 + \binom{n}{2}\alpha, \dots, x_d + \sum_{i=1}^{d-1} \binom{n}{i}x_{d-i} + \binom{n}{d}\alpha).$$

For a polynomial p and $k \in \mathbb{N}$ we will write $p^{[k]}$ for $\binom{p}{k} = \frac{1}{k!}p(p-1)\dots(p-k+1)$, and $p^{[0]} = 1$. Let $g(n) = \begin{pmatrix} \text{Id}_X \\ T^{p_1(n)} \\ \vdots \\ T^{p_r(n)} \end{pmatrix}$, $n \in \mathbb{Z}$. For $x = (x_1, x_2, \dots, x_d) \in X$, the orbit $\mathcal{O}(\hat{P}, \bar{x}) = \{g(n)\bar{x}\}_{n \in \mathbb{Z}} \subseteq X^{r+1}$ of the point $\bar{x} = \begin{pmatrix} x \\ x \\ \vdots \\ x \end{pmatrix}$ under the action of $\hat{P} = \{0, p_1, \dots, p_r\}$ is

$$\left\{ \begin{pmatrix} x_1, & x_2, & \dots, & x_d \\ x_1 + p_1(n)\alpha, & x_2 + p_1(n)x_1 + p_1(n)^{[2]}\alpha, & \dots, & x_d + \sum_{i=1}^{d-1} p_1(n)^{[i]}x_{d-i} + p_1(n)^{[d]}\alpha \\ \vdots & \vdots & & \vdots \\ x_1 + p_r(n)\alpha, & x_2 + p_r(n)x_1 + p_r(n)^{[2]}\alpha, & \dots, & x_d + \sum_{i=1}^{d-1} p_r(n)^{[i]}x_{d-i} + p_r(n)^{[d]}\alpha \end{pmatrix} \right\}_{n \in \mathbb{Z}} \quad (4.1)$$

By Corollary 3.3, when $x_1, \dots, x_d$ and α are rationally independent in $\mathbb{T}$, $\mathcal{O}(\hat{P}, \bar{x})$ is well distributed (and hence, dense) in the subtorus

$$\begin{pmatrix} x_1 \\ x_1 \\ \vdots \\ x_1 \\ x_2 \\ x_2 \\ \vdots \\ x_2 \\ \vdots \\ \vdots \\ x_{d-1} \\ x_{d-1} \\ \vdots \\ x_{d-1} \\ x_d \\ x_d \\ \vdots \\ x_d \end{pmatrix} + \text{Span} \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ p_1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ p_r & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ p_1^{[2]} & p_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ p_r^{[2]} & p_r & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 \\ p_1^{[d-1]} & p_1^{[d-2]} & p_1^{[d-3]} & \dots & p_1 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ p_r^{[d-1]} & p_r^{[d-2]} & p_r^{[d-3]} & \dots & p_r & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ p_1^{[d]} & p_1^{[d-1]} & p_1^{[d-2]} & \dots & p_1^{[2]} & p_1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ p_r^{[d]} & p_r^{[d-1]} & p_r^{[d-2]} & \dots & p_r^{[2]} & p_r \end{pmatrix} \pmod{1}$$

of $X^{r+1} \simeq (\mathbb{T}^{(r+1)})^d$ (and is contained in this subtorus if $x_1, \dots, x_d, \alpha$ are rationally dependent). The closure of the orbit $\mathcal{O}(\hat{P}, \Delta_{X^{r+1}})$ of the entire diagonal $\Delta_{X^{r+1}} = \{\bar{x}, x \in X\}$ of X^{r+1} is therefore the subtorus

$$H = \overline{\mathcal{O}(\hat{P}, \Delta_{X^{r+1}})} = \text{Span } \Theta_d \pmod{1} \subseteq X^{r+1}$$

where for each $k \in \mathbb{N}$ we define Θ_k as the $(r+1)k \times 2k$ matrix

$$\Theta_k = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 & p_1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \dots & 0 & 0 & p_r & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & p_1^{[2]} & p_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 & p_r^{[2]} & p_r & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 1 & 0 & p_1^{[k-1]} & p_1^{[k-2]} & p_1^{[k-3]} & \dots & p_1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & p_r^{[k-1]} & p_r^{[k-2]} & p_r^{[k-3]} & \dots & p_r & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & p_1^{[k]} & p_1^{[k-1]} & p_1^{[k-2]} & \dots & p_1^{[2]} & p_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & p_r^{[k]} & p_r^{[k-1]} & p_r^{[k-2]} & \dots & p_r^{[2]} & p_r \end{pmatrix}.$$

For $k \leq d$, let $H_k = \pi_k^{r+1}(H) \subseteq X_k^{r+1}$; identifying X_k with $\mathbb{T}^k$ and X_k^{r+1} with $(\mathbb{T}^{(r+1)})^k$, we get

$$H_k = \text{Span } \Theta_k \bmod 1 \subseteq X_k^{r+1}. \quad (4.2)$$

If we suppress the 1-st and the $(k+1)$ -st columns from the matrix Θ_{k+1} , we obtain the matrix $\begin{pmatrix} 0 \\ \Theta_k \end{pmatrix}$. Hence we have

$$H_{k+1} \supseteq \text{Span} \begin{pmatrix} 0 \\ \Theta_k \end{pmatrix} \bmod 1. \quad (4.3)$$

Assume now that for some $k \leq d$ one has $F_k^{r+1} \subseteq H_k$. Using formula (4.3), this implies that $F_{k+1}^{r+1} \subseteq H_{k+1}$, and of course $F_l^{r+1} \subseteq H_l$ for all $l > k$, which gives $H \supseteq L_k^{r+1}$. Let us call the minimal k with this property *the Weyl complexity* of P and denote it by $W(P)$ or $W(p_1, \dots, p_r)$.

We note that the first component X of X^{r+1} actually plays no role in determining $W(P)$. For $k \leq d$ let

$$\Lambda_k = \begin{pmatrix} p_1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_r & 0 & 0 & \dots & 0 & 0 \\ p_1^{[2]} & p_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_r^{[2]} & p_r & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_1^{[k-1]} & p_1^{[k-2]} & p_1^{[k-3]} & \dots & p_1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_r^{[k-1]} & p_r^{[k-2]} & p_r^{[k-3]} & \dots & p_r & 0 \\ p_1^{[k]} & p_1^{[k-1]} & p_1^{[k-2]} & \dots & p_1^{[2]} & p_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_r^{[k]} & p_r^{[k-1]} & p_r^{[k-2]} & \dots & p_r^{[2]} & p_r \end{pmatrix}.$$

The subtorus $M = \text{Span } \Lambda_d \bmod 1 \subseteq X^r$ is the (translated to 0) orbit of a “generic” point $\bar{x}$ of the diagonal Δ_{X^r} under the action of P : $M = \overline{\mathcal{O}(P, \bar{x})} - \bar{x}$ where $\bar{x} = \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix}$ and x is a point of X whose coordinates and α are rationally independent. For $k \leq d$ let

$$M_k = \pi_k^r(M) = \text{Span } \Lambda_k \bmod 1 \subseteq X_k^r.$$

Then $H_k = \Delta_{X_k^{r+1}} \oplus (\{0\} \times M_k) \subseteq X_k \times X_k^r$. It follows that $F_k^{r+1} \subseteq H_k$ iff $F_k^r \subseteq M_k$. Put $w_0(P) = 0$ and for each $k \in \mathbb{N}$ let

$$w_k(P) = \dim M_k = \text{Rank } \Lambda_k.$$

Since $w_k(P) = w_{k-1}(P) + \dim(M_k \cap F_k^r)$, we have $F_k^r \subseteq M_k$ iff $w_k(P) - w_{k-1}(P) = r$. We obtain:

Proposition 4.1. *The Weyl complexity $W(P)$ equals the minimal k for which $M_k \supseteq F_k^r$, and the minimal k for which $w_k(P) - w_{k-1}(P) = r$.*

We will now formally introduce the Vandermonde complexity for our system P . Let τ_k be the projection of X to F_k ; we define the *Vandermonde complexity* of P , $V(P)$ or $V(p_1, \dots, p_r)$, as the minimal k for which $\tau_k^{r+1}(H) = F_k^{r+1}$. Identifying the subtorus F_k^{r+1} of X^{r+1} with $\mathbb{T}^{r+1}$ we have

$$\tau_k^{r+1}(H_k) = \text{Span} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & p_1^{[k]} & p_1^{[k-1]} & \dots & p_1^{[2]} & p_1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 1 & p_r^{[k]} & p_r^{[k-1]} & \dots & p_r^{[2]} & p_r & 1 \end{pmatrix} \bmod 1 = \text{Span} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & p_1 & p_1^2 & \dots & p_1^{k-2} & p_1^{k-1} & p_1^k \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 1 & p_r & p_r^2 & \dots & p_r^{k-2} & p_r^{k-1} & p_r^k \end{pmatrix} \bmod 1.$$

Define $v_k(P) = \text{Rank} \begin{pmatrix} p_1 & p_1^2 & \dots & p_1^{k-2} & p_1^{k-1} & p_1^k \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ p_r & p_r^2 & \dots & p_r^{k-2} & p_r^{k-1} & p_r^k \end{pmatrix}$, $k \in \mathbb{N}$; then $\tau_k^{r+1}(H_k) = F_k^{r+1}$ iff $\tau_k^r(M_k) = F_k^r$ iff $v_k(P) = r$. We see that

Proposition 4.2. *$V(P)$ equals the minimal k for which $\tau_k^r(M_k) = F_k^r$, and the minimal k for which $v_k(P) = r$.*

5. Properties of Vandermonde and Weyl complexities and examples

We start with the Vandermonde complexity.

Lemma 5.1. *For any system $P = \{p_1, \dots, p_r\}$ of r integral polynomials with zero constant term, $V(P) \leq r$.*

Proof. We have

$$\det \begin{pmatrix} p_1 & p_1^2 & \dots & p_1^r \\ \vdots & \vdots & & \vdots \\ p_r & p_r^2 & \dots & p_r^r \end{pmatrix} = \prod_{i=1}^r (p_i \cdot \prod_{j=i+1}^r (p_j - p_i)) \neq 0.$$

(This is the Vandermonde determinant, which explains our terminology.) Thus, the vectors

$\begin{pmatrix} p_1(x) \\ \vdots \\ p_r(x) \end{pmatrix}, \dots, \begin{pmatrix} p_1^r(x) \\ \vdots \\ p_r^r(x) \end{pmatrix}$ are linearly independent for all but finitely many x , and so, $v_r(P) = r$.

Hence, $V(P) \leq r$. ■

Here are some properties of the Vandermonde complexity, which are clear from the definition and Proposition 4.2:

Proposition 5.2. *Let $\{p_1, \dots, p_r\}$ be a system of integral polynomials with zero constant term.*

- (i) *If $\{q_1, \dots, q_s\} \subseteq \{p_1, \dots, p_r\}$, then $V(q_1, \dots, q_s) \leq V(p_1, \dots, p_r)$.*
- (ii) *$V(p_1, \dots, p_r) = 1$ iff $p_1, \dots, p_r$ are linearly independent.*
- (iii) *If $p_1, \dots, p_r$ are all linear, $V(p_1, \dots, p_r) = r$.*
- (iv) *V is invariant under polynomial substitutions: for any nonzero integral polynomial h with zero constant term, $V(p_1(h(x)), \dots, p_r(h(x))) = V(p_1(x), \dots, p_r(x))$.*
- (v) *For any nonzero integer $m \neq 0$, $V(mp_1, \dots, mp_r) = V(p_1, \dots, p_r)$.*

Examples of computation of Vandermonde complexity.

Consider the system $P = \{x, 2x, x^2\}$. Using the “coefficient method” from Lemma 2.2(ii), we get

$$v_1(P) = \text{Rank} \begin{pmatrix} x \\ 2x \\ x^2 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{pmatrix} = 2$$

and

$$v_2(P) = \text{Rank} \begin{pmatrix} x & x^2 \\ 2x & 4x^2 \\ x^2 & x^4 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & 0 & 4 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} = 3,$$

so $V(P) = 2$.

For the system $P = \{x, x^2, x + x^2, x + 2x^2\}$ we have

$$v_1(P) = \text{Rank} \begin{pmatrix} x \\ x^2 \\ x+x^2 \\ x+2x^2 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} = 2$$

and

$$v_2(P) = \text{Rank} \begin{pmatrix} x & x^2 \\ x^2 & x^4 \\ x+x^2 & x^2+2x^3+x^4 \\ x+2x^2 & x^2+4x^3+4x^4 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 4 & 4 \end{pmatrix} = 4,$$

so $V(P) = 2$.

For the system $P = \{x, x^2, x + x^2, x + 2x^2, x + 3x^2\}$,

$$v_1(P) = \text{Rank} \begin{pmatrix} x \\ x^2 \\ x+x^2 \\ x+2x^2 \\ x+3x^2 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = 2,$$

$$v_2(P) = \text{Rank} \begin{pmatrix} x & x^2 \\ x^2 & x^4 \\ x+x^2 & x^2+2x^3+x^4 \\ x+2x^2 & x^2+4x^3+4x^4 \\ x+3x^2 & x^2+6x^3+9x^4 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 4 & 4 \\ 1 & 3 & 1 & 6 & 9 \end{pmatrix} = 4$$

and

$$v_3(P) = \text{Rank} \begin{pmatrix} x & x^2 & x^3 \\ x^2 & x^4 & x^6 \\ x+x^2 & x^2+2x^3+x^4 & x^3+3x^4+3x^5+x^6 \\ x+2x^2 & x^2+4x^3+4x^4 & x^3+6x^4+12x^5+8x^6 \\ x+3x^2 & x^2+6x^3+9x^4 & x^3+9x^4+27x^5+27x^6 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 2 & 1 & 1 & 3 & 3 & 1 \\ 1 & 2 & 1 & 4 & 4 & 1 & 6 & 12 & 8 \\ 1 & 3 & 1 & 6 & 9 & 1 & 9 & 27 & 27 \end{pmatrix} = 5,$$

so $V(P) = 3$.

The Weyl complexity has properties similar to those of the Vandermonde complexity:

Proposition 5.3. *Let $\{p_1, \dots, p_r\}$ be a system of integral polynomials with zero constant term.*

- (i) *If $\{q_1, \dots, q_s\} \subseteq \{p_1, \dots, p_r\}$, then $W(q_1, \dots, q_s) \leq W(p_1, \dots, p_r)$.*
- (ii) *$W(p_1, \dots, p_r) = 1$ iff $p_1, \dots, p_r$ are linearly independent.*
- (iii) *If all $p_1, \dots, p_r$ are linear, $W(p_1, \dots, p_r) = r$.*
- (iv) *For any nonzero integral polynomial h with zero constant term, $W(p_1(h(x)), \dots, p_r(h(x))) = W(p_1(x), \dots, p_r(x))$.*
- (v) *For any nonzero integer m , $W(mp_1, \dots, mp_r) = W(p_1, \dots, p_r)$.*
- (vi) *$W(p_1, \dots, p_r) \geq V(p_1, \dots, p_r)$.*

Proof. (i) is clear from the definition. (ii) and (vi) follow from Proposition 4.1. It follows from formula (4.2) and Lemma 2.2(i) that for the system $P(h) = \{p_1(h(x)), \dots, p_r(h(x))\}$ we have $\overline{\mathcal{O}(\widehat{P}(h), \Delta_{X^{r+1}})} = \overline{\mathcal{O}(\widehat{P}, \Delta_{X^{r+1}})}$; this implies (iv). We postpone the proof of (iii) and of (v) till the end of this section. ■

Remark. The fact that $W(P)$ is finite for any system of integral polynomials P is a consequence of the general study of HKZ factors; we do not prove it here. $W(P)$ may be strictly larger than $V(P)$ as an example at the end of this section demonstrates. A question that we leave open is whether $W(p_1, \dots, p_r)$ is always $\leq r$.

The definition of the Weyl complexity via the standard Weyl system is, actually, inconvenient for practical usage. We will modify it a little bit by replacing the standard Weyl system by a quasi-standard one. Let (X, T) be a quasi-standard Weyl system, let $L_k = \{(0, \dots, 0, x_k, \dots, x_d)\}$, $X_k = X/L_{k+1}$, $F_k = \{(0, \dots, 0, x_k)\} \subseteq X_k$, π_k be the projection of X to X_k and τ_k be the projection of X to F_k , $k = 1, \dots, d$; let $\Delta_{X^{r+1}} = \left\{ \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix}, x \in X \right\}$ be the diagonal of X^{r+1} , and let $\bar{x}$ be a “generic” point of the diagonal of X^r . Let $P = \{p_1, \dots, p_r\}$ be a system of integral polynomials with zero constant term and with $W(P) \leq k$, let $H = \overline{\mathcal{O}(\widehat{P}, \Delta_{X^{r+1}})} \subseteq X^{r+1}$, $M = \overline{\mathcal{O}(\widehat{P}, \bar{x})} - \bar{x} \subseteq X^r$, $H_k = \pi_k^{r+1}(H) \subseteq X_k^{r+1}$ and $M_k = \pi_k^r(M) \subseteq X_k^r$, $k = 1, \dots, d$. Let $(\tilde{X}, \tilde{T})$ be a standard Weyl system for which (X, T) is a factor, $\eta: \tilde{X} \rightarrow X$, as in Lemma 3.4, and let $\tilde{H}, \tilde{M}, \tilde{L}_k, \tilde{F}_k, \tilde{H}_k$ and $\tilde{M}_k$, $k = 1, \dots, d$, be the corresponding subtori of $\tilde{X}^{r+1}, \tilde{X}^r, \tilde{X}, \tilde{X}_k, \tilde{X}_k^{r+1}$ and $\tilde{X}_k^r$ respectively. Then for all k , $F_k = \eta(\tilde{F}_k)$, $H_k = \eta(\tilde{H}_k)$ and $M_k = \eta(\tilde{M}_k)$. So, $H \supseteq L_k^{r+1}$ iff $\tilde{H} \supseteq \tilde{L}_k^{r+1}$, $M \supseteq L_k^r$ iff $\tilde{M} \supseteq \tilde{L}_k^r$, $H_k \supseteq F_k^{r+1}$ iff $\tilde{H}_k \supseteq \tilde{F}_k^{r+1}$, and $M_k \supseteq F_k^r$ iff $\tilde{M}_k \supseteq \tilde{F}_k^r$. Since η has finite fibers, $\dim H_k = \dim \tilde{H}_k$ and $\dim M_k = \dim \tilde{M}_k$. We obtain that for computing the Weyl complexity of the system $\{p_1, \dots, p_r\}$ any quasi-standard Weyl system can be used:

Proposition 5.4. *$W(P)$ equals the minimal k for which $H \supseteq L_k^{r+1}$, the minimal k for which $M \supseteq L_k^r$, the minimal k for which $H_k \supseteq F_k^{r+1}$, and the minimal k for which $M_k \supseteq F_k^r$; $w_k(P) = \dim M_k$ for all k .*

An analogous fact holds for the Vandermonde complexity:

Proposition 5.5. *For any quasi-standard Weyl system, $v_k(P) = \dim \tau_k^r(M)$ for all k ; $V(P)$ equals the minimal k for which $\tau_k^{r+1}(H) = F_k^{r+1}$ and the minimal k for which $\tau_k^r(M) = F_k^r$.*

Proof. For any quasi-standard Weyl system (X, T) ,

$$\tau_k^{r+1}(H) = \text{Span} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & h_k(p_1) & h_{k-1}(p_1) & \dots & h_1(p_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & h_k(p_r) & h_{k-1}(p_r) & \dots & h_1(p_r) \end{pmatrix} \text{mod } 1,$$

where for each $i = 1, \dots, k$, h_i is a polynomial of degree i with zero constant term. (Under $h(p)$ we understand the polynomial $h(p(x))$.) Performing suitable column transformations of the last matrix and using Lemma 2.1, we come to $\text{Span} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & h_k(p_1) & h_{k-1}(p_1) & \dots & h_1(p_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & h_k(p_r) & h_{k-1}(p_r) & \dots & h_1(p_r) \end{pmatrix} =$

$$\text{Span} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & p_1^k & p_1^{k-1} & \dots & p_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & p_r^k & p_r^{k-1} & \dots & p_r \end{pmatrix}.$$

Similarly,

$$\tau_k^r(M) = \text{Span} \begin{pmatrix} h_k(p_1) & h_{k-1}(p_1) & \dots & h_1(p_1) \\ \vdots & \vdots & \ddots & \vdots \\ h_k(p_r) & h_{k-1}(p_r) & \dots & h_1(p_r) \end{pmatrix} \text{mod } 1 = \text{Span} \begin{pmatrix} p_1^k & p_1^{k-1} & \dots & p_1 \\ \vdots & \vdots & \ddots & \vdots \\ p_r^k & p_r^{k-1} & \dots & p_r \end{pmatrix} \text{mod } 1.$$

■

(To clarify what we mean under “suitable column transformations” let us consider an example. For the Weyl system $(\mathbb{T}^3, T)$ where

$$T(x_1, x_2, x_3) = (x_1 + \alpha, x_2 + 3x_1, x_3 + x_2 + 2x_1 + \alpha),$$

one checks that

$$T^n(x_1, x_2, x_3) = \left(x_1 + n\alpha, x_2 + 3nx_1 + \frac{3}{2}n(n-1)\alpha, x_3 + nx_2 + \frac{1}{2}(3n^2 + n)x_1 + \frac{1}{2}(n^3 - n^2 + 2n)\alpha \right),$$

$n \in \mathbb{Z}$, and so,

$$\tau_3^r(M) = \text{Span} \begin{pmatrix} \frac{1}{2}(p_1^3 - p_1^2 + 2p_1) & \frac{1}{2}(3p_1^2 + p_1) & p_1 \\ \vdots & \vdots & \vdots \\ \frac{1}{2}(p_r^3 - p_r^2 + 2p_r) & \frac{1}{2}(3p_r^2 + p_r) & p_r \end{pmatrix} \text{mod } 1 = \text{Span} \begin{pmatrix} p_1^3 & p_1^2 & p_1 \\ \vdots & \vdots & \vdots \\ p_r^3 & p_r^2 & p_r \end{pmatrix} \text{mod } 1.$$

Now we may also get:

Proof of Proposition 5.3(v). Let (X, T) be a standard Weyl system and m be a nonzero integer. Put $R = T^m$; (X, R) is then a quasi-standard Weyl system. Let $P = \{p_1, \dots, p_r\}$ be a system of integral polynomials with zero constant term and let $mP = \{mp_1, \dots, mp_r\}$. Using an index to specify what transformation we consider, we have $\mathcal{O}_R(\hat{P}, \Delta_{X^{r+1}}) = \mathcal{O}_T(m\hat{P}, \Delta_{X^{r+1}})$. Since the first orbit is responsible for $W(p_1, \dots, p_r)$ and the second orbit is responsible for $W(mp_1, \dots, mp_r)$, these two numbers coincide. ■

Consider the quasi-standard Weyl system

$$T(x_1, \dots, x_d) = \left(x_1 + \alpha, x_2 + 2x_1 + \alpha, \dots, x_d + \sum_{i=1}^{d-1} \binom{d}{i} x_i + \alpha \right).$$

For $n \in \mathbb{Z}$ one has

$$T^n(x_1, x_2, \dots, x_d) = \left(x_1 + n\alpha, x_2 + 2nx_1 + n^2\alpha, \dots, x_d + \sum_{i=1}^{d-1} \binom{d}{i} n^i x_{d-i} + n^d\alpha \right).$$

For this system,

$$M_k = \text{Span } \Lambda'_k \text{ mod } 1 \quad (5.1)$$

where

$$\Lambda'_k = \begin{pmatrix} p_1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ p_r & 0 & 0 & \dots & 0 & 0 \\ p_1^2 & 2p_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ p_r^2 & 2p_r & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ p_1^{k-1} & \binom{k-1}{k-2} p_1^{k-2} & \binom{k-1}{k-3} p_1^{k-3} & \dots & (k-1)p_1 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ p_r^{k-1} & \binom{k-1}{k-2} p_r^{k-2} & \binom{k-1}{k-3} p_r^{k-3} & \dots & (k-1)p_r & 0 \\ p_1^k & \binom{k}{k-1} p_1^{k-1} & \binom{k}{k-2} p_1^{k-2} & \dots & \binom{k}{2} p_1^2 & kp_1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ p_r^k & \binom{k}{k-1} p_r^{k-1} & \binom{k}{k-2} p_r^{k-2} & \dots & \binom{k}{2} p_r^2 & kp_r \end{pmatrix}$$

and by Proposition 5.4,

$$w_k(P) = \dim M_k = \text{Rank } \Lambda'_k \quad (5.2)$$

for all $k \in \mathbb{N}$. We will use this definition of the numbers $w_k(P)$ in our computations.

Proof of Proposition 5.3(iii). If p_i are linear, $p_i(x) = c_i x$, $i = 1, \dots, r$, formula (5.1) and Lemma 2.2 (ii) imply

$$\begin{aligned} M_k &= \text{span} \begin{pmatrix} c_1 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & & \vdots \\ c_r & 0 & \dots & 0 & 0 & \dots & 0 & \dots & \dots & 0 \\ 0 & c_1^2 & \dots & 0 & 2c_1 & \dots & 0 & \dots & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & & \vdots \\ 0 & c_r^2 & \dots & 0 & 2c_r & \dots & 0 & \dots & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & & \vdots \\ 0 & 0 & \dots & c_1^k & 0 & \dots & \binom{k}{k-1} c_1^{k-1} & \dots & \dots & kc_1 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & & \vdots \\ 0 & 0 & \dots & c_r^k & 0 & \dots & \binom{k}{k-1} c_r^{k-1} & \dots & \dots & kc_r \end{pmatrix} \text{ mod } 1 \\ &= \text{span} \begin{pmatrix} c_1 & 0 & 0 & \dots & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & & \vdots & & \vdots \\ c_r & 0 & 0 & \dots & \dots & 0 & \dots & 0 \\ 0 & c_1^2 & c_1 & \dots & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & & \vdots & & \vdots \\ 0 & c_r^2 & c_r & \dots & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \dots & c_1^k & \dots & c_1 \\ \vdots & \vdots & \vdots & & & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \dots & c_r^k & \dots & c_r \end{pmatrix} \text{ mod } 1, \end{aligned}$$

and one has $F_k^r \subseteq M_k$ iff $k \geq r$. ■

Examples of computation of Weyl complexity.

Consider the system $P = \{x, 2x, x^2\}$. Using the formula (5.2) and Lemma 2.2(ii) we get

$$\begin{aligned} w_1(P) &= \text{Rank} \begin{pmatrix} x \\ 2x \\ x^2 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{pmatrix} = 2, \\ w_2(P) &= \text{Rank} \begin{pmatrix} x & 0 \\ 2x & 0 \\ x^2 & 0 \\ x^2 & 2x \\ 4x^2 & 4x \\ x^4 & 2x^2 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 4 & 0 & 4 & 0 \\ 0 & 0 & 1 & 0 & 2 \end{pmatrix} = 4, \\ w_3(P) &= \text{Rank} \begin{pmatrix} x & 0 & 0 \\ 2x & 0 & 0 \\ x^2 & 0 & 0 \\ x^2 & 2x & 0 \\ 4x^2 & 4x & 0 \\ x^4 & 2x^2 & 0 \\ x^3 & 3x^2 & 3x \\ 8x^3 & 12x^2 & 6x \\ x^6 & 3x^4 & 3x^2 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 3 & 0 & 3 & 0 \\ 0 & 0 & 8 & 0 & 0 & 0 & 12 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 3 & 0 & 3 \end{pmatrix} = 7 = w_2 + 3, \end{aligned}$$

and thus, $W(P) = 3$. (Recall that $V(P) = 2$.)

In contrast, for the system $P = \{x, 2x, x^3\}$,

$$\begin{aligned} w_1(P) &= \text{Rank} \begin{pmatrix} x \\ 2x \\ x^3 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{pmatrix} = 2, \\ w_2(P) &= \text{Rank} \begin{pmatrix} x & 0 \\ 2x & 0 \\ x^3 & 0 \\ x^2 & 2x \\ 4x^2 & 4x \\ x^6 & 2x^3 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 4 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 & 0 & 2 \end{pmatrix} = 5, \end{aligned}$$

and therefore, $W(P) = 2$.

One can easily see that for any system P of 3 polynomials one has $W(P) \leq 3$. For example, if $a, b, c, d \in \mathbb{Z}$ with $(a, b) \neq (c, d)$, then

$V(x, ax + bx^2, cx + dx^2) = 2$ if at least one of b, d is nonzero;

$W(x, ax + bx^2, cx + dx^2) = 2$ when both b and d are nonzero, $W(x, ax, cx + dx^2) = 3$.

In a recent preprint, [Fr], N. Frantzikinakis classifies all the possibilities for the values of the Weyl complexity for a system of 3 polynomials.

6. Characteristic factors and large intersections

In this Section we investigate some dynamical properties in Weyl systems: we describe what information is given by Weyl and Vandermonde complexities of a family of polynomials on quantitative multiple recurrence properties along this family. We begin with a simple example that gives a flavor of what we develop in the sequel. Let us consider the quasi-standard Weyl system on $X = \mathbb{T}^4$,

$$T(x_1, x_2, x_3, x_4) = (x_1 + \alpha, x_2 + 2x_1 + \alpha, x_3 + 3x_2 + 3x_1 + \alpha, x_4 + 4x_3 + 6x_2 + 4x_1 + \alpha),$$

where $\alpha \in \mathbb{T}$ is irrational, and the system of polynomials $P = \{n, 2n, n^2\}$, for which we know that $V(P) = 2$ and $W(P) = 3$. Let A_0, A_1, A_2, A_3 be four measurable subsets of X with positive measure. We have the following results:

(i) If the sets A_i do not depend on the first coordinate x_1 (that is, each $A_i = \mathbb{T} \times I_i$ with $I_i \subseteq \mathbb{T}^3$), then

$$UC\text{-}\lim \mu_X(A_0 \cap T^{-n} A_1 \cap T^{-2n} A_2 \cap T^{-n^2} A_3) > 0.$$

(ii) If the sets A_i are independent of the algebra of subsets which depend only on x_1, x_2 (that is, $\mu_{\mathbb{T}^2}(A_i \cap L_{x_1, x_2}) = \mu_X(A_i)$, where $L_{x_1, x_2} = \{(x_1, x_2)\} \times \mathbb{T}^2$, for almost all $(x_1, x_2) \in \mathbb{T}^2$) then

$$UC\text{-}\lim \mu_X(A_0 \cap T^{-n} A_1 \cap T^{-2n} A_2 \cap T^{-n^2} A_3) = \mu_X(A_0) \mu_X(A_1) \mu_X(A_2) \mu_X(A_3).$$

(iii) If the sets A_i are independent of the algebra of subsets which depend only on x_1, x_2, x_3 then

$$\lim_{n \rightarrow \infty} \mu_X(A_0 \cap T^{-n} A_1 \cap T^{-2n} A_2 \cap T^{-n^2} A_3) = \mu_X(A_0) \mu_X(A_1) \mu_X(A_2) \mu_X(A_3).$$

It can be shown (see Lemmas 6.3 and 6.7 and Proposition 7.5 below) that in each of these three statements, the hypothesis of the independence of the sets A_i of the corresponding algebras cannot be weakened. Moreover, each of these results still holds if the variable n is replaced by any nonconstant integral polynomial $h(n)$ (see Propositions 6.12–6.14).

We now move to the general situation. Throughout this section let $P = \{p_1, \dots, p_r\}$ be a system of integral polynomials with zero constant term.

Let X and X' be compact commutative Lie groups with normalized Haar measures μ_X and $\mu_{X'}$ thereon, and let $\pi: X \rightarrow X'$ be a surjective (and continuous) homomorphism. Denote by F_z , $z \in X'$, the fibers of π , $F_z = \pi^{-1}(z)$. For a function $f \in L^1(X)$, the *conditional expectation* $E(f|X')$ of f with respect to X' is the function on X' defined by

$$E(f|X')(z) = \int_{F_z} f d\mu_{F_z} = \int_{F_0} f(z+x) d\mu_{F_0}(x).$$

(For $f \in L^2(X)$, $E(f|X')$ is the orthogonal projection of f onto the subspace $\pi^*(L^2(X'))$ of $L^2(X)$.) We will consider $E(f|X')$ as a function on X , as well as on X' .

We say that a measurable set $A \subseteq X$ is *independent of X'* if $E(1_A|X') = \mu_X(A)$; this is equivalent to saying that for almost all fibers F_z , $z \in Z$, of π one has $\mu_{F_z}(A) = \mu_X(A)$. If $A_1, \dots, A_r$ are subsets of X independent of X' , then $\prod_{i=1}^r A_i \subseteq X^r$ is independent of $(X')^r$.

We say that a measurable set $B \subseteq X$ *originates from X'* if $B = \pi^{-1}(B')$ for some $B' \subseteq X'$. If $B_1, \dots, B_r \subseteq X$ originate from X' then $\prod_{i=1}^r B_i \subseteq X^r$ originates from $(X')^r$. If B is a closed subgroup of X that originates from X' then $B = \pi^{-1}(B')$ for the subgroup $B' = \pi(B)$ of X' , and for any $f \in C(X)$ one has $\int_B f d\mu_B = \int_{B'} E(f|X') d\mu_{B'}$.

Let (X, T) be a quasi-standard Weyl system. We will first show that if $W(P) \leq k$, then

X_{k-1} is a characteristic factor for P . (Cf. [F1], §10 and [F3], p.54.) Put $g(n) = \begin{pmatrix} \text{Id}_X \\ T^{p_1(n)} \\ \vdots \\ T^{p_r(n)} \end{pmatrix}$

and $D_n = g(n)\Delta_{X^{r+1}}$, $n \in \mathbb{Z}$. By Proposition 5.4, the torus $H = \overline{\bigcup_{n \in \mathbb{Z}} D_n} \subseteq X^{r+1}$ contains L_k^{r+1} . Hence, the property of a point $x = (x_1, \dots, x_d) \in (\mathbb{T}^{r+1})^d$ to belong to

H does not depend on the coordinates $x_k, \dots, x_d$. Thus, $H = (\pi_{k-1}^{r+1})^{-1}(H_{k-1})$ so that H originates from X_{k-1}^{r+1} . Let $f_0, \dots, f_r$ be continuous functions on X and let $f = \bigotimes_{i=0}^r f_i$ (that is, f is the function on X^{r+1} defined by $f(x_0, \dots, x_r) = f_0(x_0) \cdot \dots \cdot f_r(x_r)$). Since H originates from X_{k-1}^{r+1} , we have

$$\int_H f d\mu_H = \int_{H_{k-1}} E(f|X_{k-1}^{r+1}) d\mu_{H_{k-1}}.$$

Since the sequence $\{D_n\}_{n \in \mathbb{Z}}$ is well distributed in H , $UC\text{-}\lim_n \int_{D_n} f d\mu_{D_n} = \int_H f d\mu_H$ and

$$UC\text{-}\lim_n \int_{D_n} E(f|X_{k-1}^{r+1}) d\mu_{D_n} = \int_H E(f|X_{k-1}^{r+1}) d\mu_H = \int_{H_{k-1}} E(f|X_{k-1}^{r+1}) d\mu_{H_{k-1}}.$$

For any $\tilde{z} = (z_0, \dots, z_r) \in X_{k-1}^{r+1}$,

$$E(f|X_{k-1}^{r+1})(\tilde{z}) = \int_{\tilde{z} + L_k^{r+1}} f_0 \otimes \dots \otimes f_r d\mu_{\tilde{z} + L_k^{r+1}} = \prod_{i=0}^r \int_{z_i + L_k} f_i d\mu_{z_i + L_k} = \prod_{i=0}^r E(f_i|X_{k-1})(z_i),$$

so $E(f|X_{k-1}^{r+1}) = \bigotimes_{i=0}^r E(f_i|X_{k-1})$. For any n we have

$$\int_{D_n} f d\mu_{D_n} = \int_{\Delta_{X^{r+1}}} g(n) f d\mu_{\Delta_{X^{r+1}}} = \int_X f_0 \cdot T^{p_1(n)} f_1 \cdot \dots \cdot T^{p_r(n)} f_r d\mu_X,$$

and similarly

$$\begin{aligned} \int_{D_n} E(f|X_{k-1}^{r+1}) d\mu_{D_n} &= \int_{\Delta_{X_{k-1}^{r+1}}} g(n) E(f|X_{k-1}^{r+1}) d\mu_{\Delta_{X_{k-1}^{r+1}}} \\ &= \int_{X_{k-1}} E(f_0|X_{k-1}) \cdot T^{p_1(n)} E(f_1|X_{k-1}) \cdot \dots \cdot T^{p_r(n)} E(f_r|X_{k-1}) d\mu_{X_{k-1}}. \end{aligned}$$

Thus,

$$\begin{aligned} UC\text{-}\lim_n \int_X f_0 \cdot T^{p_1(n)} f_1 \cdot \dots \cdot T^{p_r(n)} f_r d\mu_X \\ = UC\text{-}\lim_n \int_{X_{k-1}} E(f_0|X_{k-1}) \cdot T^{p_1(n)} E(f_1|X_{k-1}) \cdot \dots \cdot T^{p_r(n)} E(f_r|X_{k-1}) d\mu_{X_{k-1}}. \end{aligned}$$

Since continuous functions are dense in $L^\infty(X)$ in L^1 -topology, we obtain that X_{k-1} is a *characteristic factor* for the system P :

Proposition 6.1. *Let (X, T) be a quasi-standard Weyl system, let $W(P) \leq k$, and let $f_0, \dots, f_r \in L^\infty(X)$. Then*

$$\begin{aligned} UC\text{-}\lim_n \int_X f_0 \cdot T^{p_1(n)} f_1 \cdot \dots \cdot T^{p_r(n)} f_r d\mu_X \\ = UC\text{-}\lim_n \int_{X_{k-1}} E(f_0|X_{k-1}) \cdot T^{p_1(n)} E(f_1|X_{k-1}) \cdot \dots \cdot T^{p_r(n)} E(f_r|X_{k-1}) d\mu_{X_{k-1}}. \end{aligned}$$

We remark that one can derive from Proposition 6.1 and Theorem 3.5 the fact that the natural factor X_{k-1} is a characteristic factor for any (not necessarily quasi-standard) Weyl system. We do not give the details here since this fact is not needed in the sequel.

Applying Proposition 6.1 to the characteristic functions $1_{A_0}, \dots, 1_{A_r}$ of subsets $A_0, \dots, A_r$ of X that are independent of X_{k-1} we get

Proposition 6.2. *Let (X, T) be a quasi-standard Weyl system, let $W(P) \leq k$, and let $A_0, \dots, A_r$ be measurable subsets of X independent of X_{k-1} . Then $UC\text{-}\lim_n \mu_X(A_0 \cap T^{-p_1(n)} A_1 \cap \dots \cap T^{-p_r(n)} A_r) = \prod_{i=0}^r \mu_X(A_i)$.*

To be sure that X_{k-1} , the $(k-1)$ -st natural factor of X , is the “optimal” characteristic factor for a system of Weyl complexity k , we have to check that X_{k-2} is not characteristic. This is so indeed:

Lemma 6.3. *Let (X, T) be a quasi-standard Weyl system of depth $d \geq k-1$ and let $W(P) \geq k$. Then there exist functions $f_0, \dots, f_r \in L^\infty(X)$ such that*

$$\begin{aligned} UC\text{-}\lim_n \int_X f_0 \cdot T^{p_1(n)} f_1 \cdot \dots \cdot T^{p_r(n)} f_r d\mu_X \\ \neq UC\text{-}\lim_n \int_{X_{k-2}} E(f_0|X_{k-2}) \cdot T^{p_1(n)} E(f_1|X_{k-2}) \cdot \dots \cdot T^{p_r(n)} E(f_r|X_{k-2}) d\mu_{X_{k-2}}. \end{aligned}$$

Proof. By Proposition 5.4, H does not contain L_{k-1}^{r+1} . Thus there exists $\tilde{x} = (x_0, x_1, \dots, x_r) \in L_{k-1}^{r+1} \setminus H$. For each $i = 0, 1, \dots, r$ fix a nonnegative continuous function f_i on X such that $f_i(x_i) > 0$ and $f_0 \otimes f_1 \otimes \dots \otimes f_r|_H = 0$. The function $f_0 \otimes f_1 \otimes \dots \otimes f_r$ is zero on $g(n)\Delta_{X^{r+1}}$ for all n , thus

$$\int_X f_0 \cdot T^{p_1(n)} f_1 \cdot \dots \cdot T^{p_r(n)} f_r d\mu_X = 0$$

for all n . On the other hand, each function $E(f_i|X_{k-2})$ is positive in a neighborhood of the point 0 in H_{k-2} , so

$$\int_{H_{k-2}} E(f_0|X_{k-2}) \otimes E(f_1|X_{k-2}) \otimes \dots \otimes E(f_r|X_{k-2}) d\mu_{H_{k-2}} > 0.$$

Since the sequence $\{g(n)\Delta_{X_{k-2}^{r+1}}\}_{n \in \mathbb{Z}}$ is well distributed in H_{k-2} , the last expression is equal to

$$UC\text{-}\lim_n \int_{X_{k-2}} E(f_0|X_{k-2}) \cdot T^{p_1(n)} E(f_1|X_{k-2}) \cdot \dots \cdot T^{p_r(n)} E(f_r|X_{k-2}) d\mu_{X_{k-2}}.$$

■

In the case $W(P) < k$, X_{k-1} is characteristic for the system P in a stronger sense:

Proposition 6.4. *Let (X, T) be a quasi-standard Weyl system, let $W(P) < k$, and let $f_0, \dots, f_r \in L^\infty(X)$. Then*

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(\int_X f_0 \cdot T^{p_1(n)} f_1 \cdot \dots \cdot T^{p_r(n)} f_r d\mu_X \right. \\ \left. - \int_{X_{k-1}} E(f_0|X_{k-1}) \cdot T^{p_1(n)} E(f_1|X_{k-1}) \cdot \dots \cdot T^{p_r(n)} E(f_r|X_{k-1}) d\mu_{X_{k-1}} \right) = 0. \end{aligned} \quad (6.1)$$

In particular, we obtain:

Proposition 6.5. *Let (X, T) be a quasi-standard Weyl system, let $W(P) < k$, and let $A_0, \dots, A_r$ be measurable subsets of X independent of X_{k-1} . Then $\lim_{n \rightarrow \infty} \mu_X(A_0 \cap T^{-p_1(n)} A_1 \cap \dots \cap T^{-p_r(n)} A_r) = \prod_{i=0}^r \mu_X(A_i)$.*

To prove Proposition 6.4 we need to show that if at least one of the functions $f_0, \dots, f_r \in L^\infty(X)$ is orthogonal to the subspace $L^2(X_{k-1})$ of $L^2(X)$, then $\lim_{n \rightarrow \infty} \int_X f_0 \cdot T^{p_1(n)} f_1 \cdot \dots \cdot T^{p_r(n)} f_r d\mu = 0$. This follows from the following fact:

Lemma 6.6. *If $f_0, \dots, f_r$ are characters on X of which at least one is orthogonal to $L^2(X_{k-1})$ then $\int_X f_0 \cdot T^{p_1(n)} f_1 \cdot \dots \cdot T^{p_r(n)} f_r d\mu = 0$ for all but finitely many $n \in \mathbb{Z}$.*

Proof. We may and, for simplicity, will assume that (X, T) is a standard Weyl system: $X = \mathbb{T}^d$, $T(x_1, x_2, \dots, x_d) = (x_1 + \alpha, x_2 + x_1, \dots, x_d + x_{d-1})$ (though the proof is the same for any quasi-standard system). Let $f_i = \exp(2\pi i(m_{i,1}x_1 + \dots + m_{i,d}x_d))^{(*)}$, $m_{i,j} \in \mathbb{Z}$, $i = 0, \dots, r$. Then

$$T^{p_i(n)} f_i = \exp(2\pi i(m_{i,1}(x_1 + p_i(n)\alpha) + m_{i,2}(x_2 + p_i(n)x_1 + p_i(n)^{[2]}\alpha) + \dots + m_{i,d}(x_d + \sum_{j=1}^{d-1} p_i(n)^{[j]} x_{d-j} + p_i(n)^{[d]}\alpha))),$$

$i = 0, \dots, r$, and

$$f_0 \cdot \prod_{i=1}^r T^{p_i(n)} f_i = \exp\left(2\pi i(x_1(m_{0,1} + \sum_{i=1}^r \sum_{j=1}^d m_{i,j} p_i(n)^{[j-1]}) + x_2(m_{0,2} + \sum_{i=1}^r \sum_{j=2}^d m_{i,j} p_i(n)^{[j-2]}) + \dots + x_d(m_{0,d} + \sum_{i=1}^r m_{i,d}) + \alpha \sum_{i=1}^r \sum_{j=1}^d m_{i,j} p_i(n)^{[j]})\right),$$

$n \in \mathbb{Z}$. Thus, $\int_X f_0 \cdot T^{p_1(n)} f_1 \cdot \dots \cdot T^{p_r(n)} f_r d\mu = 0$ whenever at least one of the coefficients $m_{0,1} + \sum_{i=1}^r \sum_{j=1}^d m_{i,j} p_i(n)^{[j-1]}$, $m_{0,2} + \sum_{i=1}^r \sum_{j=2}^d m_{i,j} p_i(n)^{[j-2]}$, $\dots$, $m_{0,d} + \sum_{i=1}^r m_{i,d}$

is nonzero, that is, when $N(n)\overline{m} \neq 0$ where $\overline{m}$ is the vector $\begin{pmatrix} m_{0,1} \\ \vdots \\ m_{r,1} \\ \vdots \\ m_{0,d} \\ \vdots \\ m_{r,d} \end{pmatrix}$ and $N(n)$ is the matrix

$$\begin{pmatrix} 1 & 1 & \dots & 1 & 0 & p_1(n) & \dots & p_r(n) & 0 & p_1(n)^{[2]} & \dots & p_r(n)^{[2]} & 0 & p_1(n)^{[3]} & \dots & p_r(n)^{[3]} & \dots & \dots & 0 & p_1(n)^{[d-1]} & \dots & p_r(n)^{[d-1]} \\ 0 & 0 & \dots & 0 & 1 & \dots & 1 & 0 & p_1(n) & \dots & p_r(n) & 0 & p_1(n)^{[2]} & \dots & p_r(n)^{[2]} & \dots & \dots & 0 & p_1(n)^{[d-2]} & \dots & p_r(n)^{[d-2]} \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \dots & 1 & 0 & p_1(n) & \dots & p_r(n) & \dots & \dots & 0 & p_1(n)^{[d-3]} & \dots & p_r(n)^{[d-3]} \\ \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & \dots & 1 & 1 & \dots & 1 \end{pmatrix},$$

$n \in \mathbb{Z}$. Assume that $N(n)\overline{m} = 0$ for infinitely many $n \in \mathbb{Z}$. Since the condition $N(n)\overline{m} = 0$ is polynomial in n , we then have $N(n)\overline{m} = 0$ for all $n \in \mathbb{Z}$. In particular, $N(0)\overline{m} = 0$,

(*) In the expression “ $2\pi i$ ”, π is not a projection but $3.14\dots$, and i is not the index appearing in the rest of the formula but $\sqrt{-1}$.

therefore $(N(n) - N(0))\overline{m} = 0$ and $\begin{pmatrix} N(0) \\ N(n) - N(0) \end{pmatrix} \overline{m}$ for all $n \in \mathbb{Z}$. We have

$$N(0) = \begin{pmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 1 & 1 & \dots & 1 \end{pmatrix},$$

so

$$\begin{pmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 1 & 1 & \dots & 1 \\ 0 & 0 & \dots & 0 & 0 & p_1(n) & \dots & p_r(n) & 0 & p_1(n)^{[2]} & \dots & p_r(n)^{[2]} & \dots & \dots & 0 & p_1(n)^{[d-1]} & \dots & p_r(n)^{[d-1]} & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & p_1(n) & \dots & p_r(n) & \dots & \dots & 0 & p_1(n)^{[d-2]} & \dots & p_r(n)^{[d-2]} & \dots & \dots & \dots \\ \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 0 & p_1(n) & \dots & p_r(n) & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 0 & 0 & \dots & 0 & \dots & \dots & \dots \end{pmatrix} \overline{m} = 0$$

and thus (after erasing the $r + 1$ first columns, and the first and the last lines)

$$\begin{pmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & \dots & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 1 & 1 & \dots & 1 \\ 0 & p_1(n) & \dots & p_r(n) & 0 & p_1(n)^{[2]} & \dots & p_r(n)^{[2]} & \dots & \dots & 0 & p_1(n)^{[d-1]} & \dots & p_r(n)^{[d-1]} \\ 0 & 0 & \dots & 0 & 0 & p_1(n) & \dots & p_r(n) & \dots & \dots & 0 & p_1(n)^{[d-2]} & \dots & p_r(n)^{[d-2]} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 0 & p_1(n) & \dots & p_r(n) \end{pmatrix} \tilde{m} = 0$$

for all $n \in \mathbb{Z}$, where $\tilde{m} = \begin{pmatrix} m_{0,2} \\ \vdots \\ m_{r,2} \\ \vdots \\ m_{0,d} \\ \vdots \\ m_{r,d} \end{pmatrix} \in \mathbb{R}^{(d-1)(r+1)}$. After introducing the standard inner

product on the space $\mathbb{R}^{(d-1)(r+1)}$, we interpret this identity as the fact that the vector $\tilde{m}$ is orthogonal to the subspace

$$\mathcal{H} = \text{Span} \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & p_1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \dots & 0 & p_r & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & p_1^{[2]} & p_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & p_r^{[2]} & p_r & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 1 & p_1^{[d-1]} & p_1^{[d-2]} & \dots & p_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & p_r^{[d-1]} & p_r^{[d-2]} & \dots & p_r \end{pmatrix} \quad (6.2)$$

of $\mathbb{R}^{(d-1)(r+1)}$. Comparing the formulas (6.2) and (4.2) we see that $\mathcal{H} \bmod 1 = H_{d-1}$. If $W(p_1, \dots, p_r) \leq k - 1$ then H contains L_{k-1}^{r+1} , and thus $\mathcal{H}$ contains the subspace $\mathcal{L}_{k-1} = \{(0, \dots, 0, u_{0,k-1}, \dots, u_{r,k-1}, \dots, u_{0,d-1}, \dots, u_{r,d-1}), u_{i,j} \in \mathbb{R}\}$ of $\mathbb{R}^{(d-1)(r+1)}$. But if f_i is orthogonal to X_{k-1} for some i then $m_{i,j} \neq 0$ for some $j \geq k$ and thus $\tilde{m}$ is not orthogonal to $\mathcal{L}_{k-1}$. ■

The following lemma shows that the assumption $W(P) < k$ in Proposition 6.4 cannot be weakened:

Lemma 6.7. *Let (X, T) be a quasi-standard Weyl system of depth $d \geq k$ and let $W(P) \geq k$. Then there exist functions $f_0, \dots, f_r \in L^1(X)$ such that identity (6.1) does not hold.*

Proof. In the notation of the proof of Lemma 6.6, if $W(P) \geq k$ then H does not contain

L_{k-1}^{r+1} , and thus there exists a vector $\tilde{m} = \begin{pmatrix} m_{0,2} \\ \vdots \\ m_{r,2} \\ \vdots \\ m_{0,d} \\ \vdots \\ m_{r,d} \end{pmatrix} \in \mathbb{Z}^{(d-1)(r+1)}$ with $m_{i_0, j_0} \neq 0$ for

some i_0 and some $j_0 \geq k$ and orthogonal to $\mathcal{H}$. For the vector $\bar{m} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ m_{0,2} \\ \vdots \\ m_{r,2} \\ \vdots \\ m_{0,d} \\ \vdots \\ m_{r,d} \end{pmatrix} \in \mathbb{Z}^{d(r+1)}$

we then have $N(n)\bar{m} = 0$ for all n . Put $f_i = \exp(2\pi i(m_{i,1}x_1 + \dots + m_{i,d}x_d))$, $i = 0, \dots, r$; then, for any $n \in \mathbb{Z}$, $f_0 \cdot T^{p_1(n)}f_1 \cdot \dots \cdot T^{p_r(n)}f_r$ is a constant of modulus 1. On the other hand, since $m_{i_0, j_0} \neq 0$ with $j_0 \geq k$, we have $E(f_{i_0}|X_{k-1}) = 0$, and thus $E(f_0|X_{k-1}) \cdot T^{p_1(n)}E(f_1|X_{k-1}) \cdot \dots \cdot T^{p_r(n)}E(f_r|X_{k-1}) = 0$ for all $n \in \mathbb{Z}$. ■

We now turn to the Vandermonde complexity; let us assume that $V(P) \leq k$. To simplify notation, assume also that (X, T) is a quasi-standard Weyl system of depth k (that is, $X = X_k = \mathbb{T}^k$). Let functions $f_i \in L^1(X)$, $i = 0, \dots, r$, be nonnegative, with $\int_X f_i d\mu_X > 0$, and independent of X_{k-1} in a very strong sense: assume that they only depend on the last, k th coordinate of X . Then

Proposition 6.8. *If $V(p_1, \dots, p_r) \leq k$ then $UC\text{-}\lim_n \int_X f_0 \cdot T^{p_1(n)}f_1 \cdot \dots \cdot T^{p_r(n)}f_r d\mu_X > 0$.*

Proposition 6.8 is equivalent to the following:

Proposition 6.9. *Let $V(p_1, \dots, p_r) \leq k$ and let $A_i = X_{k-1} \times I_i$, $i = 0, \dots, r$, where $I_0, \dots, I_r$ are subsets of F_k of positive measure. Then $UC\text{-}\lim_n \mu_X(A_0 \cap T^{-p_1(n)}A_1 \cap \dots \cap T^{-p_r(n)}A_r) > 0$.*

Proof. We may assume that $I_0, \dots, I_r$ are open intervals in $F_k \simeq \mathbb{T}$. Let $A = \prod_{i=0}^r A_i$. Since $\tau_k^{r+1}(H) = F_k^{r+1}$, there is a point $\tilde{x} \in H$ with $\tau_k^{r+1}(\tilde{x}) \in I_0 \times \dots \times I_r$. Thus, $A \cap H \neq \emptyset$. Since $A \cap H$ is open in H , $\mu_H(A \cap H) > 0$. Since the sequence D_n is well distributed in H ,

$$UC\text{-}\lim_n \mu_X(A_0 \cap T^{-p_1(n)}A_1 \cap \dots \cap T^{-p_r(n)}A_r) = UC\text{-}\lim_n \mu_{D_n}(A \cap D_n) = \mu_H(A \cap H) > 0.$$

■

Remark. If a set A is independent of X_{k-1} then so is the set $T^c A$ for any $c \in \mathbb{Z}$; this implies that the assertions of Proposition 6.2 and Proposition 6.5 remain true if the intersection $A_0 \cap T^{-p_1(n)} A_1 \cap \dots \cap T^{-p_r(n)} A_r$ is replaced by $A_0 \cap T^{-p_1(n)-c_1} A_1 \cap \dots \cap T^{-p_r(n)-c_r} A_r$ with arbitrary $c_i \in \mathbb{Z}$ (see also Proposition 6.15 and Proposition 6.17 below). A similar extension of Proposition 6.9 does not hold: one can construct an explicit example of a system of integer polynomials $P = \{p_1, \dots, p_r\}$ with $V(P) = 2$, integers $c_1, \dots, c_r$, a quasi-standard Weyl system $(\mathbb{T}^2, T)$ and intervals $I_0, \dots, I_r$ in $\mathbb{T}$ such that for $A_i = \mathbb{T} \times I_i \subseteq \mathbb{T}^2$ one has $(A_0 \times \prod_{i=1}^r T^{-c_i} A_i) \cap H = \emptyset$.

We will now obtain further refinements of the preceding results by considering the system of integral polynomials $P(h) = \{p_1(h(n)), \dots, p_r(h(n))\}$, where h is any nonconstant integral polynomial. If $h(0) = 0$, by Proposition 5.3 the system $P(h)$ has the same Weyl complexity as P . Let, again, $g(n) = \begin{pmatrix} \text{Id}_X \\ T^{p_1(n)} \\ \vdots \\ T^{p_r(n)} \end{pmatrix}$, $n \in \mathbb{Z}$; then even if $h(0) \neq 0$, we have:

Lemma 6.10. *For any nonconstant integral polynomial h the sequence $\{g(h(n))\Delta_{X^{r+1}}\}_{n \in \mathbb{Z}}$ is well distributed in H .*

Proof. By Proposition 3.2, the sequence $g(h(n))\Delta_{X^{r+1}}$, $n \in \mathbb{Z}$, is well distributed in H if it is dense in H . Let x be a “generic” point of X , that is, let the coordinates of x and the elements $\alpha_i \in \mathbb{T}$ in the definition of the quasi-standard Weyl system be rationally independent. Then the closure $H_x = \overline{\{g(n)\bar{x}, n \in \mathbb{Z}\}} \subseteq X^{r+1}$ of the orbit of $\bar{x}$ under g has form $\bar{x} + \text{Span } Q \bmod 1$ for some polynomial matrix Q . The closure $K_x = \overline{\{g(h(n))\bar{x}, n \in \mathbb{Z}\}} \subseteq X^{r+1}$ of the orbit of $\bar{x}$ under $g(h)$ has form $g(h(0))\bar{x} + \text{Span}(Q(h) - Q(h(0))) \bmod 1$. Since $g(h(0))\bar{x} \in H_x$ and $\text{Span}(Q(h) - Q(h(0))) = \text{Span } Q$ by Lemma 2.3, we obtain $K_x = H_x$. Hence,

$$\bigcup_{n \in \mathbb{Z}} \overline{g(h(n))\Delta_{X^{r+1}}} = \bigcup_{x \in X} \overline{K_x} = \bigcup_{x \in X} \overline{H_x} = H.$$

■

We may now strengthen Proposition 6.1:

Proposition 6.11. *Let $W(P) \leq k$ and let $f_0, \dots, f_r \in L^\infty(X)$. For any nonconstant integral polynomial h one has*

$$\begin{aligned} UC\text{-}\lim_n \int_X f_0 \cdot T^{p_1(h(n))} f_1 \cdot \dots \cdot T^{p_r(h(n))} f_r d\mu_X \\ = UC\text{-}\lim_n \int_{X_{k-1}} E(f_0|X_{k-1}) \cdot T^{p_1(h(n))} E(f_1|X_{k-1}) \cdot \dots \cdot T^{p_r(h(n))} E(f_r|X_{k-1}) d\mu_{X_{k-1}}. \end{aligned}$$

Applying this to $f_i = 1_{A_i}$, we get

Proposition 6.12. *Let $W(P) \leq k$ and let $A_0, \dots, A_r$ be measurable subsets of X independent of X_{k-1} . For any nonconstant integral polynomial h one has*

$$UC\text{-}\lim_n \mu_X(A_0 \cap T^{-p_1(h(n))} A_1 \cap \dots \cap T^{-p_r(h(n))} A_r) = \prod_{i=0}^r \mu_X(A_i).$$

When $W(P) < k$, Lemma 6.6 immediately implies:

Proposition 6.13. *If $W(P) < k$ and $A_0, \dots, A_r$ are measurable subsets of X independent of X_{k-1} , then for any nonconstant integral polynomial h one has*

$$\lim_{n \rightarrow \infty} \mu_X(A_0 \cap T^{-p_1(h(n))} A_1 \cap \dots \cap T^{-p_r(h(n))} A_r) = \prod_{i=0}^r \mu_X(A_i).$$

(We will no longer deal with functions on X but only with subsets of X , though our statements can, of course, be easily reformulated in the language of functions.) Assuming that X has depth k , directly from Lemma 6.10 we get

Proposition 6.14. *Let $V(P) \leq k$ and let $A_i = X_{k-1} \times I_i$, $i = 0, \dots, r$, where $I_0, \dots, I_r$ are subsets of F_k of positive measure. Then for any nonconstant integral polynomial h ,*

$$UC\text{-}\lim_n \mu_X(A_0 \cap T^{-p_1(h(n))} A_1 \cap \dots \cap T^{-p_r(h(n))} A_r) > 0.$$

Let $m \geq 2$ and $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$. We will now investigate the *non-connected Weyl system* (Y, R) where $Y = X \times \mathbb{Z}_m$ and $R: Y \rightarrow Y$ is defined by $R(x, j) = (Tx, j+1)$; for further generality, we will also add, when possible, “shifting constants” c_i in the formulation of our results:

Proposition 6.15. *Let $W(P) \leq k$, and let B be a measurable subset of Y independent of $X_{k-1} \times \mathbb{Z}_m$. Then for any nonconstant integral polynomial h and any $c_1, \dots, c_r \in \mathbb{Z}$,*

$$UC\text{-}\lim_n \mu_Y(B \cap R^{-p_1(h(n)) - c_1} B \cap \dots \cap R^{-p_r(h(n)) - c_r} B) = \mu_Y(B)^{r+1}.$$

Proof. Let $B = \bigcup_{j \in \mathbb{Z}_m} (A_j \times \{j\})$. Then $A_0, \dots, A_{m-1}$ are independent of X_{k-1} , and $\mu_X(A_0) = \dots = \mu_X(A_{m-1}) = \mu_Y(B)$. Put $B_j = A_j \times \{j\}$, $j \in \mathbb{Z}_m$.

For $l \in \{0, \dots, m-1\}$, consider the system $P(h(mn+l)) = \{p_1(h(mn+l)) + c_1, \dots, p_r(h(mn+l)) + c_r\}$ of polynomials in the variable n . For any i and j , $R^{-p_i(h(mn+l)) - c_i} B_j \subseteq X \times \{j - d_i\}$ for all $n \in \mathbb{Z}$, where $d_i = p_i(h(l)) + c_i \bmod m$. Thus, for $j_0, j_1, \dots, j_r \in \mathbb{Z}_m$, if $j_i = j_0 + d_i$ for all $i = 0, \dots, r$ then by Proposition 6.12

$$\begin{aligned} UC\text{-}\lim_n \mu_Y(B_{j_0} \cap R^{-p_1(h(mn+l)) - c_1} B_{j_1} \cap \dots \cap R^{-p_r(h(mn+l)) - c_r} B_{j_r}) \\ &= \frac{1}{m} UC\text{-}\lim_n \mu_X(A_{j_0} \cap T^{-p_1(h(mn+l)) - c_1} A_{j_1} \cap \dots \cap T^{-p_r(h(mn+l)) - c_r} A_{j_r}) \\ &= \frac{1}{m} UC\text{-}\lim_n \mu_X(A_{j_0} \cap T^{-p_1(h(mn+l))} (T^{-c_1} A_{j_1}) \cap \dots \cap T^{-p_r(h(mn+l))} (T^{-c_r} A_{j_r})) \\ &= \frac{1}{m} \prod_{i=0}^r \mu_X(T^{-c_i} A_{j_i}) = \frac{1}{m} \prod_{i=0}^r \mu_X(A_{j_i}) = \frac{1}{m} \mu_Y(B)^{r+1}, \end{aligned}$$

and otherwise $B_{j_0} \cap R^{-p_1(h(mn+l)) - c_1} B_{j_1} \cap \dots \cap R^{-p_r(h(mn+l)) - c_r} B_{j_r} = \emptyset$ for all n . Thus,

$$\begin{aligned} UC\text{-}\lim_n \mu_Y(B \cap R^{-p_1(h(mn+l)) - c_1} B \cap \dots \cap R^{-p_r(h(mn+l)) - c_r} B) \\ &= \sum_{j_0, \dots, j_r \in \mathbb{Z}_m} UC\text{-}\lim_n \mu_Y(B_{j_0} \cap R^{-p_1(h(mn+l)) - c_1} B_{j_1} \cap \dots \cap R^{-p_r(h(mn+l)) - c_r} B_{j_r}) \\ &= \sum_{j_0 \in \mathbb{Z}_m} UC\text{-}\lim_n \mu_Y(B_{j_0} \cap R^{-p_1(h(mn+l)) - c_1} B_{j_0+d_1} \cap \dots \cap R^{-p_r(h(mn+l)) - c_r} B_{j_0+d_r}) \\ &= m \cdot \frac{1}{m} \mu_Y(B)^{r+1} = \mu_Y(B)^{r+1}. \end{aligned}$$

Since this is true for every $l = 0, 1, \dots, m-1$, we get

$$UC\text{-}\lim_n \mu_Y(B \cap R^{-p_1(h(n))-c_1} B \cap \dots \cap R^{-p_r(h(n))-c_r} B) = \mu_Y(B)^{r+1}.$$

■

Remark. A similar proof allows one to get a more general result:

Proposition 6.16. *Let $W(P) \leq k$, and let $B_{(0)}, \dots, B_{(r)}$ be measurable subsets of Y independent of $X_{k-1} \times \mathbb{Z}_m$. Then for any nonconstant integral polynomial h*

$$UC\text{-}\lim_n \mu_Y(B_{(0)} \cap R^{-p_1(h(n))} B_{(1)} \cap \dots \cap R^{-p_r(h(n))} B_{(r)}) = \prod_{i=0}^r \mu_Y(B_{(i)}).$$

(This proposition gives Proposition 6.15 if we apply it to $B_{(0)} = B$ and $B_{(i)} = R^{-c_i} B$, $i = 1, \dots, r$.) Same remark applies also to Proposition 6.17 and, with certain modifications, to Proposition 6.18 below.

If in the proof of Proposition 6.15 we replace $UC\text{-}\lim$ by $\lim$ and Proposition 6.12 by Proposition 6.13, we get

Proposition 6.17. *If $W(P) < k$ and B is a measurable subset of Y independent of $X_{k-1} \times \mathbb{Z}_m$, then for any nonconstant integral polynomial h and any $c_1, \dots, c_r \in \mathbb{Z}$,*

$$\lim_{n \rightarrow \infty} \mu_Y(B \cap R^{-p_1(h(n))-c_1} B \cap \dots \cap R^{-p_r(h(n))-c_r} B) = \mu_Y(B)^{r+1}.$$

In the notation of Proposition 6.15, assume now that $V(P) \leq k$, X has depth k and $B = \bigcup_{j \in \mathbb{Z}_m} B_j = \bigcup_{j \in \mathbb{Z}_m} (A_j \times \{j\})$ where each A_j has form $X_{k-1} \times I_j$ for $I_j \subseteq F_k$, $j \in \mathbb{Z}_m$, of positive measure. Then, in the same way, we obtain from Proposition 6.14

Proposition 6.18. *For any nonconstant integral polynomial h ,*

$$UC\text{-}\lim_n \mu_Y(B \cap R^{-p_1(h(n))} B \cap \dots \cap R^{-p_r(h(n))} B) > 0.$$

7. Combinatorics

Let $E \subseteq \mathbb{Z}$ with $UD(E) > 0$ (that is, the uniform density $UD(E)$ exists and is positive) and let $P = \{p_1, \dots, p_r\}$ be a system of integral polynomials (with not necessarily zero constant term). For $n \in \mathbb{Z}$, define

$$E_n = \{a \in \mathbb{Z} : a, a + p_1(n), \dots, a + p_r(n) \in E\} = E \cap (E - p_1(n)) \cap \dots \cap (E - p_r(n)).$$

Recall that E is *UC-positive* with respect to P if $UC\text{-}\lim_n UD(E_n) > 0$; that E is *UC-balanced* with respect to P if $UC\text{-}\lim_n UD(E_n) = UD(E)^{r+1}$; and that E is *balanced* with respect to P if $\lim_{n \rightarrow \infty} UD(E_n) = UD(E)^{r+1}$.

Theorem 7.1. *For any $k \in \mathbb{N}$ and any system $\{p_1, \dots, p_r\}$ of integral polynomials with zero constant term and with $V(p_1, \dots, p_r) > k$, there is a set $E \subset \mathbb{Z}$ of positive uniform density such that*

- (i) *for any system of integral polynomials $\{q_1, \dots, q_s\}$ with zero constant term and with $V(q_1, \dots, q_s) \leq k$ and any nonconstant integral polynomial h the set E is UC-positive with respect to the system $\{q_1(h(n)), \dots, q_s(h(n))\}$;*
- (ii) *for any system $\{q_1, \dots, q_s\}$ of integral polynomials with zero constant term and with $W(q_1, \dots, q_s) \leq k$, any nonconstant integral polynomial h and any integers $c_1, \dots, c_s$ the set E is UC-balanced with respect to the system $\{q_1(h(n)) + c_1, \dots, q_s(h(n)) + c_s\}$;*
- (iii) *for any system $\{q_1, \dots, q_s\}$ of integral polynomials with zero constant term and with $W(q_1, \dots, q_s) < k$, any nonconstant integral polynomial h and any integers $c_1, \dots, c_s$ the set E is balanced with respect to the system $\{q_1(h(n)) + c_1, \dots, q_s(h(n)) + c_s\}$;*
- (iv) *there exist nonzero integers m and l such that E contains no configuration of the form $\{a, a + p_1(mn + l), \dots, a + p_r(mn + l)\}$, $a, n \in \mathbb{Z}$.*

It is possible to characterize the pairs (m, l) that work for (iv). For fixed m and l , consider the vectors $u_j = (u_j^{(0)}, \dots, u_j^{(r)}) \in \mathbb{R}^{r+1}$, $j = 0, \dots, m-1$, where $u_j^{(i)} = 1$ if $p_i(l) \equiv j \pmod{m}$ and $u_j^{(i)} = 0$ otherwise; we assume here $p_0 = 0$. Let us say that the pair

(m, l) separates $p_1, \dots, p_r$ on the level k if $u_j \notin \text{Span} \left(\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & p_1 & p_1^2 & \dots & p_1^k \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & p_r & p_r^2 & \dots & p_r^k \end{pmatrix} \right)$ for some j . This may

only be the case when $k < V(p_1, \dots, p_r)$. On the other hand, for m large enough there exists l such that $0, p_1(l), \dots, p_r(l)$ are all different modulo m , and then the pair (m, l) separates $p_1, \dots, p_r$ on the level k for all $k < V(p_1, \dots, p_r)$.

Example. For the system $\{x, x^2, x + x^2, x + 2x^2\}$, $m = 2$ and $l = 1$ we have $u_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ and $u_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$. Since

$$u_0 \notin \text{Span} \left(\begin{pmatrix} 1 & 0 \\ 1 & x \\ 1 & x^2 \\ 1 & x+x^2 \end{pmatrix} \right) = \text{span} \left(\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} \right),$$

the pair $(2, 1)$ separates $x, x^2, x + x^2, x + 2x^2$ on the level 1.

For the same system and $(m, l) = (3, 1)$ we have $u_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, $u_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $u_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, and so, the pair $(3, 1)$ also separates $x, x^2, x + x^2, x + 2x^2$ on the level 1.

We may now strengthen Theorem 7.1:

Theorem 7.2. *For any $k \in \mathbb{N}$, any system $\{p_1, \dots, p_r\}$ of integral polynomials with zero constant term and with $V(p_1, \dots, p_r) > k$, and any integers $m, l_1, \dots, l_\nu$ such that each of the pairs $(m, l_1), \dots, (m, l_\nu)$ separates $p_1, \dots, p_r$ on the level k there is a set $E \subset \mathbb{Z}$ of positive uniform density such that*

(i), (ii), (iii) as in Theorem 7.1;

(iv) E contains no configuration of the form $\{a, a + p_1(mn + l), \dots, a + p_r(mn + l)\}$, $a, n \in \mathbb{Z}$, with $l \in \{l_1, \dots, l_\nu\}$.

Example. Consider the system $\{x, 2x, \dots, rx\}$. The Vandermonde complexity of this

system equals r , and $\text{Span} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & p_1 & p_1^2 & \dots & p_1^{r-1} \\ 1 & p_2 & p_2^2 & \dots & p_2^{r-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & p_r & p_r^2 & \dots & p_r^{r-1} \end{pmatrix} = \text{span} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2^2 & \dots & 2^{r-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & r & r^2 & \dots & r^{r-1} \end{pmatrix}$ is the hyperplane de-

scribed by the equation $\sum_{i=0}^r (-1)^i \binom{r}{i} u^{(i)} = 0$. For $m = 2$ and $l = 1$, the vector $u_0 = (1, 0, 1, 0, \dots, 1, 0)$ is not contained in this hyperplane, and thus the pair $(2, 1)$ separates the polynomials $x, 2x, \dots, rx$ on the level $r - 1$. Actually, the following holds:

Lemma 7.3. *For any r, m, l with l not divisible by m , the pair (m, l) separates the polynomials $x, 2x, \dots, rx$ on the level $r - 1$.*

Proof. For any m and any l not divisible by m , the nonzero vectors u_i corresponding to these m, l and the system $\{x, 2x, \dots, rx\}$ are

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix};$$

they are periodic with same distance $b = m / \gcd(l, m)$ between “1”s. We therefore have to check that for some $j \in \{0, \dots, b - 1\}$ the number $e_j = \sum_{\substack{0 \leq i \leq r \\ i \equiv j \pmod{b}}} (-1)^i \binom{r}{i}$ is nonzero.

Let λ be a primitive root of unity of degree b ; then we have

$$0 \neq (\lambda - 1)^r = \sum_{i=0}^r (-1)^i \binom{r}{i} \lambda^{r-i} = \sum_{j=0}^{b-1} \sum_{\substack{0 \leq i \leq r \\ i \equiv j \pmod{b}}} (-1)^i \binom{r}{i} \lambda^{r-j} = \sum_{j=0}^{b-1} e_j \lambda^{r-j},$$

and thus one of e_j must be nonzero. $\blacksquare$

Since $V(x, 2x, \dots, rx) = r$, we obtain:

Corollary 7.4. *For any $r, m \geq 2$ there is a set $E \subset \mathbb{Z}$ of positive uniform density such that*

(i) *for any system $\{q_1, \dots, q_s\}$ of integral polynomials with zero constant term and with $V(q_1, \dots, q_s) \leq r - 1$ and any nonconstant integral polynomial h , the set E is UC-positive with respect to the system $\{q_1(h(n)), \dots, q_s(h(n))\}$;*

(ii) *for any system $\{q_1, \dots, q_s\}$ of integral polynomials with zero constant term and with $W(q_1, \dots, q_s) \leq r - 1$, any nonconstant integral polynomial h and any integers $c_1, \dots, c_s$, the set E is UC-balanced with respect to the system $\{q_1(h(n)) + c_1, \dots, q_s(h(n)) + c_s\}$;*

- (iii) for any system $\{q_1, \dots, q_s\}$ of integral polynomials with zero constant term and with $W(q_1, \dots, q_s) < r - 1$, any nonconstant integral polynomial h and any integers $c_1, \dots, c_s$, the set E is balanced with respect to the system $\{q_1(h(n)) + c_1, \dots, q_s(h(n)) + c_s\}$;
- (iv) E contains no arithmetic progressions of the form $\{a, a + (mn + l), \dots, a + r(mn + l)\}$, $a, n, l \in \mathbb{Z}$, with l not divisible by m .

Let $\{p_1, \dots, p_r\}$ be a system of integral polynomials with zero constant term and with $V(p_1, \dots, p_r) > k$ and let integers $m, l_1, \dots, l_\nu$ be such that each of the pairs $(m, l_1), \dots, (m, l_\nu)$ separates these polynomials on the level k . Consider the system (Y, R) introduced at the end of the previous section, namely, $Y = X \times \mathbb{Z}_m$ and $R(x, i) = (Tx, i + 1)$, where (X, T) is a quasi-standard Weyl system of depth k . We keep the notation of the preceding section. Let, again, $B = \bigcup_{j \in \mathbb{Z}_m} B_j = \bigcup_{j \in \mathbb{Z}_m} (A_j \times \{j\})$ where $A_j = X_{k-1} \times I_j$, I_j are open intervals in F_k , $j \in \mathbb{Z}_m$. The dynamical reason for Theorem 7.2 being true is the following proposition:

Proposition 7.5. *The intervals I_j can be chosen so that $B \cap R^{-p_1(mn+l)}B \cap \dots \cap R^{-p_r(mn+l)}B = \emptyset$ for all $n \in \mathbb{Z}$ and all $l = l_1, \dots, l_\nu$.*

Proof. Let l be one of $l_1, \dots, l_\nu$. Let $d_i = p_i(l) \bmod m \in \mathbb{Z}_m$, $i = 0, \dots, r$. Let $u_j = (u_j^{(0)}, \dots, u_j^{(r)}) \in \mathbb{R}^{r+1}$, $j \in \mathbb{Z}_m$, where $u_j^{(i)} = 1$ if $d_i \equiv j$ and $u_j^{(i)} = 0$ otherwise. We have

$$U = \text{span}(u_j, j \in \mathbb{Z}_m) \bmod 1 = \left\{ \sum_{j \in \mathbb{Z}_m} u_j t_j, t_j \in \mathbb{T} \right\} \\ = \{(t_0, t_{d_1}, \dots, t_{d_r}), t_j \in \mathbb{T}, j \in \mathbb{Z}_m\} \subseteq \mathbb{T}^{r+1}.$$

Since (m, l) separates $\{p_1, \dots, p_r\}$ on the level k , U is not contained in the subtorus $\tau_k^{r+1}(H)$ of F_k^{r+1} (where τ_k is, again, the projection of X to F_k , and where we identify F_k with $\mathbb{T}$). Hence, the preimage of $\tau_k^{r+1}(H)$ in $\mathbb{T}^m$ under the mapping $(t_0, t_1, \dots, t_{m-1}) \mapsto (t_0, t_{d_1}, \dots, t_{d_r}) \in F_k^{r+1}$ is a proper subtorus of $\mathbb{T}^m$, so it is “negligible” in $\mathbb{T}^m$. Thus, there exist elements $t_j \in \mathbb{T}$ for $j \in \mathbb{Z}_m$, such that not only $(t_0, t_{d_1}, \dots, t_{d_r}) \notin \tau_k^{r+1}(H)$, but also $(t_{\sigma(0)}, t_{\sigma(d_1)}, \dots, t_{\sigma(d_r)}) \notin \tau_k^{r+1}(H)$ for any permutation σ of $\mathbb{Z}_m$. Moreover, t_j , $j \in \mathbb{Z}_m$, can be chosen so that this will hold for each $l = l_1, \dots, l_\nu$. Put $I_j = (t_j - \delta, t_j + \delta) \subset \mathbb{T} = F_k$, $j \in \mathbb{Z}_m$, where $\delta > 0$ is small enough to ensure $I_{\sigma(0)} \times I_{\sigma(d_1)} \times \dots \times I_{\sigma(d_r)} \cap \tau_k^{r+1}(H) = \emptyset$ for all σ , and, again, for any choice of $l \in \{l_1, \dots, l_\nu\}$. Now define $A_j = X_{k-1} \times I_j \subset X$. Then for any permutation σ of $\mathbb{Z}_m$, $A_{\sigma(0)} \times A_{\sigma(d_1)} \times \dots \times A_{\sigma(d_r)} \cap H = \emptyset$.

Let $l \in \{l_1, \dots, l_\nu\}$. Put $B_j = A_j \times \{j\} \subset Y$, $j \in \mathbb{Z}_m$, and $B = \bigcup_{j \in \mathbb{Z}_m} B_j$. Let $\Delta_{Y^{r+1}}$ be the diagonal in Y^{r+1} . One has $B \cap R^{-p_1(mn+l)}B \cap \dots \cap R^{-p_r(mn+l)}B = \emptyset$ for all $n \in \mathbb{Z}$ if $B^{r+1} \cap \left(\begin{smallmatrix} \text{Id}_Y \\ R^{p_1(mn+l)} \\ \vdots \\ R^{p_r(mn+l)} \end{smallmatrix} \right) \Delta_{Y^{r+1}} = \emptyset$ for all $n \in \mathbb{Z}$. Let $K = \bigcup_{n \in \mathbb{Z}} \left(\begin{smallmatrix} \text{Id}_Y \\ R^{p_1(mn+l)} \\ \vdots \\ R^{p_r(mn+l)} \end{smallmatrix} \right) \Delta_{Y^{r+1}}$. Put $\bar{z} = (z, \dots, z) \in (\mathbb{Z}_m)^{r+1}$ and $\Delta_z = \Delta_{X^{r+1}} \times \bar{z}$, $z \in \mathbb{Z}_m$. Then $\Delta_{Y^{r+1}} = \bigcup_{z \in \mathbb{Z}_m} \Delta_z$ and $K = \bigcup_{z \in \mathbb{Z}_m} H_z$, where $H_z = \bigcup_{n \in \mathbb{Z}} \left(\begin{smallmatrix} \text{Id}_Y \\ R^{p_1(mn+l)} \\ \vdots \\ R^{p_r(mn+l)} \end{smallmatrix} \right) \Delta_z$, $z \in \mathbb{Z}_m$. By Lemma 6.10, for each

$z \in \mathbb{Z}_m$, $H_z = H \times (z, d_1 + z, \dots, d_r + z)$, and so

$$\begin{aligned} B^{r+1} \cap H_z &= (B_z \times B_{d_1+z} \times \dots \times B_{d_r+z}) \cap H_z \\ &= \left((A_z \times A_{d_1+z} \times \dots \times A_{d_r+z}) \cap H \right) \times (z, d_1 + z, \dots, d_r + z) = \emptyset. \end{aligned}$$

■

Example. Consider the polynomial system $\{x, 2x, 3x, 4x\}$. We have $V(x, 2x, 3x, 4x) = 4$, and the pair $(2, 1)$ separates $x, 2x, 3x, 4x$ on the level 3. (Indeed, the vectors

$$u_0 = (1, 0, 1, 0, 1) \text{ and } u_1 = (0, 1, 0, 1, 0) \text{ are not contained in } \text{Span} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & x & x^2 & x^3 \\ 1 & 2x & (2x)^2 & (2x)^3 \\ 1 & 3x & (3x)^2 & (3x)^3 \\ 1 & 4x & (4x)^2 & (4x)^3 \end{pmatrix} = \text{span} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{pmatrix}.)$$

Consider the Weyl system (Y, R) where $Y = X \times \mathbb{Z}_2$ with $X = \mathbb{T}^3$ and $R(x_1, x_2, x_3, i) = (x_1 + \alpha, x_2 + x_1, x_3 + x_2, i + 1)$ with an irrational α . For this system we have

$$\tau_3^5(H) = \text{Span} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & x & x^{[2]} & x^{[3]} \\ 1 & 2x & (2x)^{[2]} & (2x)^{[3]} \\ 1 & 3x & (3x)^{[2]} & (3x)^{[3]} \\ 1 & 4x & (4x)^{[2]} & (4x)^{[3]} \end{pmatrix} = \text{Span} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & x & x^2 & x^3 \\ 1 & 2x & (2x)^2 & (2x)^3 \\ 1 & 3x & (3x)^2 & (3x)^3 \\ 1 & 4x & (4x)^2 & (4x)^3 \end{pmatrix} = \text{span} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{pmatrix} \subset F_3^5 = \mathbb{T}^5.$$

The space $U = \text{span}\{u_0, u_1\} = \text{span} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \left\{ \begin{pmatrix} t_0 \\ t_1 \\ t_0 \\ t_1 \\ t_0 \end{pmatrix}, t_0, t_1 \in \mathbb{T} \right\}$ is not contained in

$\tau_3^5(H)$, thus we can choose $t_0, t_1 \in \mathbb{T}$ such that $\begin{pmatrix} t_0 \\ t_1 \\ t_0 \\ t_1 \\ t_0 \end{pmatrix}, \begin{pmatrix} t_1 \\ t_0 \\ t_1 \\ t_0 \\ t_1 \end{pmatrix} \notin \tau_3^5(H)$ (this is so when $t_0 \neq t_1$). Next we choose $\delta > 0$ such that for $I_0 = (t_0 - \delta, t_0 + \delta)$ and $I_1 = (t_1 - \delta, t_1 + \delta)$ one has

$$(I_0 \times I_1 \times I_0 \times I_1 \times I_0) \cap \tau_3^5(H) = \emptyset \quad \text{and} \quad (I_1 \times I_0 \times I_1 \times I_0 \times I_1) \cap \tau_3^5(H) = \emptyset.$$

For $A_0 = X_2 \times I_0 \subset X$ and $A_1 = X_2 \times I_1 \subset X$ we then have $(A_0 \times A_1 \times A_0 \times A_1 \times A_0) \cap H = \emptyset$ and $(A_1 \times A_0 \times A_1 \times A_0 \times A_1) \cap H = \emptyset$. Finally, we put $B = (A_0 \times \{0\}) \cup (A_1 \times \{1\}) \subset Y$.

Let now $K = \bigcup \overline{\begin{pmatrix} \text{Id}_Y \\ R^{2n+1} \\ R^{2(2n+1)} \\ R^{3(2n+1)} \\ R^{4(2n+1)} \end{pmatrix}} \Delta_{Y^{r+1}} \subset Y^5 = X^5 \times \mathbb{Z}_2^5$. Then $K = H_0 \cup H_1$ where $H_0 = H \times \{0, 1, 0, 1, 0\}$ and $H_1 = H \times \{1, 0, 1, 0, 1\}$. Since

$$\begin{aligned} B^5 \cap (X^5 \times \{0, 1, 0, 1, 0\}) &= (A_0 \times A_1 \times A_0 \times A_1 \times A_0) \times \{0, 1, 0, 1, 0\} \\ \text{and } B^5 \cap (X^5 \times \{1, 0, 1, 0, 1\}) &= (A_1 \times A_0 \times A_1 \times A_0 \times A_1) \times \{1, 0, 1, 0, 1\}, \end{aligned}$$

we obtain $B^5 \cap K = \emptyset$. This implies $B \cap R^{-(2n+1)}B \cap R^{-2(2n+1)}B \cap R^{-3(2n+1)}B \cap R^{-4(2n+1)}B = \emptyset$ for all $n \in \mathbb{Z}$.

Proof of Theorem 7.2. In the notation of Proposition 7.5, take any point $y \in Y$. Since (Y, R) is an ergodic Weyl system, it is uniquely ergodic and so, $UD(\{a \in \mathbb{Z} : R^a y \in C\}) = \mu_Y(C)$ for any open $C \subseteq Y$ with $\mu_Y(\partial C) = 0$. Define $E = \{a \in \mathbb{Z} : R^a y \in B\}$. Then $UD(E) = \mu_Y(B) > 0$, and for any $s \in \mathbb{N}$ and $a, c_1, \dots, c_s \in \mathbb{Z}$ one has $a, a + c_1, \dots, a + c_s \in E$ iff $R^a y, R^{a+c_1} y, \dots, R^{a+c_s} y \in B$ iff $R^a y \in B \cap R^{-c_1} B \cap \dots \cap R^{-c_s} B$.

Since $B \cap R^{-p_1(mn+l)} B \cap \dots \cap R^{-p_r(mn+l)} B = \emptyset$ for all $n \in \mathbb{Z}$ and $l \in \{l_1, \dots, l_\nu\}$, the set E does not contain configurations of the form $a, a + p_1(mn + l), \dots, a + p_r(mn + l)$.

Let $\{q_1, \dots, q_s\}$ be a system of integral polynomials with zero constant term. Let $W(q_1, \dots, q_s) \leq k$, let h be a nonconstant integral polynomial, and let $c_1, \dots, c_s \in \mathbb{Z}$. By Proposition 6.15

$$UC\text{-}\lim_n \mu_Y(B \cap R^{-q_1(h(n)) - c_1} B \cap \dots \cap R^{-q_s(h(n)) - c_s} B) = \mu_Y(B)^{s+1}.$$

For $n \in \mathbb{Z}$ put $E_n = \{a \in \mathbb{Z} : a, a + q_1(h(n)) + c_1, \dots, a + q_s(h(n)) + c_s \in E\}$. We have $UD(E_n) = \mu_Y(B \cap R^{-q_1(h(n)) - c_1} B \cap \dots \cap R^{-q_s(h(n)) - c_s} B)$, and so, $UC\text{-}\lim_n UD(E_n) = \mu_Y(B)^{s+1} = UD(E)^{s+1}$.

If $W(q_1, \dots, q_s) < k$, then by Proposition 6.17 we obtain $\lim_{n \rightarrow \infty} UD(E_n) = \mu_Y(B)^{s+1} = UD(E)^{s+1}$.

Finally, let $V(q_1, \dots, q_s) \leq k$, and let h be a nonconstant integral polynomial. Put $E_n = \{a \in \mathbb{Z} : a, a + q_1(h(n)), \dots, a + q_s(h(n)) \in E\}$, $n \in \mathbb{Z}$. By Proposition 6.18,

$$UC\text{-}\lim_n UD(E_n) = UC\text{-}\lim_n \mu_Y(B \cap R^{-q_1(h(n))} B \cap \dots \cap R^{-q_s(h(n))} B) > 0.$$

■

Bibliography

- [B] V. Bergelson, Weakly mixing PET, *Ergod. Th. and Dynam. Sys.* **7** (1987), 337–349.
- [BL] V. Bergelson and A. Leibman, Polynomial extensions of van der Waerden’s and Szemerédi’s theorems, *J. of AMS* **9** (1996), 725–753.
- [BM] V. Bergelson and R. McCutcheon, Uniformity in polynomial Szemerédi theorem, *Ergodic Theory of $\mathbb{Z}^d$ -actions*, London Math. Soc. Lecture Note Series **228** (1996), 273–296.
- [Fr] N. Frantzikinakis, Multiple ergodic averages for three polynomials and applications, preprint. (Arxiv, reference math.DS/0606567.)
- [FrK] N. Frantzikinakis and B. Kra, Polynomial averages converge to the product of integrals, *Israel J. of Math* **148** (2005), 267–276.
- [F1] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. d’Analyse Math.* **31** (1977), 204–256.
- [F2] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton Univ. Press, 1981.
- [F3] H. Furstenberg, Nonconventional ergodic averages, *Proceedings of Symposia in Pure Math.* **50** (1990), 43–56.
- [HK1] B. Host and B. Kra, Non-conventional ergodic averages and nilmanifolds, *Annals of Math.* **161** (2005), no. 1, 397–488.

- [HK2] B. Host and B. Kra, Convergence of polynomial ergodic averages, *Israel J. of Math* **149** (2005), 1–19.
- [L1] A. Leibman, Convergence of multiple ergodic averages along polynomials of several variables, *Israel J. of Math* **146** (2005), 303–322.
- [L2] A. Leibman, Host-Kra and Ziegler factors, and convergence of multiple averages, appendix to “Combinatorial and Diophantine Applications of Ergodic Theory” by V. Bergelson, *Handbook of Dynamical Systems*, vol. 1B, B. Hasselblatt and A. Katok, eds., Elsevier, pp. 841–853.
- [Sz] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 199–245.
- [We] H. Weyl, Über die Gleichverteilung von Zahlen mod Eins, *Math. Ann.* **77** (1916), 313–352.
- [Z] T. Ziegler, Universal characteristic factors and Furstenberg averages, *J. Amer. Math. Soc.* **20** (2007), no. 1, 53–97.

V. Bergelson, Department of Mathematics, the Ohio State University, OH 43210, USA.
e-mail: vitaly@math.ohio-state.edu

A. Leibman, Department of Mathematics, the Ohio State University, OH 43210, USA.
e-mail: leibman@math.ohio-state.edu

E. Lesigne, Laboratoire de Mathématiques et Physique Théorique, UMR CNRS 6083.
 Université François Rabelais, Tours, France. *e-mail:* emmanuel.lesigne@univ-tours.fr