

Orbit of the diagonal in the power of a nilmanifold

A. Leibman

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

August 31, 2009

Abstract

Let X be a nilmanifold, that is, a compact homogeneous space of a nilpotent Lie group G , and let $a \in G$. We study the closure of the orbit of the diagonal of X^r under the action $(a^{p_1(n)}, \dots, a^{p_r(n)})$, where p_i are integer-valued polynomials in m integer variables. (Knowing this closure is crucial for finding limits of the form $\lim_{N \rightarrow \infty} \frac{1}{N^m} \sum_{n \in \{1, \dots, N\}^m} \mu(T^{p_1(n)} A_1 \cap \dots \cap T^{p_r(n)} A_r)$ where T is a measure preserving transformation of a finite measure space (Y, μ) and A_i are subsets of Y , and limits of the form $\lim_{N \rightarrow \infty} \frac{1}{N^m} \sum_{n \in \{1, \dots, N\}^m} d((A_1 + p_1(n)) \cap \dots \cap (A_r + p_r(n)))$ where A_i are subsets of \mathbb{Z} and $d(A)$ is the density of A in \mathbb{Z} .) We give a simple description of the closure of the orbit of the diagonal in the case all p_i are linear, in the case G is connected, and in the case the identity component of G is commutative; in the general case our description of the orbit is not explicit.

0. Introduction

0.1. *Multiple ergodic averages* $M_N = \frac{1}{N} \sum_{n=1}^N \mu_X(A \cap T^{-n} A \cap \dots \cap T^{-rn} A)$, where T is a measure preserving transformation of a probability measure space (X, \mathcal{B}, μ_X) and $A \in \mathcal{B}$, $\mu_X(A) > 0$, had appeared in Furstenberg's proof of Szemerédi's theorem ([Fu]). Furstenberg proved in [Fu] that $\liminf_{N \rightarrow \infty} M_N > 0$; the question whether $\lim_{N \rightarrow \infty} M_N$ exists remained open for a long time. A way of solving this problem was already outlined in [Fu]: one has to find a *characteristic factor* of (X, T) with respect to the system of actions $\{\text{Id}_X, T^n, \dots, T^{rn}\}$, $n \in \mathbb{Z}$, that is, a factor (X', T) of (X, T) such that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \left| \int_X f_0 T^n f_1 \dots T^{rn} f_r d\mu_X - \int_{X'} E(f_0|X') T^n E(f_1|X') \dots T^{rn} E(f_r|X') d\mu_{X'} \right| = 0$$

for any $f_0, f_1, \dots, f_r \in L^\infty(X)$. If the system (X', T) has a simple enough structure so that the limits $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \int_{X'} h_0 T^n h_1 \dots T^{rn} h_r d\mu_{X'}$ with $h_i \in L^\infty(X')$ can be easily determined, then $\lim_{N \rightarrow \infty} M_N$ will also be found by taking $h_i = E(1_A|X')$, $i = 0, 1, \dots, r$.

Supported by NSF grants DMS-0345350 and DMS-0600042.

2000 Mathematical Subject Classification: Primary 37C99; Secondary 22E25.

0.2. A *d*-step nilsystem is a measure preserving system realized by a translation on a *d*-step nilmanifold, a compact homogeneous space of a *d*-step nilpotent Lie group equipped with the Haar measure. A *d*-step pro-nilsystem is the inverse limit of a sequence of *d*-step nilsystems. It was proven by Host and Kra in [HK1] and, independently, by Ziegler in [Z2] that any ergodic probability measure preserving system (X, T) possesses a factor characteristic with respect to $\{\text{Id}_X, T^n, \dots, T^{rn}\}$ and isomorphic to an $(r - 1)$ -step pro-nilsystem. It was then shown in [HK2] and [L5] that pro-nilsystems are also characteristic for any system of “polynomial actions” of the form $\{T^{p_1(n)}, \dots, T^{p_r(n)}\}$, $n \in \mathbb{Z}^m$, where p_i are integer-valued polynomials on \mathbb{Z}^m . Thus, the problem of identifying the limit of polynomial multiple ergodic averages $\frac{1}{N} \sum_{n=1}^N \mu_X(T^{-p_1(n)}A \cap \dots \cap T^{-p_r(n)}A)$ is reduced to the case where (X, T) is a nilsystem.

0.3. Let (X, a) be a nilsystem, that is, X is a nilmanifold and a is a translation on X , and let μ_X be the Haar measure on X . Let p_1, \dots, p_r be integer-valued polynomials on \mathbb{Z}^m . It is proved in [L4] that under any polynomial action φ of \mathbb{Z}^m on X , the closure $Y = \overline{\text{Orb}}(D)$ of the orbit $\text{Orb}_D = \{\varphi(n)D\}_{n \in \mathbb{Z}^m}$ of a subnilmanifold D of X is either a subnilmanifold or a finite union of subnilmanifolds of X . Moreover, “the sequence” $\{\varphi(n)D\}_{n \in \mathbb{Z}^m}$ is well distributed on Y , that is, $\text{C-lim}_n \varphi(n)\mu_D = \mu_Y$ where μ_D is the Haar measure on D , μ_Y is a Haar measure on Y (more exactly, a linear combination of the Haar measures on the connected components of Y), and $\text{C-lim}_n = \lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{n \in \Phi_N}$ where $\{\Phi_N\}$ is any Følner sequence in \mathbb{Z}^m . This, in particular, is applicable to the closure $Y = \overline{\text{Orb}}_P(\Delta_{X^r})$ of the orbit of the diagonal $\Delta_{X^r} = \left\{ \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix} : x \in X \right\}$ under the polynomial action

$$\varphi(n) \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} a^{p_1(n)}x_1 \\ \vdots \\ a^{p_r(n)}x_r \end{pmatrix}, \quad n \in \mathbb{Z}^m, \quad (0.1)$$

of \mathbb{Z}^m on the nilmanifold X^r . It follows that for any $f_1, \dots, f_r \in C(X)$,

$$\begin{aligned} \text{C-lim}_n \int_X a^{p_1(n)}f_1 \cdot \dots \cdot a^{p_r(n)}f_r d\mu_X &= \text{C-lim}_n \int_{\Delta_{X^r}} a^{p_1(n)}f_1 \otimes \dots \otimes a^{p_r(n)}f_r d\mu_{\Delta_{X^r}} \\ &= \text{C-lim}_n \int_{\Delta_{X^r}} \varphi(n)(f_1 \otimes \dots \otimes f_r) d\mu_{\Delta_{X^r}} = \text{C-lim}_n \int_{\varphi(n)\Delta_{X^r}} f_1 \otimes \dots \otimes f_r d\mu_{\varphi(n)\Delta_{X^r}} \\ &= \int_Y f_1 \otimes \dots \otimes f_r d\mu_Y, \end{aligned}$$

where μ_Y is the Haar measure on $Y = \overline{\text{Orb}}_P(\Delta_{X^r})$. It then follows that

$$\text{C-lim}_n \int_X a^{p_1(n)}f_1 \cdot \dots \cdot a^{p_r(n)}f_r d\mu_X = \int_Y f_1 \otimes \dots \otimes f_r d\mu_Y \quad (0.2)$$

for any $f_1, \dots, f_r \in L^\infty(X)$. Hence, for any measurable sets $A_1, \dots, A_r \subseteq X$,

$$\text{C-lim}_n \mu_X(a^{-p_1(n)}A_1 \cap \dots \cap a^{-p_r(n)}A_r) = \mu_Y((A_1 \times \dots \times A_r) \cap Y),$$

and thus the problem of evaluating $\lim_{N \rightarrow \infty} \frac{1}{|\Phi_N|} \sum_{n \in \Phi_N} \mu_X(a^{-p_1(n)}A_1 \cap \dots \cap a^{-p_r(n)}A_r)$ is reduced to the problem of finding $Y = \overline{\text{Orb}}_P(\Delta_{X^r})$.

0.4. Once $\overline{\text{Orb}}_P(\Delta_{X^r})$ with respect to the action defined by a system $P = \{p_1, \dots, p_r\}$ of integer-valued polynomials has been found, it is immediately possible to determine the “optimal” characteristic factor corresponding to P . Any nilmanifold X has a natural tower of factors, $X = X_s \rightarrow X_{s-1} \rightarrow \dots \rightarrow X_1 \rightarrow X_0 = \{\cdot\}$, where for each d , X_d is a d -step nilmanifold. It is of interest to learn what is the minimal d for which X_d is characteristic for P . In terms of the orbit $\overline{\text{Orb}}_P(\Delta_{X^r})$ of the diagonal in X^r , characteristic factors of X are easily characterizable: X_d is characteristic for the action defined by P iff $\overline{\text{Orb}}_P(\Delta_{X^r})$ contains the fiber of the projection $X \rightarrow X_d$.

If $d \in \mathbb{N}$ is such that the factor X_d is characteristic for P for *any* nilsystem (X, a) , we call d *the complexity of P* ; by [HK1] and [Z2], if d is the complexity of P , then *any* ergodic measure preserving system (X, T) has a factor characteristic for P that is isomorphic to a d -step pro-nilsystem.

Remark. The complexity, as well as the W-complexity (Weyl complexity) of P that we introduce in this paper differ from those in [BLLe1]: the complexities defined here are less by 1 than the corresponding complexities in [BLLe1].

Knowing $\overline{\text{Orb}}_P(\Delta_{X^r})$, one can also determine conditions on the sets A_1, \dots, A_r to guarantee that $\mu_X(a^{-p_1(n)}A_1 \cap \dots \cap a^{-p_r(n)}A_r)$ is positive, or is greater than a certain constant, for some n . Examples of such application of $\overline{\text{Orb}}_P(\Delta_{X^r})$ can be found in [BHK], [BLLe1], [FK], or [F].

0.5. Here is a list of earlier results.

$\overline{\text{Orb}}_P(\Delta_{X^r})$ for the “linear” system $P = \{0, n, \dots, rn\}$ was first found in [Z1], and, in a much simpler way, in [BHK]. For the system $P = \{n_1, \dots, n_l, n_1+n_2, \dots, n_1+\dots+n_l\}$ of linear functions in l integer variables n_1, \dots, n_l consisting of all sums of distinct variables, $\overline{\text{Orb}}_P(\Delta_{X^r})$ was described in [HK1]. The complexity of any system of $r+1$ (distinct) linear functions in one variable $\{0, c_1n, \dots, c_rn\}$ was shown to be $\leq r$ in [HK1] and $= r-1$ in [Z2]; the complexity of any system of $r+1$ linear functions in several variables was proven to be $\leq r-1$ in [L6].

In [FK], the complexity of a system of linearly independent polynomials in one variable was shown to be 0 with respect to any totally-ergodic system and ≤ 1 in the general case.

In [BLLe1], $\overline{\text{Orb}}_P(\Delta_{X^r})$ for a polynomial system $P = \{0, p_1(n), \dots, p_r(n)\}$ was found in the case where (X, a) is a Weyl system, that is, when X is a torus and a is a unipotent linear transformation of X .

In [F], $\overline{\text{Orb}}_P(\Delta_{X^r})$ was found in the case where (X, a) is a general nilsystem and P is a system with three nonzero polynomials, $P = \{0, p_1, p_2, p_3\}$.

A finitary analog of the problem of determining the complexity of a system of linear forms is considered in [GW]. (In [GW], G is taken to be a finite group, specifically \mathbb{Z}_p^n for an odd prime p , and, given a system $P = \{p_1, \dots, p_r\}$ of linear forms $G^m \rightarrow G$, it is asked what the degree of “uniformity” of a subset $A \subseteq G$ has to be in order to guarantee that it contains approximately “the right number”, $|A|^r |G|^{m-r}$, of configurations of the form $\{p_1(n), \dots, p_r(n)\}$, $n \in G^m$.) The case of systems of complexity 2 is studied and the obtained result agrees with the results of the present paper.

0.6. Our goal in this paper is to find $\overline{\text{Orb}}_P(\Delta_{X^r})$ for a system $P = \{p_1, \dots, p_r\}$ of polynomials in several variables. It is natural to confine ourselves to ergodic transformations, and thus assume that the translation a of the nilmanifold X is ergodic. We will go farther and assume throughout the paper that X is connected; this implies that the action of a is totally ergodic on X . Doing this, we do not lose much generality. Indeed, if X is disconnected, it consists of several connected components, permuted by a ; a power a^k of a preserves the components, and after replacing a by a^k we may treat each component of X individually.

0.7. Let $\pi: G \rightarrow X$ be the natural projection; we will denote by $\pi^{\times r}$ the corresponding projection $G^r \rightarrow X^r$. “To find $\overline{\text{Orb}}_P(\Delta_{X^r})$ ” means to find a subgroup H of G^r (or a finite union of right cosets of H , – if $\overline{\text{Orb}}_P(\Delta_{X^r})$ is disconnected) such that $\pi^{\times r}(H) = \overline{\text{Orb}}_P(\Delta_{X^r})$. In the case where the polynomials $p_j: \mathbb{Z}^m \rightarrow \mathbb{Z}$ are linear we give a simple explicit description of such H . (See Theorem 6.3). Let $p_j(n_1, \dots, n_m) = \sum_{i=1}^m c_{j,i} n_m$, $j = 1, \dots, r$; put $v_0 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ and $v_i = \begin{pmatrix} c_{1,i} \\ \vdots \\ c_{r,i} \end{pmatrix}$, $i = 1, \dots, m$. For two vectors $u = \begin{pmatrix} c_1 \\ \vdots \\ c_r \end{pmatrix}$ and $v = \begin{pmatrix} e_1 \\ \vdots \\ e_r \end{pmatrix}$ in \mathbb{Z}^r we define the product $uv = \begin{pmatrix} c_1 e_1 \\ \vdots \\ c_r e_r \end{pmatrix}$. Let \hat{V} be the subgroup of \mathbb{Z}^r generated by v_0, v_1, \dots, v_r . For each $k \in \mathbb{N}$ let $\text{Pol}_k(\mathbb{Z}^m)$ be the space of polynomials $\mathbb{Z}^m \rightarrow \mathbb{Z}$ (with rational coefficients) of degree $\leq k$, and let $\mathcal{P}_k \hat{V}$ be the subgroup of \mathbb{Z}^r given by $\mathcal{P}_k \hat{V} = \{q(v_1, \dots, v_m), q \in \text{Pol}_k(\hat{V})\}$. For a vector $u = \begin{pmatrix} c_1 \\ \vdots \\ c_r \end{pmatrix} \in \mathbb{Z}^r$ and an element $b \in G$ we define $b^u = \begin{pmatrix} b^{c_1} \\ \vdots \\ b^{c_r} \end{pmatrix} \in G^r$. Let $G = G_1 \supset G_2 \supset \dots \supset G_{s+1} = \{\mathbf{1}_G\}$ be the lower central series of G . We define H as the subgroup of G^r generated by Δ_{G^r} and the elements of the form b^v with $b \in G_k$ and $v \in \mathcal{P}_k \hat{V}$ for some $k \in \mathbb{N}$. If, for each $k = 1, \dots, s$, $\{q_{k,1}, \dots, q_{k,l_k}\}$ is a set of generators in $\mathcal{P}_k \hat{V}$, then H can also be defined in the following way:

$$H = \left\{ b_0^{v_0} \prod_{k=1}^s \prod_{i=1}^{l_k} b_{k,i}^{q_{k,i}(v_1, \dots, v_m)}, \quad b_{k,i} \in G_k \text{ for every } k, i \right\}. \quad (0.3)$$

(In the case $m = 1$, $p_j(n) = jn$, $j = 1, \dots, r$, we have $v = \begin{pmatrix} 1 \\ \vdots \\ r \end{pmatrix}$ and $H = \{b_0 \prod_{k=1}^s b_k^{(v)}\}$, $b_0 \in G$, $b_k \in G_k$, $k = 1, \dots, s$ }, which coincides with the result in [Z1].)

Our proof of the fact that $\pi^{\times r}(H) = \overline{\text{Orb}}_P(\Delta_{X^r})$ is easy; it consists of three parts. First, we show that the subset H defined by (0.3) is a subgroup of G^r such that the set $\pi^{\times r}(H)$ is closed in X^r . (We say that H is a *rational* subgroup of G^r .) Then we notice that H contains Δ_{G^r} and all elements of the form $\begin{pmatrix} a^{p_1(n)} \\ \vdots \\ a^{p_r(n)} \end{pmatrix}$, thus $\pi^{\times r}(H) \supseteq \overline{\text{Orb}}_P(\Delta_{X^r})$. The nilmanifold $X_1 = G_2 \backslash X$ is a torus on which G acts by rotations, thus it is easy to find the closure $\overline{\text{Orb}}_P(\Delta_{X_1^r})$ of the orbit of the diagonal of X_1^r under the action $\begin{pmatrix} a^{p_1(n)} \\ \vdots \\ a^{p_r(n)} \end{pmatrix}$: this is the projection to X_1^r of the subgroup $(G/G_2)^{[V]} = \{b^v, b \in G/G_2, v \in V\}$ of $(G/G_2)^r$. Comparing dimensions, we show that H is a minimal rational subgroup of G^r with the property that the group H/G_2^r contains the identity component of the preimage

of $(G/G_2)^{[V]}$ in $(G/G_2)^r$ and the elements $a^v G_2$, $v \in \hat{V}$. It follows that $\pi^{\times r}(H) = \overline{\text{Orb}}_P(\Delta_{X^r})$.

Using the above description of H , we see that the complexity of P is the minimal integer d for which $\mathcal{P}_{d+1}\hat{V} \otimes \mathbb{R} = \mathbb{R}^r$. As it is noticed in [GW], this is equivalent to saying that d is the minimal integer such that the polynomial vectors $(1, p_1, \dots, p_1^{d+1}), \dots, (1, p_r, \dots, p_r^{d+1})$ are linearly independent.

0.8. When p_j are not necessarily linear but the group G is connected (or, rather, a is contained in the identity component of G), the description of H remains as simple as in the linear case, – with distinct monomials playing the role of independent variables. (See Theorem 8.3). We simply represent $p_j = \sum_{i=1}^m c_{j,i} u_i$, $c_{j,i} \in \mathbb{Q}$, where u_1, \dots, u_m are distinct (nonconstant) monomials (with trivial coefficient), put V to be the subspace of \mathbb{R}^r spanned by the vectors $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ and $v_i = \begin{pmatrix} c_{1,i} \\ \vdots \\ c_{r,i} \end{pmatrix}$, $i = 1, \dots, m$, and use V instead of \hat{V} . The complexity of a polynomial system P with respect to nilsystems corresponding to connected G can therefore be easily computed; we call it the *C-complexity* of P .

0.9. For a polynomial system P and disconnected G we first consider the case where the identity component \mathcal{G} of G is commutative. In this case $X = \mathcal{G}/(\Gamma \cap \mathcal{G})$ is a torus on which G acts by skew-product transformations. We call the system (X, a) a *Weyl system*. (The 2-dimensional torus with the transformation $a(x, y) = (x + \alpha, y + x)$ is the simplest example of such a system.) For Weyl systems the problem of determining $\overline{\text{Orb}}_P(\Delta_{X^r})$ is an easy linear algebra problem; but the answer we obtain is not elegant (see Theorem 9.4). Based on Theorem 9.7, the complexity of any polynomial system P with respect to Weyl systems (we call it the *W-complexity* of P) can, in principle, be computed, but these computations are cumbersome, and leave some questions unanswered. In particular, it is not even clear to us whether the complexity of a polynomial system P is always $\leq |P| - 2$.

0.10. The situation with the Weyl systems demonstrates that a simple description of $\overline{\text{Orb}}_P(\Delta_{X^r})$ is hardly possible in the general case, – when P is a polynomial system and no assumptions about G have been made. We attempt to write out a formula defining the subgroup H of \mathcal{G}^r for which $\pi^{\times r}(H) = \overline{\text{Orb}}_P(\Delta_{X^r})$, similar to that for Weyl systems (see subsections 11.2–11.4). However, for this formula to work it must be that H is “defined by its linear part”, as it is when P is a system of linear functions, and we don’t know whether this is so (see Conjecture 11.4); in addition, the polynomials involved in this formula are not explicit and are only defined recurrently. We do not also know whether $\overline{\text{Orb}}_P(\Delta_{X^r})$ is always connected; these question can also be resolved by affirming or refuting Conjecture 11.4.

We also describe a method of finding $\overline{\text{Orb}}_P(\Delta_{X^r})$ for concrete (X, a) and P (see subsection 11.10); it is based on successive construction of the orbits of the diagonal on certain factor-tori of certain subgroups of G . In principle, this method can be practically used to find $\overline{\text{Orb}}_P(\Delta_{X^r})$ when the complexity of P is relatively small, but such a computation for a concrete system does not look to be of any interest. It can also be utilized to establish some properties of $\overline{\text{Orb}}_P(\Delta_{X^r})$. (We do this in [BLLe2] to prove that if the polynomials constituting P have nonzero constant terms but are jointly intersective, then $\overline{\text{Orb}}_P(\Delta_{X^r})$

contains Δ_{X^r} .)

Finally, we obtain a rough estimate of the complexity of P (see Theorem 13.13). As a corollary we get that it does not exceed $c_W c(c+1)/2$, where c_W and c are the W-complexity and the C-complexity of P respectively. (We, actually, suspect that the complexity of P is always equal to c_W , but cannot confirm this conjecture either.)

0.11. We only deal with “discrete” linear and polynomial actions, that is, actions of the group \mathbb{Z}^m . The obtained results are also applicable to continuous (polynomial) actions of the group \mathbb{R}^m . Moreover, in the case of continuous actions the Lie group G acting on X can be assumed to be connected, which essentially simplifies the situation, and the case of polynomial actions of \mathbb{R}^m is completely covered by Section 8.

0.12. Here is the plan of the paper: Sections 1 and 2 are preparatory; we consider there, in more detail, properties of the orbits of points and of subnilmanifolds in a nilsystem, characteristic factors of a nilsystem, and relations between them. In Section 3 we introduce some algebraic notation. In Sections 4 and 5 we introduce the group $G^{[V]}$. In Section 6 we find $\overline{\text{Orb}}_P(\Delta_{X^r})$ in the case where P is a system of linear polynomials. In Section 7 we remind the reader how polynomial orbits on a torus look. In Section 8 we find $\overline{\text{Orb}}_P(\Delta_{X^r})$ for a general polynomial system P in the case where the Lie group G acting on X is connected. In Section 9 we find $\overline{\text{Orb}}_P(\Delta_{X^r})$ for the case of a Weyl system. (The results of this section partially repeat results from [BLLe1].) In Section 10 we describe how the closure of the orbit of a point of a nilmanifold under a polynomial action can be found. In Section 11 we describe a method of constructing $\overline{\text{Orb}}_P(\Delta_{X^r})$ for a general nilsystem. In Section 12 we introduce a filtration in the group \mathcal{G} that we use in the next section. In Section 13 we find an estimate of the complexity of a general polynomial system.

1. Orbits in a nilmanifold

In this section we introduce some notation and collect some facts related to nilmanifolds. For more details see [M], [L2], [L4], [L7] and [L8].

1.1. Throughout the paper, X will be a connected *nilmanifold*, that is, a connected compact homogeneous space of a nilpotent Lie group. X can be represented as a factor $X = G/\Gamma$, where G is a simply-connected (not necessarily connected) nilpotent Lie group and Γ is a discrete uniform subgroup of G . The group G acts on X by left translations: for $a \in G$ and $x \in X$ we have $ax \in X$.

By π we will denote the natural projection $G \longrightarrow X$, and by 1_X the point $\pi(\mathbf{1}_G) \in X$.

1.2. We will denote by \mathcal{G} the connected component of G . (We have chosen such an unusual symbol instead of the standard G^c or G° in order to be able to attach sub- and superscripts to \mathcal{G} , like \mathcal{G}_d or \mathcal{G}^r .) Since X is connected, it is a homogeneous space of \mathcal{G} as well, $X = \mathcal{G}/(\mathcal{G} \cap \Gamma)$.

Here is an example of a disconnected nilpotent Lie group G with a connected nilmanifold $X = G/\Gamma$. Let $G = \left\{ \begin{pmatrix} 1 & k & y \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix}, k \in \mathbb{Z}, x, y \in \mathbb{R} \right\}$ and $\Gamma = \left\{ \begin{pmatrix} 1 & k & l \\ 0 & 1 & m \\ 0 & 0 & 1 \end{pmatrix}, k, l, m \in \mathbb{Z} \right\}$.

Then G is disconnected, with $\mathcal{G} = \left\{ \begin{pmatrix} 1 & 0 & y \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix}, x, y \in \mathbb{R} \right\}$ isomorphic to \mathbb{R}^2 . The nilmanifold $X = G/\Gamma = \mathcal{G}/(\mathcal{G} \cap \Gamma)$ is isomorphic to the 2-dimensional torus $\mathbb{R}^2/\mathbb{Z}^2$, with G acting on X by skew-product transformations: $\begin{pmatrix} 1 & k & b \\ 0 & 1 & a \\ 0 & 0 & 1 \end{pmatrix}(x, y) = (x + a, y + kx + b)$.

1.3. The connected simply-connected nilpotent group \mathcal{G} is exponential, which implies that for every $b \in \mathcal{G}$ there exists a unique homomorphism $\mathbb{R} \rightarrow \mathcal{G}$, $t \mapsto b^t$, such that $b^1 = b$. We will constantly use this notation.

1.4. The commutator of two elements $b_1, b_2 \in G$ is $[b_1, b_2] = b_1^{-1}b_2^{-1}b_1b_2$. For $b_1, b_2, \dots, b_d \in G$ we put $[b_1, b_2, \dots, b_d] = [\dots [[b_1, b_2], b_3], \dots], b_d]$.

If B_1, B_2, \dots, B_d are subsets of G , we denote by $[B_1, B_2, \dots, B_d]$ the subgroup of G generated by the elements $[b_1, b_2, \dots, b_d]$ with $b_1 \in B_1, b_2 \in B_2, \dots, b_d \in B_d$.

1.5. We will denote by G_d , $d \in \mathbb{N}$, the terms of the lower central series of G , $G_1 = G$, $G_2 = [G, G]$, $G_3 = [G, G, G]$, \dots . If G is s -step nilpotent, $G_{s+1} = \{\mathbf{1}_G\}$. The groups G_d are normal closed Lie subgroups of G , and for each d , G_d/G_{d-1} is a commutative simply-connected Lie group.

The terms of the lower central series of \mathcal{G} will be denoted by \mathcal{G}_d , $d \in \mathbb{N}$.

1.6. When A is a group and C is a normal subgroup of A , for an element $b \in A$ or for a subset $B \subseteq A$ we will denote by $b \bmod C$, respectively by $B \bmod C$, the image of b , respectively of B , in A/C . We will write $b \in B \bmod C$ if $b \bmod C \in B \bmod C$; $B_1 \subseteq B_2 \bmod C$ if $B_1 \bmod C \subseteq B_2 \bmod C$, etc.

1.7. One can find a Malcev basis in G compatible with Γ and with the lower central series of G , namely, elements $e_{d,j} \in \Gamma$, $d = 1, \dots, s$, $j = 1, \dots, k_d$, such that for each $d = 1, \dots, s$, the elements $e_{d,j} \bmod G_{d+1}$, $j = 1, \dots, k_d$, form a basis in G_d/G_{d+1} . Every element $b \in \mathcal{G}$ is then uniquely representable in the form $b = \prod_{d=1}^s \prod_{j=1}^{k_d} e_{d,j}^{u_{d,j}}$ where the coordinate $u_{i,j} \in \mathbb{R}$ if $e_{i,j} \in \mathcal{G}$ and $u_{i,j} \in \mathbb{Z}$ otherwise. Thus, Malcev coordinates define a diffeomorphism $G \simeq \mathbb{Z}^m \times \mathbb{R}^l$.

1.8. The multiplication in G , written in Malcev coordinates, is polynomial: if $b = \prod_{d=1}^s \prod_{j=1}^{k_d} e_{d,j}^{u_{d,j}}$ and $c = \prod_{d=1}^s \prod_{j=1}^{k_d} e_{d,j}^{v_{d,j}}$, then $bc = \prod_{d=1}^s \prod_{j=1}^{k_d} e_{d,j}^{w_{d,j}}$ where for each d, j , the coordinate $w_{d,j}$ is a polynomial in the variables $u_{d',j'}$ and $v_{d'',j''}$ with $d', d'' \leq d$, taking on integer values when the arguments are integer, and having total degree d if it is assumed that each variable $u_{d',j'}$ has degree d' and each variable $v_{d'',j''}$ has degree d'' . (In the case of connected G this follows from Campbell-Hausdorff's formula; a disconnected G can be considered as a subgroup of a connected nilpotent Lie group.)

1.9. A subnilmanifold Y of X is a closed subset of X of the form $Y = Hx$, where H is a closed subgroup of G and $x \in X$; one may always assume that $H \subseteq \mathcal{G}$. A subnilmanifold Y is a nilmanifold, since $Y \simeq H/((b\Gamma b^{-1}) \cap H)$ where b is any element of G with $\pi(b) = x$. We will denote by μ_Y the normalized Haar measure on Y .

Given a closed subgroup H of G and a point $x \in X$, the set Hx may not be closed and so, be a subnilmanifold of X ; Hx is closed iff $(b\Gamma b^{-1}) \cap H$ is a uniform subgroup of

H , where b is any element of $\pi^{-1}(x)$. In particular, $H1_X = \pi(H)$ is closed iff $H \cap \Gamma$ is uniform in H ; we will say that H is *rational* in this case.

1.10. Any mapping $\varphi: \mathbb{Z}^m \rightarrow G$ defines an “action” $x \mapsto \varphi(n)x$, $n \in \mathbb{Z}^m$, $x \in X$, of \mathbb{Z}^m on X . (This notion of action we use is not too conventional.) For a point $x \in X$ the orbit of x under φ is the set $\text{Orb}_\varphi x = \{\varphi(n)x : n \in \mathbb{Z}^m\}$; for a subset $D \subseteq X$ the orbit of D under φ is $\text{Orb}_\varphi D = \bigcup_{n \in \mathbb{Z}^m} \varphi(n)D$.

1.11. If φ is a homomorphism, we say that φ is a *linear action* of \mathbb{Z}^m on X . Under a linear action, the closure $Z = \overline{\text{Orb}_\varphi(x)}$ of the orbit of any point $x \in X$ is a subnilmanifold of X , and “the sequence” $\varphi(n)x$ is well distributed on Z , $\text{C-lim}_n f(\varphi(n)x) = \int_Z f d\mu_Z$ for any $f \in C(Z)$. (See, for example, [L3].) If D is a subnilmanifold of X , the closure $Y = \overline{\text{Orb}_\varphi(D)}$ of the orbit $\text{Orb}_\varphi(D) = \bigcup_{n \in \mathbb{Z}^m} \varphi(n)D$ of D is also a subnilmanifold of X , and “the sequence” $\varphi(n)D$ is well distributed on Y : $\text{C-lim}_n \int_{\varphi(n)D} f d\mu_{\varphi(n)D} = \int_Y f d\mu_Y$ for any $f \in C(Y)$.

1.12. A mapping $\varphi: \mathbb{Z}^m \rightarrow G$ is *polynomial*, and defines a *polynomial action* of \mathbb{Z}^m on X , if it has the form $\varphi(n) = \prod_{i=1}^k b_i^{p_i(n)}$, $n \in \mathbb{Z}^m$, where $b_i \in G$ and p_i are polynomials on \mathbb{Z}^m . (When writing such an expression we always assume that it makes sense, that is, that the polynomial p_i takes only integer values whenever $b_i \notin \mathcal{G}$.)

We will call a closed subset of X of the form $\bigcup_{j=1}^l Hx_j$ where $x_1, \dots, x_l \in X$ and H is a closed subgroup of G , a *fu-subnilmanifold* of X .

Under a polynomial action, the closure $\overline{\text{Orb}_\varphi(x)}$ of the orbit of a point $x \in X$ is a fu-subnilmanifold of X , and $\varphi(n)x$ is well distributed on the connected components of $\overline{\text{Orb}_\varphi(x)}$. (See [L4]. This fact is in complete analogy with Weyl’s theorem about distribution of polynomial sequences on tori; see subsection 7.2.) The simplest example where the orbit of a polynomial sequence is not a subtorus but a finite union of subtori is given by the sequence $\{\frac{1}{3}n^2\}$ in the torus $X = \mathbb{R}/\mathbb{Z}$.

For a subnilmanifold D of X an analogous fact holds: $\overline{\text{Orb}_\varphi(D)}$ is a fu-subnilmanifold of X , and $\varphi(n)D$ is well distributed on $\overline{\text{Orb}_\varphi(D)}$.

1.13. Let φ be a linear action of \mathbb{Z}^m on X and let $x \in X$; for simplicity, let $x = 1_X$. Then $Y = \overline{\text{Orb}_\varphi(x)}$ is a (not necessarily connected) subnilmanifold of X , and has the form $Y = \pi(H)$ where H is a rational subgroup of G . The maximal subgroup of G with this property is $M = \{b \in G : bY = Y\}$; since Y is $\varphi(\mathbb{Z}^m)$ invariant, we have $\varphi(\mathbb{Z}^m) \subseteq M$. Thus, in order to determine Y , we may look for the minimal rational subgroup H of G that contains $\varphi(\mathbb{Z}^m)$.

Similarly, if D is a subnilmanifold of X with $1_X \in D$, then D has the form $D = \pi(K)$ where K is a rational subgroup of G , and in order to determine $\overline{\text{Orb}_\varphi(D)}$ we have to find the minimal rational subgroup H of G that contains both K and $\varphi(\mathbb{Z}^m)$.

1.14. If φ is a polynomial action, the problem of finding $Y = \overline{\text{Orb}_\varphi(x)}$ or $Y = \overline{\text{Orb}_\varphi(D)}$ is more difficult. In this case Y may not be a subnilmanifold of X , but only a fu-subnilmanifold. And even in the case where Y is a connected subnilmanifold, the elements $\varphi(n)$ may not preserve Y , and hence the group $H = \{b \in G : bY = Y\}$ may not contain

$\varphi(\mathbb{Z}^m)$. (See [L8].) Thus, even in this case (and assuming $\varphi(0) = \mathbf{1}_G$, $x = 1_X$, $D \ni 1_X$), in order to find Y we need to find the minimal rational subgroup H of G such that $\pi(H)$ contains $\varphi(n)x$, or respectively $\varphi(n)D$, for all $n \in \mathbb{Z}^m$.

1.15. Let φ be a polynomial action of \mathbb{Z}^m on X and let D be a connected subnilmanifold of X . Then almost all points of D have the same orbit under φ in the following sense: there is a fu-subnilmanifold Z such that for almost all $x \in D$, $\overline{\text{Orb}}_\varphi(x) = bZ$ where $b \in \pi^{-1}(x)$, and $\overline{\text{Orb}}_\varphi(x) \subseteq bZ$, $b \in \pi^{-1}(x)$, for the other points $x \in D$. We will call Z *the generic orbit* of points of D under the action φ . (Let us also remark that in the case where φ is a linear action, Z is a subnilmanifold and $\overline{\text{Orb}}_\varphi(x) = Z$ for all $x \in D$.)

Assume that $Z \ni 1_X$, let Z^c be the identity component of Z and let K be the connected rational subgroup of G such that $\pi(K) = Z^c$. Assume that $D \ni 1_X$, let L be the connected rational subgroup of G such that $\pi(L) = D$. Assume that $\varphi(0) = \mathbf{1}_G$, let $Y = \overline{\text{Orb}}_\varphi(D)$, let Y^c be the identity component of Y and let H be the connected rational subgroup of G such that $\pi(H) = Y^c$. Then K is a normal subgroup of H , $H = LK$, and thus $Y^c = LZ^c$. (See [L8].)

2. Characteristic factors, natural factors of nilmanifolds, and complexity

In this section we define characteristic factors of dynamical systems related to a system of polynomials P and the complexity of P .

2.1. We fix $a \in G$ such that the action of a on X is ergodic. Consider the nilsystem (X, a) . The subgroup of G generated by \mathcal{G} and a acts transitively on X , thus we may and will assume that G is generated by \mathcal{G} and a . Under this assumption, the groups G_d are connected for $d \geq 2$.

2.2. Given a set S and $r \in \mathbb{N}$, we will denote by Δ_{S^r} the diagonal $\left\{ \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix} : x \in S \right\}$ of S^r . Given a mapping $\tau: B \rightarrow C$ and $r \in \mathbb{N}$, we will denote by $\tau^{\times r}$ the mapping $B^r \rightarrow C^r$ defined by $\tau^{\times r}(b_1, \dots, b_r) = (\tau(b_1), \dots, \tau(b_r))$.

2.3. Let $P = \{p_1, \dots, p_r\}$ be a system of (distinct) polynomials $\mathbb{Z}^m \rightarrow \mathbb{Z}$ with zero constant term. We are interested in the closure $Y = \overline{\text{Orb}}_P(\Delta_{X^r})$ of the orbit of the diagonal of X^r under the action (0.1). In the case where p_i are linear polynomials Y is a subnilmanifold of X^r , and has the form $Y = \pi(H)$ where H is a rational subgroup of G^r containing Δ_{G^r} . If (some of) p_i are non-linear, Y is a fu-subnilmanifold of X^r . Let Y^c be the connected component of Y that contains Δ_{X^r} ; then Y^c has the form $Y^c = \pi(H)$ where H is a rational subgroup of G^r containing Δ_{G^r} and Y is a finite union of subnilmanifolds of the form Hx , $x \in X$.

2.4. Let L be a normal rational subgroup of G , let $D = \pi(L)$, and let $X' = L \backslash X = G/(L\Gamma)$. Then the nilmanifold X' is a *factor* of X , and the fibers of the natural projection $\eta: X \rightarrow X'$ are subnilmanifolds of X of the form bD , $b \in G$. We will call D *the fiber* of η .

2.5. The factor X' is *characteristic* for P if for any $f_1, \dots, f_r \in L^\infty(X)$ one has

$$\begin{aligned} \text{C-lim}_n \int_X T^{p_1(n)} f_1 \cdots T^{p_r(n)} f_r d\mu_X \\ = \text{C-lim}_n \int_{X'} T^{p_1(n)} E(f_1|X') \cdots T^{p_r(n)} E(f_r|X') d\mu_{X'}. \end{aligned}$$

Let $Y' = \eta^{\times r}(Y) = \overline{\text{Orb}}_P(\Delta_{(X')^r}) \subseteq (X')^r$. By formula (0.2), X' is characteristic for P iff

$$\int_Y f_1 \otimes \cdots \otimes f_r d\mu_Y = \int_{Y'} E(f_1|X') \otimes \cdots \otimes E(f_r|X') d\mu_{Y'},$$

for any $f_1, \dots, f_r \in L^\infty(X)$, and thus iff $\int_Y \tilde{f} d\mu_Y = \int_{Y'} E(\tilde{f}|(X')^r) d\mu_{Y'}$ for any $\tilde{f} \in L^1(X^r)$. This means that $Y = \pi^{-1}(Y')$, that is, Y consists of entire fibers $\tilde{b}D^r$, $\tilde{b} \in G^r$, of the projection $\eta^{\times r}: X^r \rightarrow (X')^r$. If L is connected, this is so iff $H \supseteq L^r$.

2.6. By X_d , $d = 0, 1, \dots$, we will denote the d^{th} natural factor $G_{d+1} \backslash X = G/(\Gamma G_{d+1})$ of X . If G is s -step nilpotent, one has the tower $X = X_s \rightarrow X_{s-1} \rightarrow \dots \rightarrow X_1 \rightarrow X_0 = \{\cdot\}$ of natural factors of X . The fiber of the projection $\pi_d: X \rightarrow X_d$ is the subnilmanifold $D_d = \pi(G_{d+1}) = G_{d+1}/(\Gamma \cap G_{d+1})$, $d = 0, 1, \dots, s-1$. Since, by our assumption, the groups G_d are connected for $d \geq 2$, the fibers D_d are connected for $d \geq 1$. The fiber of $\pi_0: X \rightarrow X_0$ is $D_0 = X$, which is also connected.

The fiber of the projection $X_d \rightarrow X_{d-1}$ is $F_d = G_d/((\Gamma \cap G_d)G_{d+1})$, $d = 1, \dots, s$. For each d , F_d is a factor of G_{d-1} , and is therefore connected. Since G_d/G_{d+1} is a commutative group, F_d is a torus.

2.7. We will call the minimal integer d for which X_d is characteristic for P , that is, the minimal d for which $G_{d+1}^r \subseteq H$, the *complexity of P with respect to (X, a)* .

If d is the minimal integer for which P has complexity $\leq d$ with respect to all connected nilsystems, we will say that P has *complexity d* . (It is proved in [HK2] and [L5] that the complexity of P is always finite.)

2.8. Here are some evident properties of the complexity (with respect to a nilsystem (X, a)) as a function of a polynomial system P :

- (i) If $P' \subseteq P$, the complexity of $P' \leq$ the complexity of P .
- (ii) If a polynomial systems P' is such that $\overline{\text{Orb}}_{P'}(\Delta_{X^r}) \subseteq \overline{\text{Orb}}_P(\Delta_{X^r})$, then the complexity of $P' \geq$ the complexity of P .
- (iii) The complexity of the system $P = \{p_1, \dots, p_r\}$ is equal to the complexity of the system $\{0, p_2 - p_1, \dots, p_r - p_1\}$.

2.9. The notions of a characteristic factor for and of the complexity of a system of polynomials $P = \{p_1, \dots, p_r\}$ can also be introduced with respect to the “non-integrated” product $T^{p_1(n)} f_1 \cdots T^{p_r(n)} f_r$. Let us say that a factor X' of X is *p -characteristic* (point characteristic) for P if for any $f_1, \dots, f_r \in L^\infty(X)$ one has

$$\text{C-lim}_n T^{p_1(n)} f_1 \cdots T^{p_r(n)} f_r = \text{C-lim}_n T^{p_1(n)} E(f_1|X') \cdots T^{p_r(n)} E(f_r|X')$$

in $L^1(X)$, and call the minimal d such that X_d is p -characteristic for P the p -complexity of P with respect to (X, a) . We are going to show that there is no need in studying the “ p ”-versions of the notions of the characteristic factor and of the complexity, since they are closely related to the ones already introduced.

2.10. For a system of polynomials P , let $Z \subseteq X^r$ be the generic orbit of points of Δ_{X^r} under the action (0.1). Let Z^c be the identity component of Z and let K be the connected rational subgroup of G^r such that $\pi(K) = Z^c$. By subsection 1.15, $H = \Delta_{G^r} K$ and $Y^c = \Delta_{G^r} Z^c$.

Thus, knowing the generic orbit of points of Δ_{X^r} , one can easily find $\overline{\text{Orb}}(\Delta_{X^r})$. On the other hand, the generic orbit Z of points of Δ_{X^r} can be found by determining the orbit $\overline{\text{Orb}}_{\widehat{P}}(\Delta_{X^{r+1}})$ where $\widehat{P} = \{0\} \cup P$. Indeed, for every point $\bar{x} = \begin{pmatrix} x \\ x \\ \vdots \\ x \end{pmatrix} = \begin{pmatrix} x \\ \bar{x} \end{pmatrix} \in \Delta_{X^{r+1}}$ one has $\overline{\text{Orb}}_{\widehat{P}}(\bar{x}) = \left(\frac{x}{\text{Orb}_P(\bar{x})} \right)$; thus for almost all $x \in X$, $\overline{\text{Orb}}_{\widehat{P}}(\bar{x}) = \begin{pmatrix} b_x \\ \bar{b}_x Z \end{pmatrix}$ and for all $x \in X$, $\overline{\text{Orb}}_{\widehat{P}}(\bar{x}) \subseteq \begin{pmatrix} b_x \\ \bar{b}_x Z \end{pmatrix}$, where $b_x \in G$ with $\pi(b_x) = x$, and $\bar{b}_x = \begin{pmatrix} b_x \\ \vdots \\ b_x \end{pmatrix} \in \Delta_{G^r}$. It follows that $\overline{\text{Orb}}_{\widehat{P}}(\Delta_{X^{r+1}}) = \bigcup_{x \in X} \begin{pmatrix} x \\ \bar{b}_x Z \end{pmatrix}$, and in particular, $\begin{pmatrix} 1_X \\ Z \end{pmatrix} = \overline{\text{Orb}}_{\widehat{P}}(\Delta_{X^{r+1}}) \cap \begin{pmatrix} 1_X \\ X^r \end{pmatrix}$.

2.11. Now let $X' = L \backslash X$ where L is a connected normal rational subgroup of G ; the fiber of the projection $X \rightarrow X'$ is then $D = \pi(L)$. Then X' is p -characteristic for P iff for almost every point $\bar{x} \in \Delta_{X^r}$, $\overline{\text{Orb}}_P(\bar{x})$ consists of entire fibers $\tilde{b}D^r$, $\tilde{b} \in G^r$, of the projection $X^r \rightarrow (X')^r$. It follows that X' is p -characteristic for P iff $K \supseteq L^r$.

It now follows that X' is p -characteristic for P iff it is characteristic for \widehat{P} . Indeed, let \widehat{H} be the connected rational subgroup of G such that $\pi(\widehat{H})$ is the identity component of $\overline{\text{Orb}}_{\widehat{P}}(\Delta_{X^{r+1}})$; then $\widehat{H} = \bigcup_{b \in G} \begin{pmatrix} b \\ \bar{b}K \end{pmatrix}$, where, for $b \in G$, $\bar{b} = \begin{pmatrix} b \\ \vdots \\ b \end{pmatrix} \in \Delta_{G^r}$. If X' is p -characteristic for P then $K \supseteq L^r$, so $\widehat{H} \supseteq \bigcup_{b \in G} \begin{pmatrix} b \\ \bar{b}L^r \end{pmatrix} \supseteq \bigcup_{b \in L} \begin{pmatrix} b \\ \bar{b}L^r \end{pmatrix} = L^{r+1}$, and X' is characteristic for \widehat{P} . If X' is characteristic for \widehat{P} , then $\widehat{H} \supseteq L^{r+1}$, so $\begin{pmatrix} 1_G \\ L^r \end{pmatrix} \subseteq \begin{pmatrix} 1_G \\ K \end{pmatrix}$, so $L^r \subseteq K$, and hence, X' is p -characteristic for P .

As a corollary, we obtain that the p -complexity of a polynomial system P equals the complexity of the system $\widehat{P} = \{0\} \cup P$.

2.12. Let us also clarify why one may confine oneself to the case of a connected nilmanifold. Assume that the nilmanifold X is disconnected; since X is compact, it consists of finitely many connected components, $X = X_1 \cup \dots \cup X_k$. Since the translation by a is ergodic, it cyclically permutes the components X_j ; thus $a^k X_j = X_j$ for all j . Consider the nilsystems $(X_j, a^k|_{X_j})$, $j = 1, \dots, k$, and the polynomial systems $P_i = \left\{ \frac{1}{k}(p_1(kn+i) - p_1(i)), \dots, \frac{1}{k}(p_r(kn+i) - p_r(i)) \right\}$, $i \in \{0, \dots, k-1\}^m$, of polynomials

in n with zero constant term. Then

$$\overline{\text{Orb}}_P(\Delta_{X^r}) = \bigcup_{j=1}^k \bigcup_{i=0}^{k-1} \begin{pmatrix} a^{p_1(i)} \\ \vdots \\ a^{p_r(i)} \end{pmatrix} \overline{\text{Orb}}_{P_i}(\Delta_{X_j^r}),$$

and the study of the orbit of the diagonal in the disconnected nilsystem (X, a) is reduced to that in the connected nilsystems $(X_j, a^k|_{X_j})$.

In particular, if the fiber D of a projection $X \rightarrow X'$ is connected, it is contained in $\overline{\text{Orb}}_P(\Delta_{X^r})$ if it is contained in $\overline{\text{Orb}}_{P_0}(\Delta_{X_1^r})$ (assuming $X_1 \ni 1_X$). Since even in the case of disconnected X the fibers D_d are connected for $d \geq 1$, we see that the complexity of P with respect to (X, a) , if positive, does not exceed the complexity of the system $P_0 = \{\frac{1}{k}p_1(kn), \dots, \frac{1}{k}p_r(kn)\}$ with respect to $(X_1, a^k|_{X_1})$, and thus does not exceed the complexity of P_0 . One can show that the complexity of P_0 with respect to any connected nilsystem is equal to the complexity of P with respect to this nilsystem. Hence, the complexity of P with respect to all, including disconnected, nilsystems, if positive, equals the complexity of P with respect to connected nilsystems only.

3. Spans

In this section we introduce some linear algebra notation.

3.1. Given a set $S \subseteq \mathbb{Z}^r$ we will denote by $\text{span}_{\mathbb{Z}} S$ the subgroup of \mathbb{Z}^r generated by S . For an integer matrix $B = \begin{pmatrix} b_{1,1} & \dots & b_{1,l} \\ \vdots & & \vdots \\ b_{r,1} & \dots & b_{r,l} \end{pmatrix}$ we will denote by $\text{span}_{\mathbb{Z}} B$ the subgroup of \mathbb{Z}^r generated by the columns of B ,

$$\text{span}_{\mathbb{Z}} B = \text{span}_{\mathbb{Z}} \left\{ \begin{pmatrix} b_{1,1} \\ \vdots \\ b_{r,1} \end{pmatrix}, \dots, \begin{pmatrix} b_{1,l} \\ \vdots \\ b_{r,l} \end{pmatrix} \right\}.$$

If B is a polynomial matrix, $B = \begin{pmatrix} p_{1,1} & \dots & p_{1,l} \\ \vdots & & \vdots \\ p_{r,1} & \dots & p_{r,l} \end{pmatrix}$ where $p_{i,j}$ are integer-valued polynomials on \mathbb{Z}^m with zero constant term, we will denote by $\text{Span}_{\mathbb{Z}} B$ the subgroup of \mathbb{Z}^r generated by the values of the columns of B ,

$$\text{Span}_{\mathbb{Z}} B = \sum_{n \in \mathbb{Z}^m} \text{span}_{\mathbb{Z}} \begin{pmatrix} p_{1,1}(n) & \dots & p_{1,l}(n) \\ \vdots & & \vdots \\ p_{r,1}(n) & \dots & p_{r,l}(n) \end{pmatrix}.$$

3.2. Similarly, for $S \subseteq \mathbb{R}^r$ we will denote by $\text{span}_{\mathbb{R}} S$ the subspace of \mathbb{R}^r spanned by S , for a real matrix $B = \begin{pmatrix} b_{1,1} & \dots & b_{1,m} \\ \vdots & & \vdots \\ b_{r,1} & \dots & b_{r,m} \end{pmatrix}$ we define

$$\text{span}_{\mathbb{R}} B = \text{span}_{\mathbb{R}} \left\{ \begin{pmatrix} b_{1,1} \\ \vdots \\ b_{r,1} \end{pmatrix}, \dots, \begin{pmatrix} b_{1,m} \\ \vdots \\ b_{r,m} \end{pmatrix} \right\},$$

and for a matrix $B = \begin{pmatrix} p_{1,1} & \dots & p_{1,l} \\ \vdots & & \vdots \\ p_{r,1} & \dots & p_{r,l} \end{pmatrix}$ with $p_{i,j}$ being polynomials $\mathbb{Z}^m \rightarrow \mathbb{R}$ we put

$$\text{Span}_{\mathbb{R}} B = \sum_{n \in \mathbb{Z}^m} \text{span}_{\mathbb{R}} \begin{pmatrix} p_{1,1}(n) & \dots & p_{1,l}(n) \\ \vdots & & \vdots \\ p_{r,1}(n) & \dots & p_{r,l}(n) \end{pmatrix}.$$

Note that if $p_{i,j} = \sum_{t=1}^k b_{i,j,t} q_t$, $i = 1, \dots, r$, $j = 1, \dots, l$, where q_1, \dots, q_t are distinct monomials with trivial coefficient (or just linearly independent polynomials), then

$$\text{Span}_{\mathbb{R}} \begin{pmatrix} p_{1,1}(n) & \dots & p_{1,l}(n) \\ \vdots & & \vdots \\ p_{r,1}(n) & \dots & p_{r,l}(n) \end{pmatrix} = \text{span}_{\mathbb{R}} \begin{pmatrix} b_{1,1,1} & \dots & b_{1,m,1} & \dots & b_{1,1,k} & \dots & b_{1,m,k} \\ \vdots & & \vdots & & \vdots & & \vdots \\ b_{r,1,1} & \dots & b_{r,m,1} & \dots & b_{r,1,k} & \dots & b_{r,m,k} \end{pmatrix}.$$

(Since \mathbb{Z} is disconnected, an analogous fact does not hold for $\text{Span}_{\mathbb{Z}}$.)

3.3. More generally, let Z be an abelian group. For a vector $v = \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix} \in \mathbb{Z}^r$, we will denote by $\text{span}_Z v$ the subgroup

$$\{vz : z \in Z\} = \left\{ \begin{pmatrix} b_1 z \\ \vdots \\ b_r z \end{pmatrix} : z \in Z \right\}$$

of \mathbb{Z}^r . For an integer matrix $B = \begin{pmatrix} b_{1,1} & \dots & b_{1,m} \\ \vdots & & \vdots \\ b_{r,1} & \dots & b_{r,m} \end{pmatrix}$ we define

$$\text{span}_Z B = \text{span}_Z \left\{ \begin{pmatrix} b_{1,1} \\ \vdots \\ b_{r,1} \end{pmatrix}, \dots, \begin{pmatrix} b_{1,m} \\ \vdots \\ b_{r,m} \end{pmatrix} \right\}.$$

If B is a polynomial matrix, $B = \begin{pmatrix} p_{1,1} & \dots & p_{1,l} \\ \vdots & & \vdots \\ p_{r,1} & \dots & p_{r,l} \end{pmatrix}$ where $p_{i,j}$ are integer-valued polynomials on \mathbb{Z}^m with zero constant term, we put

$$\text{Span}_Z B = \sum_{n \in \mathbb{Z}^m} \text{span}_Z \begin{pmatrix} p_{1,1}(n) & \dots & p_{1,l}(n) \\ \vdots & & \vdots \\ p_{r,1}(n) & \dots & p_{r,l}(n) \end{pmatrix}.$$

3.4. Now let A_1, \dots, A_r be endomorphisms of Z and let $v = \begin{pmatrix} A_1 \\ \vdots \\ A_r \end{pmatrix}$; then for $z \in Z$ the

vector $vz = \begin{pmatrix} A_1 z \\ \vdots \\ A_r z \end{pmatrix} \in \mathbb{Z}^r$ makes sense, and we again may define $\text{span}_Z v = \{vz : z \in Z\}$.

For $B = \begin{pmatrix} A_{1,1} & \dots & A_{1,m} \\ \vdots & & \vdots \\ A_{r,1} & \dots & A_{r,m} \end{pmatrix}$ where $A_{i,j} \in \text{End}(Z)$ we define

$$\text{span}_Z B = \text{span}_Z \left\{ \begin{pmatrix} A_{1,1} \\ \vdots \\ A_{r,1} \end{pmatrix}, \dots, \begin{pmatrix} A_{1,m} \\ \vdots \\ A_{r,m} \end{pmatrix} \right\},$$

and if $B = \begin{pmatrix} Q_{1,1} & \dots & Q_{1,l} \\ \vdots & & \vdots \\ Q_{r,1} & \dots & Q_{r,l} \end{pmatrix}$ where $Q_{i,j}$ are mappings $\mathbb{Z}^m \rightarrow \text{End}(Z)$, we put

$$\text{Span}_Z B = \sum_{n \in \mathbb{Z}^m} \text{span}_Z \begin{pmatrix} Q_{1,1}(n) & \dots & Q_{1,l}(n) \\ \vdots & & \vdots \\ Q_{r,1}(n) & \dots & Q_{r,l}(n) \end{pmatrix}.$$

4. A group of polynomial mappings to a nilpotent group

In this section we describe the group $\text{Pol}(K, G)$ of polynomial mappings from $K = \mathbb{Z}^l$ or \mathbb{R}^l to the nilpotent Lie group G ; it will be used in the next section to construct the group $G^{[V]}$ whose projection to X^r is $\overline{\text{Orb}}_P(\Delta_{X^r})$.

4.1. Let K be one of the groups \mathbb{Z}^l or \mathbb{R}^l , $l \in \mathbb{N}$. We will denote by $\text{Pol}(K)$ the set of polynomials on K , and for $d \in \mathbb{N}$ we will denote by $\text{Pol}_d(K)$ the set of polynomials on K of degree $\leq d$.

Given $b \in G$ and $q \in \text{Pol}(K)$, we may consider the mapping $b^q: K \rightarrow G$, $u \mapsto b^{q(u)}$ for $u \in K$. To avoid complicated notation, let us make the following agreement: when we write b^q with $b \notin G$, we assume that $K = \mathbb{Z}^l$ and $q(K) \subseteq \mathbb{Z}$; then $q(u) \in \mathbb{Z}$ and $b^{q(u)}$ makes sense for all $u \in K$.

A *polynomial mapping* $\varphi: K \rightarrow G$ is a mapping of the form $\varphi(u) = \prod_{i=1}^k b_i^{q_i(u)}$, $u \in K$, where $b_i \in G$ and $q_i \in \text{Pol}(K)$, $i = 1, \dots, k$. With respect to the element-wise multiplication, polynomial mappings from K to G form a group $\text{Pol}(K, G)$.

4.2. Given subsets S_1, \dots, S_l of a group R , we will denote by $\langle S_1, \dots, S_l \rangle$ the subgroup of R generated by S_1, \dots, S_l . We have $\text{Pol}(K, G) = \langle b^q : b \in G, q \in \text{Pol}(K) \rangle$.

4.3. We define a subgroup $\text{Pol}^*(K, G)$ of the group $\text{Pol}(K, G)$ in the following way:

$$\text{Pol}^*(K, G) = \langle b^q : b \in G_d, q \in \text{Pol}_d(K) \text{ for some } d \rangle.$$

Let $\{e_{d,j} \in \Gamma : d = 1, \dots, s, j = 1, \dots, k_d\}$ be a Malcev basis in G compatible with Γ and with the lower central series of G , such that for each $d = 1, \dots, s$, the elements $e_{d,j} \bmod G_{d+1}$, $j = 1, \dots, k_d$, form a basis in G_d/G_{d+1} . The following proposition clearly follows from subsection 1.8:

Proposition. *Every polynomial mapping $\varphi \in \text{Pol}^*(K, G)$ is uniquely representable in the form $\varphi = \prod_{d=1}^s \prod_{j=1}^{k_d} e_{d,j}^{q_{d,j}}$ with $q_{d,j} \in \text{Pol}_d(K)$ for all d and j . In particular, if $\varphi \in \text{Pol}^*(K, G)$ satisfies $\text{Im}(\varphi) \subseteq G_{d_0}$ for some $d_0 \in \mathbb{N}$, then φ has the form $\varphi = \prod_{d=d_0}^s \prod_{j=1}^{k_d} e_{d,j}^{q_{d,j}}$ with $q_{d,j} \in \text{Pol}_d(K)$ for all d and j .*

4.4. Let $u_1, \dots, u_r \in K$; consider the subgroup

$$\text{Pol}^*(K, G) \begin{pmatrix} u_1 \\ \vdots \\ u_r \end{pmatrix} = \left\{ \begin{pmatrix} \varphi(u_1) \\ \vdots \\ \varphi(u_r) \end{pmatrix} : \varphi \in \text{Pol}^*(K, G) \right\}$$

of G^r .

Corollary. Any $\tilde{b} \in \text{Pol}^*(K, G) \binom{u_1}{u_r}$ is uniquely representable in the form

$$\tilde{b} = \begin{pmatrix} \prod_{d=1}^s \prod_{j=1}^{k_d} e_{d,j}^{q_{d,j}(u_1)} \\ \vdots \\ \prod_{d=1}^s \prod_{j=1}^{k_d} e_{d,j}^{q_{d,j}(u_r)} \end{pmatrix} \quad (4.1)$$

with $q_{d,j} \in \text{Pol}_d(K)$ for all d and j . In particular, if $\tilde{b} \in \text{Pol}^*(K, G) \binom{u_1}{u_r} \cap G_{d_0}^r$ for some $d_0 \in \mathbb{N}$, then

$$\tilde{b} = \begin{pmatrix} \prod_{d=d_0}^s \prod_{j=1}^{k_d} e_{d,j}^{q_{d,j}(u_1)} \\ \vdots \\ \prod_{d=d_0}^s \prod_{j=1}^{k_d} e_{d,j}^{q_{d,j}(u_r)} \end{pmatrix} \quad (4.2)$$

with $q_{d,j} \in \text{Pol}_d(K)$ for all d and j .

4.5. A polynomial sequence in G is a polynomial mapping $\mathbb{Z} \rightarrow G$. For a mapping $g: \mathbb{Z} \rightarrow G$ we define the derivative $Dg: \mathbb{Z} \rightarrow G$ of φ by $Dg(n) = g(n)^{-1}g(n+1)$. It is easy to see that φ is polynomial iff $D^k g = \text{const}$ for some $k \in \mathbb{N}$ (see [L1] or [L2]); we call the minimal k with this property the *degree* of g . If g is an element of $\text{Pol}^*(\mathbb{Z}, G)$ and the nilpotency class of G is $\leq s$, then the degree of g is $\leq s$.

4.6. The following fact can be easily proved by induction on the degree of a polynomial sequence:

Lemma. Any polynomial sequence g of degree $\leq k$ is uniquely determined by its initial values $g(0), g(1), \dots, g(k)$.

4.7. Any polynomial sequence $g \in \text{Pol}^*(\mathbb{Z}, G)$ is uniquely representable in the form $g(n) = \prod_{k=0}^s b_k \binom{n}{k}$, $n \in \mathbb{Z}$, with $b_k \in G_k$ for all k (where $\binom{n}{k} = \frac{1}{k!}n(n-1)\dots(n-k+1)$). The elements b_k can be found inductively in the following way. We put $b_0 = g(0)$. Assume that elements $b_0, b_1, \dots, b_{d-1} \in G$ have already been found so that $g(n) = \prod_{k=0}^{d-1} b_k \binom{n}{k} \text{ mod } G_d$, $n \in \mathbb{Z}$. We then define

$$b_d = \left(\prod_{k=0}^{d-1} b_k \binom{n}{k} \right)^{-1} g(d),$$

and get $b_d \in G_d$. The sequence $\prod_{k=0}^d b_k \binom{n}{k} \text{ mod } G_{d+1}$ has degree $\leq d$ and coincides with $g(n) \text{ mod } G_{d+1}$ for $n = 0, \dots, d$; by Lemma 4.6, $g(n) = \prod_{k=0}^d b_k \binom{n}{k} \text{ mod } G_{d+1}$ for all $n \in \mathbb{Z}$. After determining all of b_0, b_1, \dots, b_s , we obtain $g(n) = \prod_{k=0}^s b_k \binom{n}{k}$, $n \in \mathbb{Z}$.

Remark. An analogous fact holds for polynomial mappings $\varphi \in \text{Pol}^*(K, G)$ with $l > 1$ and can be derived similarly, but we do not need it in this paper.

5. A subgroup of G^r associated with a subgroup of \mathbb{Z}^r

In this section we construct and study the subgroup $G^{[V]} \subseteq G^r$, for V being a subgroup of \mathbb{Z}^r or a subspace of \mathbb{R}^r ; this group can be seen as an analogue of $\text{span}_G V$ for non-abelian G .

5.1. We introduce the coordinate-wise multiplication on \mathbb{R}^r : for vectors $v_1 = \begin{pmatrix} c_{1,1} \\ \vdots \\ c_{1,r} \end{pmatrix}$ and $v_2 = \begin{pmatrix} c_{2,1} \\ \vdots \\ c_{2,r} \end{pmatrix}$ in \mathbb{R}^r let $v_1 v_2 = \begin{pmatrix} c_{1,1} c_{2,1} \\ \vdots \\ c_{1,r} c_{2,r} \end{pmatrix}$. For vectors $v_1 = \begin{pmatrix} c_{1,1} \\ \vdots \\ c_{1,r} \end{pmatrix}, \dots, v_l = \begin{pmatrix} c_{l,1} \\ \vdots \\ c_{l,r} \end{pmatrix}$ in \mathbb{R}^r and a polynomial $q \in \text{Pol}(\mathbb{R}^l)$ we then have $q(v_1, \dots, v_l) = \begin{pmatrix} q(c_{1,1}, \dots, c_{l,1}) \\ \vdots \\ q(c_{1,r}, \dots, c_{l,r}) \end{pmatrix} \in \mathbb{R}^r$.

5.2. Let \hat{V} be a subgroup of \mathbb{Z}^r . For $d \in \mathbb{N}$ we define

$$\mathcal{P}_d \hat{V} = \{q(v_1, \dots, v_l) : l \in \mathbb{N}, q \in \text{Pol}_d(\mathbb{Z}^l), q(\mathbb{Z}^l) \subseteq \mathbb{Z}, \text{ and } v_1, \dots, v_l \in \hat{V}\}.$$

$\mathcal{P}_d \hat{V}$ is a subgroup of \mathbb{Z}^r . If \hat{V} is generated by vectors v_1, \dots, v_l , then $\mathcal{P}_d \hat{V}$ is generated by the vectors $\prod_{i=1}^l \binom{v_i}{d_i}$ with $\sum_{i=1}^l d_i \leq d$ (where $\binom{v}{k} = \frac{1}{k!} v(v-1) \dots (v-k+1)$).

Note that if $\hat{V} \supseteq \Delta_{\mathbb{Z}^r}$, then $\mathcal{P}_1 \hat{V} = \hat{V}$; we will only need to deal with subgroups \hat{V} having this property.

We also define

$$\hat{V}^{*d} = \text{span}_{\mathbb{Z}} \{v_1 v_2 \dots v_l : 0 \leq l \leq d, v_1, \dots, v_l \in \hat{V}\}$$

(where we assume that the empty product is equal to $\binom{1}{1}$). If \hat{V} is generated by vectors v_1, \dots, v_l , then \hat{V}^{*d} is generated by the vectors $\prod_{i=1}^l v_i^{d_i}$ with $\sum_{i=1}^l d_i \leq d$. Therefore, \hat{V}^{*d} is a subgroup of finite index in $\mathcal{P}_d \hat{V}$.

5.3. Now let V be a subspace of \mathbb{R}^r . For $d \in \mathbb{N}$ let

$$\mathcal{P}_d V = \{q(v_1, \dots, v_l) : l \in \mathbb{N}, q \in \text{Pol}_d(\mathbb{R}^l) \text{ and } v_1, \dots, v_l \in V\}$$

and

$$V^{*d} = \text{span}_{\mathbb{R}} \{v_1 v_2 \dots v_l : 0 \leq l \leq d, v_1, \dots, v_l \in V\}.$$

Then $\mathcal{P}_d V = V^{*d}$. If V is spanned by vectors v_1, \dots, v_l , then V^{*d} is spanned by the vectors $\prod_{i=1}^l v_i^{d_i}$ with $\sum_{i=1}^l d_i \leq d$.

If \hat{V} is a subgroup of \mathbb{Z}^r and V is the subspace of \mathbb{R}^r spanned by \hat{V} , then, for any $d \in \mathbb{N}$, $\text{span}_{\mathbb{R}} \mathcal{P}_d \hat{V} = \mathcal{P}_d V = V^{*d}$.

5.4. Let b be an element of a group M and let $v = \begin{pmatrix} c_1 \\ \vdots \\ c_r \end{pmatrix} \in \mathbb{Z}^r$. We define $b^v = \begin{pmatrix} b^{c_1} \\ \vdots \\ b^{c_r} \end{pmatrix} \in M^r$.

If M is a Lie group and $b \in M$ is such that b^t is defined for all $t \in \mathbb{R}$, then $b^v = \begin{pmatrix} b^{c_1} \\ \vdots \\ b^{c_r} \end{pmatrix} \in M^r$

is defined for $v = \begin{pmatrix} c_1 \\ \vdots \\ c_r \end{pmatrix} \in \mathbb{R}^r$.

5.5. Let M be a commutative group and let \hat{V} be a subgroup of \mathbb{Z}^r . We define a subgroup $M^{[\hat{V}]}$ of M^r in the following way:

$$M^{[\hat{V}]} = \langle b^v : b \in M, v \in \hat{V} \rangle.$$

($M^{[\hat{V}]}$ is the same as $\text{span}_M(\hat{V})$ in additive notation.)

If M is a connected simply-connected commutative Lie group (that is, a vector space), then the subgroup

$$M^{[V]} = \langle b^v : b \in M, v \in V \rangle$$

of M^r is also defined for any subspace V of \mathbb{R}^r . In this case, if \hat{V} is a subgroup of \mathbb{Z}^r and $V = \text{span}_{\mathbb{R}} \hat{V}$, we have $M^{[V]} = M^{[\hat{V}]}$. The dimension of $M^{[V]}$ is $\dim M \cdot \dim V$, and $M^{[V]} = M^r$ iff $V = \mathbb{R}^r$.

We would prefer to always deal with the \mathbb{R} -space V instead of the \mathbb{Z} -module \hat{V} , but we have to use \hat{V} when our group M is disconnected.

5.6. Now let G be a simply-connected nilpotent Lie group with a discrete uniform subgroup Γ , and let \hat{V} be a subgroup of \mathbb{Z}^r . We define

$$\begin{aligned} G^{[\hat{V}]} &= \langle b^w : w \in \mathcal{P}_d \hat{V} \text{ and } b \in G_d \text{ for some } d \rangle \\ &= \langle b^{q(v_1, \dots, v_l)} : l \in \mathbb{N}, q \in \text{Pol}_d(\mathbb{Z}^l), b \in G_d \text{ for some } d, \text{ and } v_j \in \hat{V}, j = 1, \dots, l \rangle. \end{aligned}$$

5.7. Proposition. $G^{[\hat{V}]}$ is a closed rational subgroup of G^r ,

$$G^{[\hat{V}]} = \left\{ \prod_{d=1}^s \prod_{j=1}^{l_d} b_{d,j}^{w_{d,j}} : b_{d,j} \in G_d \text{ and } w_{d,j} \in \mathcal{P}_d \hat{V} \text{ for all } d \text{ and } j \right\}, \quad (5.1)$$

and for any $d_0 \in \mathbb{N}$,

$$G^{[\hat{V}]} \cap G_{d_0}^r = \left\{ \prod_{d=d_0}^s \prod_{j=1}^{l_d} b_{d,j}^{w_{d,j}} : b_{d,j} \in G_d \text{ and } w_{d,j} \in \mathcal{P}_d \hat{V} \text{ for all } d \text{ and } j \right\}. \quad (5.2)$$

Proof. Let \hat{V} be generated by vectors $v_1 = \begin{pmatrix} c_{1,1} \\ \vdots \\ c_{1,r} \end{pmatrix}, \dots, v_l = \begin{pmatrix} c_{l,1} \\ \vdots \\ c_{l,r} \end{pmatrix}$. Put $u_1 = (c_{1,1}, \dots, c_{l,1}), \dots, u_r = (c_{1,r}, \dots, c_{l,r}) \in \mathbb{Z}^l$. Then

$$\begin{aligned} G^{[\hat{V}]} &= \langle b^{q(v_1, \dots, v_l)} : q \in \text{Pol}_d(\mathbb{Z}^l) \text{ and } b \in G_d \text{ for some } d \rangle \\ &= \left\langle \begin{pmatrix} b^{q(u_1)} \\ \vdots \\ b^{q(u_r)} \end{pmatrix} : q \in \text{Pol}_d(\mathbb{Z}^l) \text{ and } b \in G_d \text{ for some } d \right\rangle = \text{Pol}^*(\mathbb{Z}^l, G) \begin{pmatrix} u_1 \\ \vdots \\ u_r \end{pmatrix} \end{aligned}$$

in the notation of subsection 4.4. Hence, $G^{[\hat{V}]}$ is a subgroup of G^r , and (5.1), (5.2) follow from (4.1), (4.2) respectively; it also follows from (4.1) that $G^{[\hat{V}]}$ is closed. Clearly, $\Gamma^{[\hat{V}]}$ is uniform in $G^{[\hat{V}]}$, and thus $G^{[\hat{V}]}$ is rational with respect to Γ^r . ■

5.8. Corollary. *Let, for each d , the group $\mathcal{P}_d \hat{V}$ is generated modulo $\mathcal{P}_{d-1} \hat{V}$ by elements $w_{d,1}, \dots, w_{d,l_d} \in \mathcal{P}_d \hat{V}$. (We assume here that $\mathcal{P}_0 \hat{V} = \{0\}$.) Then*

$$G^{[\hat{V}]} = \left\{ \prod_{d=1}^s \prod_{j=1}^{l_d} b_{d,j}^{w_{d,j}} : b_{d,j} \in G_d \text{ for all } d \text{ and } j \right\}.$$

Proof. Let $\tilde{b} \in G^{[\hat{V}]}$. Since the vectors $w_{1,1}, \dots, w_{1,l_1}$ generate $\mathcal{P}_1 \hat{V}$, the element $\tilde{b} \bmod G_2^r$ can be written in the form

$$\tilde{b} = \prod_{j=1}^{l_1} b_{1,j}^{w_{1,j}} \bmod G_2^r,$$

with $b_{1,1}, \dots, b_{1,l_1} \in G$.

Consider the element $\tilde{b}_1 = \prod_{j=1}^{l_1} b_{1,j}^{w_{1,j}}$. We have $\tilde{b}\tilde{b}_1^{-1} \in G^{[\hat{V}]} \cap G_2^r$, thus, by Proposition 5.7 and since the vectors $w_{1,1}, \dots, w_{1,l_1}, w_{2,1}, \dots, w_{2,l_2}$ generate $\mathcal{P}_2 \hat{V}$, we can write

$$\tilde{b}\tilde{b}_1^{-1} = \prod_{j=1}^{l_1} d_{1,j}^{w_{1,j}} \prod_{j=1}^{l_2} b_{2,j}^{w_{2,j}} \bmod G_3^r$$

with $d_{1,1}, \dots, d_{1,l_1}, b_{2,1}, \dots, b_{2,l_2} \in G_2$. Put

$$\tilde{b}_2 = \prod_{j=1}^{l_1} (b_{1,j} d_{1,j})^{w_{1,j}} \prod_{j=1}^{l_2} b_{2,j}^{w_{2,j}},$$

then $\tilde{b} = \tilde{b}_2 \bmod G_3$.

And so on, by induction, until we get $\tilde{b}_s \in G^{[\hat{V}]}$ representable in the desired form and such that $\tilde{b} = \tilde{b}_s \bmod G_{s+1}$. Since $G_{s+1} = \{\mathbf{1}_G\}$, $\tilde{b} = \tilde{b}_s$. ■

5.9. If G is connected, the closed rational subgroup

$$\begin{aligned} G^{[V]} &= \left\langle b^w : w \in V^{*d} \text{ and } b \in G_d \text{ for some } d \right\rangle \\ &= \left\langle b^{q(v_1, \dots, v_l)} : l \in \mathbb{N}, q \in \text{Pol}_d(\mathbb{Z}^l), b \in G_d \text{ for some } d, \text{ and } v_j \in V, j = 1, \dots, l \right\rangle \end{aligned}$$

of G^r is defined for any subspace V of \mathbb{R}^r . If $V = \text{span}_{\mathbb{R}} \hat{V}$ for a subgroup \hat{V} of \mathbb{Z}^r , then $G^{[V]} = G^{[\hat{V}]}$.

5.10. Let \hat{V} be a subgroup of \mathbb{Z}^r . For $d \in \mathbb{N}$ let $M_d = G_d/G_{d+1}$; M_d is then a commutative Lie group. It follows from (5.2) that for any d ,

$$(G^{[\hat{V}]} \cap G_d^r) \bmod G_{d+1}^r = M_d^{[\mathcal{P}_d \hat{V}]}.$$

If, for some d , M_d is connected (and thus is a vector space), we have

$$(G^{[\hat{V}]} \cap G_d^r) \bmod G_{d+1}^r = M_d^{[\mathcal{P}_d V]} = M_d^{[V^{*d}]}$$

where $V = \text{span}_{\mathbb{R}}(\hat{V})$.

5.11. Let \hat{V} be a subgroup of \mathbb{Z}^r and let $V = \text{span}_{\mathbb{R}}(\hat{V})$. If, for some d , $\mathcal{P}_d(\hat{V}) = \mathbb{Z}^r$ or M_d is connected and $V^{*d} = \mathbb{R}^r$, then $M_k^{[\mathcal{P}_d(\hat{V})]} = M_k^r$ for all $k \geq d$, and so

$$G^{[\hat{V}]} = \left\langle b^w : w \in \mathcal{P}_d \hat{V} \text{ and } b \in G_k \text{ for some } k < d \right\rangle \cdot G_d^r.$$

5.12. The following technical lemma says that if H is a subgroup of G^r , then knowing “the linear part” H/G_2^r of H one may “estimate from below” the other “graduated components” $(H \cap G_d^r) \bmod G_{d+1}^r$, $d = 2, 3, \dots$, of H .

Lemma. *Let \hat{V} be a subgroup of \mathbb{Z}^r with $\hat{V} \supseteq \Delta_{\mathbb{Z}^r}$ and let H be a subgroup of G^r such that $H \bmod G_2^r \supseteq M_1^{[\hat{V}]}$. Then $(H \cap G_d^r) \bmod G_{d+1}^r \supseteq M_d^{[V^{*d}]}$ for all d .*

Proof. Let $d \geq 2$, $k \leq d$ and $v_1, \dots, v_k \in \hat{V}$; put $v_{k+1} = \dots = v_d = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$. Then for any $b = [b_1, b_2, \dots, b_d] \in G_d$ we have

$$b^{v_1 v_2 \dots v_k} = b^{v_1 v_2 \dots v_d} = [b_1^{v_1}, b_2^{v_2}, \dots, b_d^{v_d}] \bmod G_{d+1}^r.$$

Since $b_i^{v_i} \bmod G_2^r \in M_1^{[\hat{V}]} \subseteq H \bmod G_2^r$ for $i = 1, \dots, d$, we obtain $b^{v_1 v_2 \dots v_k} \in H \bmod G_{d+1}^r$. ■

5.13. The elements b^v , $b \in G$, $v \in \hat{V}$, of G^r belong to the group $G^{[\hat{V}]}$. Actually, $G^{[\hat{V}]}$ is generated by these elements; this fact is algebraic and is true for a general nilpotent group G without any topology assumed on it. We do not need this fact here; what we will need is the following weaker proposition:

Proposition. *Assume that the subgroups G_d are connected for $d \geq 2$. Let \hat{V} be a subgroup of \mathbb{Z}^r with $\hat{V} \supseteq \Delta_{\mathbb{Z}^r}$ and let H be a subgroup of G^r such that $H \subseteq G^{[\hat{V}]}$ and $H \bmod G_2^r \supseteq M_1^{[\hat{V}]}$. Then $H = G^{[\hat{V}]}$.*

Proof. For any $d \geq 2$, by Lemma 5.12 and since M_d is connected,

$$(H \cap G_d^r) \bmod G_{d+1}^r \supseteq M_d^{[\hat{V}^{*d}]} = M_d^{[V^{*d}]} = M_d^{[\mathcal{P}_d V]} = M_d^{[\mathcal{P}_d \hat{V}]} = (G^{[\hat{V}]} \cap G_d^r) \bmod G_{d+1}^r,$$

where $V = \text{span}_{\mathbb{R}} \hat{V}$. Since also $H \bmod G_2^r \supseteq M_1^{[\hat{V}]} = G^{[\hat{V}]} \bmod G_2^r$, we obtain $H \supseteq G^{[\hat{V}]}$. ■

5.14. Below we will also need the following lemma:

Lemma. *Let \hat{V} be a subgroup of \mathbb{Z}^r . Then $[G_d^{[\hat{V}]}, \Delta_{G^r}] = G_{d+1}^{[\hat{V}]} \bmod G_{d+2}^r$ for any $d \in \mathbb{N}$.*

Proof. Put $e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \Delta_{\mathbb{Z}^r}$. For any $b_1 \in G_d$, $b_2 \in G$, and $v \in \hat{V}$,

$$[b_1, b_2]^v = [b_1^v, b_2^e] \bmod G_{d+2}^r \in [G_d^{[\hat{V}]}, \Delta_{G^r}] \bmod G_{d+2}^r,$$

and $G_{d+1}^{[\hat{V}]} \bmod G_{d+2}^r$ is generated by these elements. ■

6. The orbit of the diagonal under a system of linear actions

In this section we find $\overline{\text{Orb}}_P(\Delta_{X^r})$ in the case where $P = \{p_1, \dots, p_r\}$ is a system of linear functions.

6.1. We assume that G is a simply-connected (but not necessarily connected) nilpotent Lie group, Γ is a discrete uniform subgroup of G , X is the nilmanifold G/Γ and a is an element of G whose action on X is ergodic. We also assume that X is connected and that G is generated by \mathcal{G} and a ; this implies that the groups G_d with $d \geq 2$ are connected.

We denote by π is the natural projection $G \rightarrow X$, by X_d , $d = 0, 1, 2, \dots$, the factor $G_{d+1} \backslash X$ of X , by $\pi_d: X \rightarrow X_d$ the corresponding projection, and by D_d the fiber $\pi(G_d)$ of π_{d-1} . We also denote by M_d , $d = 1, 2, \dots$, the factor group G_d/G_{d+1} and by F_d the fiber of the projection $X_d \rightarrow X_{d-1}$, so that $F_d = M_d/((\Gamma \cap G_d) \bmod G_{d+1})$. Let also $M_0 = (M_1)^c = \mathcal{G}/(\mathcal{G} \cap G_2)$.

6.2. Let $P = \{p_1, \dots, p_r\}$ be a system of linear functions $\mathbb{Z}^m \rightarrow \mathbb{Z}$, $p_j(n_1, \dots, n_m) = c_{j,1}n_1 + \dots + c_{j,m}n_m$, $c_{j,i} \in \mathbb{Z}$, $j = 1, \dots, r$. We need to find a rational subgroup H of G^r such that $\pi(H) = \overline{\text{Orb}}_P(\Delta_{X^r})$. We will thus look for the minimal rational subgroup H of G^r that contains all elements $\begin{pmatrix} a^{p_1(n)}b \\ \vdots \\ a^{p_r(n)}b \end{pmatrix}$, $n \in \mathbb{Z}^m$, $b \in G$. Let $\hat{V} = \text{Span}_{\mathbb{Z}} \begin{pmatrix} 1 & p_1 \\ \vdots & \vdots \\ 1 & p_r \end{pmatrix}$; then H is the minimal rational subgroup of G^r that contains all the elements a^v , $v \in \hat{V}$.

6.3. $G^{[\hat{V}]}$ is a rational subgroup of G^r and contains a^v , $v \in \hat{V}$. Thus, $H \subseteq G^{[\hat{V}]}$.

Theorem. $H = G^{[\hat{V}]}$, and so, $\overline{\text{Orb}}_P(\Delta_{X^r}) = \pi^{\times r}(G^{[\hat{V}]})$.

Proof. X_1 is a torus on which G acts by translations. Since the action of a is ergodic on X , the induced action is ergodic on X_1 , and thus the closure of $\overline{\text{Orb}}_P(\Delta_{X_1^r})$ of the orbit of $\Delta_{X_1^r}$ under the action (0.1) is $X_1^{[\hat{V}]}$. $M_0^{[\hat{V}]}$ is the identity component of the group $(\pi^{\times r})^{-1}(X_1^{[\hat{V}]})$, thus $H \bmod G_2^r$ must contain $M_0^{[\hat{V}]}$. Since G is generated by \mathcal{G} and a , M_1 is generated by M_0 and by the element $a \bmod G_2$. Since both $M_0^{[\hat{V}]} \subseteq H \bmod G_2^r$ and $a^v \in H$ for all $v \in \hat{V}$, we obtain that $M_1^{[\hat{V}]} \subseteq H \bmod G_2^r$. By Proposition 5.13, $H = G^{[\hat{V}]}$. ■

6.4. For every $d \in \mathbb{N}$ we have $(G^{[\hat{V}]} \cap G_d^r) \bmod G_{d+1}^r = M_d^{[\mathcal{P}_d \hat{V}]}$, thus $\overline{\text{Orb}}_P(\Delta_{X^r})$ is a (finite) tower of extensions whose fibers are the groups $\pi^{\times r}(M_d^{[\mathcal{P}_d \hat{V}]}) = F_d^{[\mathcal{P}_d \hat{V}]}$. Let $V = \text{span}_{\mathbb{R}} \hat{V}$; since F_d are connected, $\pi^{\times r}(M_d^{[\mathcal{P}_d \hat{V}]}) = F_d^{[V^{*d}]}$ for all d .

6.5. Corollary. $\overline{\text{Orb}}_P(\Delta_{X^r})$ is connected.

6.6. X_d is a characteristic factor for P iff $\overline{\text{Orb}}_P(\Delta_{X^r}) \supseteq D_{d+1}^r$, iff $F_k^{[V^{*k}]} = F_k^r$ for all $k \geq d+1$, iff $V^{*k} = \mathbb{R}^r$ for all $k \geq d+1$ (such that F_k is nontrivial), iff $V^{*(d+1)} = \mathbb{R}^r$. We therefore obtain:

Corollary. The complexity of a system $P = \{p_1, \dots, p_r\}$ is the minimal integer d for which $V^{*(d+1)} = \mathbb{R}^r$.

6.7. Examples. (1) Let P be a system of r different linear functions in 1 variable, $P = \{c_1n, \dots, c_rn\}$. For this system we have $V = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & c_1 \\ \vdots & \vdots \\ 1 & c_r \end{pmatrix}$, and thus

$$V^{*d} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & c_1 & \dots & c_1^d \\ \vdots & \vdots & & \vdots \\ 1 & c_r & \dots & c_r^d \end{pmatrix}$$

for any $d \in \mathbb{N}$. Since c_1, \dots, c_r are all distinct, $\dim V^{*d} = d+1$, $d = 1, 2, \dots$, and $V^{*(d)} = \mathbb{R}^r$ iff $d \geq r-1$. Thus, the complexity of P is $r-2$ (which agrees with the results from [Z2] and [BHK]), and the subgroup $H \subset G^r$ for which we have $\overline{\text{Orb}}_P(\Delta_{X^r}) = \pi^{\times r}(H)$ is, by Corollary 5.8 and subsection 5.11

$$\left\{ \begin{pmatrix} b_0 b_1^{c_1} b_2^{c_1^2} \dots b_{r-2}^{c_1^{r-2}} \\ b_0 b_1^{c_2} b_2^{c_2^2} \dots b_{r-2}^{c_2^{r-2}} \\ \vdots \\ b_0 b_1^{c_r} b_2^{c_r^2} \dots b_{r-2}^{c_r^{r-2}} \end{pmatrix} : b_0 \in G, b_1 \in G_1, \dots, b_{r-2} \in G_{r-2} \right\} \cdot G_{r-1}.$$

(2) Consider a system $P = \{c_1n_1 + d_1n_2, \dots, c_7n_1 + d_7n_2\}$ of 7 linear functions in 2 variables. Here $V = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & c_1 & d_1 \\ \vdots & \vdots & \vdots \\ 1 & c_7 & d_7 \end{pmatrix}$, $V^{*2} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & c_1 & d_1 & c_1^2 & c_1d_1 & d_1^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & c_7 & d_7 & c_7^2 & c_7d_7 & d_7^2 \end{pmatrix}$, and $V^{*3} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & c_1 & d_1 & c_1^2 & c_1d_1 & d_1^2 & c_1^3 & c_1^2d_1 & c_1d_1^2 & d_1^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & c_7 & d_7 & c_7^2 & c_7d_7 & d_7^2 & c_7^3 & c_7^2d_7 & c_7d_7^2 & d_7^3 \end{pmatrix}$, which equals \mathbb{R}^7 if the vectors $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, $\begin{pmatrix} c_1 \\ \vdots \\ c_7 \end{pmatrix}$ and $\begin{pmatrix} d_1 \\ \vdots \\ d_7 \end{pmatrix}$ are in the “general position” (in the sense, clear from the context). Thus, the complexity of a generic system of 7 linear equations in 2 variables is equal to 2 and

$$H = \left\{ \begin{pmatrix} b_0 b_1^{c_1} b_2^{d_1} b_3^{c_1^2} b_4^{c_1d_1} b_5^{d_1^2} \\ \vdots \\ b_0 b_1^{c_7} b_2^{d_7} b_3^{c_7^2} b_4^{c_7d_7} b_5^{d_7^2} \end{pmatrix} : b_0, b_1, b_2 \in G, b_3, b_4, b_5 \in G_2 \right\} \cdot G_3^7.$$

The complexity of P may vary if the vectors $e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, $v = \begin{pmatrix} c_1 \\ \vdots \\ c_7 \end{pmatrix}$ and $w = \begin{pmatrix} d_1 \\ \vdots \\ d_7 \end{pmatrix}$ are not in the general position. If these vectors are linearly dependent, say, w is a linear combination of e and v , then $V = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & c_1 \\ \vdots & \vdots \\ 1 & c_7 \end{pmatrix}$, and the complexity of P equals 5 as in Example (1) above. If, say, $w = v^2$, one checks that the complexity of P is 3, etc.

(3) It can be observed from the preceding example that the complexity of the generic system of r linear functions in m variables is equal to the minimal integer d such that $\binom{m+d+1}{m} \geq r$.

(4) Let $m \in \mathbb{N}$ and let $P = \{p_S : S \subseteq \{1, \dots, m\}\}$ be the system of 2^m linear functions $p_S = \sum_{i \in S} n_i$ in the variables n_1, \dots, n_m . (Systems of this form, introduced in [B] under the name of *Khinchine's systems*, play a crucial role in [HK1].) Let \mathcal{S} be

the power set of $\{1, \dots, m\}$ and let $\{u_S : S \in \mathcal{S}\}$ be the natural basis in $\mathbb{R}^{\mathcal{S}}$. Then $V = \text{span}_{\mathbb{R}}\{v_{\emptyset}, v_{\{1\}}, \dots, v_{\{m\}}\}$, where $v_{\{i\}} = \sum_{S \ni i} u_S$ and $v_{\emptyset} = \sum_{S \in \mathcal{S}} u_S$. For any $d \in \mathbb{N}$, $V^{*d} = \text{span}_{\mathbb{R}}\{v_R : R \subseteq \{1, \dots, m\}, |R| \leq d\}$, where $v_R = \sum_{S \supseteq R} u_S$. For $d < m$, $V^{*(d)}$ is spanned by less than 2^m vectors, and thus is a proper subspace of $\mathbb{R}^{\mathcal{S}}$. On the other hand, the set of all vectors v_R , $R \in \mathcal{S}$, spans $\mathbb{R}^{\mathcal{S}}$; indeed, for any $S \in \mathcal{S}$ one has $u_S = \sum_{R \supseteq S} (-1)^{|R|-|S|} v_R$. Hence, the complexity of P is $m - 1$ (which agrees with the result from [BHK]) and

$$\begin{aligned} H &= \left\{ \left(\prod_{R \subseteq S} b_R \right)_{S \in \mathcal{S}} : b_{\emptyset} \in G, b_R \in G_{|R|}, R \in \mathcal{S} \right\} \\ &= \left\{ \left(\prod_{R \subseteq S} b_R \right)_{S \in \mathcal{S}} : b_{\emptyset} \in G, b_R \in G_{|R|}, R \in \mathcal{S}, |R| \leq m - 1 \right\} \cdot G_m^r \end{aligned}$$

(where, in the products $\prod_{R \subseteq S} b_R$, the order of elements b_R is arbitrary, but is, of course, the same for all entries of the vector $(\prod_{R \subseteq S} b_R)_{S \in \mathcal{S}}$).

For instance, for $m = 2$ we have $P = \{0, n_1, n_2, n_1 + n_2\}$ and $\mathcal{S} = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$. Ordering the basis in $\mathbb{R}^{\mathcal{S}}$ accordingly we have $V = \text{span}_{\mathbb{R}}\{v_{\emptyset}, v_{\{1\}}, v_{\{2\}}\} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ and $V^{*2} = \text{span}_{\mathbb{R}}\{v_{\emptyset}, v_{\{1\}}, v_{\{2\}}, v_{\{1,2\}}\} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$. Thus,

$$G^{[\hat{V}]} = \left\{ \begin{pmatrix} a_{\emptyset} \\ a_{\emptyset} a_{\{1\}} \\ a_{\emptyset} a_{\{2\}} \\ a_{\emptyset} a_{\{1\}} a_{\{2\}} a_{\{1,2\}} \end{pmatrix} : a_{\emptyset} \in G, a_{\{1\}}, a_{\{2\}} \in G_1, a_{\{1,2\}} \in G_2 \right\}.$$

(5) The following two concrete examples are called upon to demonstrate that the complexity of a system of linear functions may depend not only on “the linear” but also on “the polynomial” properties of the corresponding space \hat{V} . Let

$$P = \{n_1 + n_2, 2n_1 + 4n_2, 3n_1 + 9n_2, 4n_1 + 16n_2, 5n_1 + 25n_2, 6n_1 + 36n_2\}.$$

For this system $V = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{pmatrix} \neq \mathbb{R}^6$, $V^{*2} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 16 \\ 1 & 3 & 9 & 27 & 81 \\ 1 & 4 & 16 & 64 & 256 \\ 1 & 5 & 25 & 125 & 625 \\ 1 & 6 & 36 & 216 & 1296 \end{pmatrix} \neq \mathbb{R}^6$, and finally $V^{*3} = \mathbb{R}^6$, so that the complexity of P equals 2.

Now let

$$P = \{n_1 + n_2, 2n_1 + 4n_2, 3n_1 + 9n_2, 4n_1 + 16n_2, 5n_1 + 25n_2, 6n_1 + 37n_2\}.$$

For this system $V = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 37 \end{pmatrix}$ and already $V^{*2} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 4 & 8 & 16 \\ 1 & 3 & 9 & 9 & 27 & 81 \\ 1 & 4 & 16 & 16 & 64 & 256 \\ 1 & 5 & 25 & 25 & 125 & 625 \\ 1 & 6 & 36 & 37 & 222 & 1369 \end{pmatrix} = \mathbb{R}^6$, so that the complexity of P is 1.

7. Polynomial orbits on tori

In this section we remind the reader how polynomial orbits on tori look; we will use this information in subsequent sections.

7.1. In this section let X be a torus, $X = G/\Gamma$ where G is a finite dimensional \mathbb{R} -vector space and Γ is a uniform lattice in G . We will use additive notation for G and X . Any element $\alpha \in X$ acts on X by a translation, and the system (X, α) is a special case of a nilsystem. An advantage of a torus in comparison with a general nilmanifold is that it is a group, which allows one to easily compute linear and polynomial orbits of its points. We are going to discuss here some details related to this.

7.2. If φ is a linear (conventional) action of \mathbb{Z}^m by translations on X , then $Y = \overline{\text{Orb}}_\varphi(x)$ where x is a point of X (or $Y = \overline{\text{Orb}}_\varphi(D)$ where D is a subtorus of X) is a coset of a closed subgroup of X , and so, is either a subtorus or a disjoint finite union of subtori of X . If φ is a polynomial action, then Y is a “fu-subtorus” of X , and so, again, is either a subtorus or a disjoint finite union of subtori of X . This is so since X is a nilmanifold, but also can be easily seen independently. Indeed, let φ be a polynomial mapping of \mathbb{Z}^m to X ; assume that $\varphi(0) = 1_X$ and take $x = 1_X$. For each character χ on X , $\chi \circ \varphi$ is a polynomial $\mathbb{Z}^m \rightarrow \mathbb{T} = \mathbb{R}/\mathbb{Z}$, and thus, by a theorem of Weyl, either $\chi(\varphi(\mathbb{Z}^m))$ is dense and uniformly distributed in \mathbb{T} , or $\chi(\varphi(\mathbb{Z}^m))$ is a finite set of rational points in \mathbb{T} . If the first option holds for all characters on X , then $\overline{\text{Orb}}_\varphi(x) = \overline{\varphi(\mathbb{Z}^m)} = X$. If the second option holds for a nonzero character on X , then it also holds for a primitive character χ (that is, a nonzero character that is not a multiple of any other character). In this case there exists a subgroup F of finite index in \mathbb{Z}^m such that $\chi \circ \varphi$ is constant on the cosets of F , and for each $c \in \text{Im}(\chi \circ \varphi)$ one can consider the restriction of φ on each of the cosets of F on which $\chi \circ \varphi$ takes the value c , replace X by the torus $\chi^{-1}(c)$, and use induction on the dimension of X .

7.3. An element $\alpha \in X$ will be said to be *irrational* if α is not contained in any proper closed subgroup of X ; or equivalently, if $\chi(\alpha)$ is irrational for any character χ on X ; or equivalently, if the coordinates of α are linearly independent over \mathbb{Q} . Clearly, almost all (in any sense) elements of X are irrational. If α is irrational, then the set $\mathbb{Z}\alpha$ is dense in X .

If Y is a closed subgroup of X , we will say that an element $\alpha \in X$ is *irrational modulo* Y if $\alpha \bmod Y$ is an irrational element of X/Y . If this is the case, $\alpha + \beta$ is an irrational element of X for almost all $\beta \in Y$.

7.4. Lemma. *Let α be an irrational element of X and let $p_{i,j}$ be polynomials $\mathbb{Z}^m \rightarrow \mathbb{Z}$ with zero constant term, $i = 1, \dots, k$, $j = 1, \dots, r$. Then*

$$\left\{ \overline{\left(\begin{array}{c} \sum_{j=1}^k p_{1,j}(n)\alpha \\ \vdots \\ \sum_{j=1}^k p_{r,j}(n)\alpha \end{array} \right)}_{n \in \mathbb{Z}^m} \right\} = \text{Span}_X \left(\begin{array}{c} \sum_{j=1}^k p_{1,j} \\ \vdots \\ \sum_{j=1}^k p_{r,j} \end{array} \right).$$

Proof. Denote the left-hand side of this equality by Y and the right-hand side by Z . Clearly, $Y \subseteq Z$. Hence, Y is either a closed subgroup of Z or a union of several components

of a closed subgroup of Z . On the other hand, since $\mathbb{Z}\alpha$ is dense in X , $\mathbb{Z}Y$ is dense in Z ; this is only possible if $Y = Z$. ■

7.5. In the notation from subsection 3.4, same argument proves the following more general statement:

Lemma. *Let α be an irrational element of X , let $p_{i,j}$ be polynomials $\mathbb{Z}^m \rightarrow \mathbb{Z}$ with zero constant term, and let $A_{i,j}$ be linear transformations of X , $i = 1, \dots, k$, $j = 1, \dots, r$. Then*

$$\overline{\left\{ \begin{pmatrix} \sum_{j=1}^k p_{1,j}(n) A_{1,j} \alpha \\ \vdots \\ \sum_{j=1}^k p_{r,j}(n) A_{r,j} \alpha \end{pmatrix} \right\}_{n \in \mathbb{Z}^m}} = \text{Span}_X \begin{pmatrix} \sum_{j=1}^k p_{1,j} A_{1,j} \\ \vdots \\ \sum_{j=1}^k p_{r,j} A_{r,j} \end{pmatrix}.$$

8. A system of polynomial actions – the case of a connected group

In this section we find $\overline{\text{Orb}}_P(\Delta_{X^r})$ in the case where P is a polynomial system and G is connected.

8.1. Let $P = \{p_1, \dots, p_r\}$ where p_i are distinct polynomials $\mathbb{Z}^m \rightarrow \mathbb{Z}$ with $p_i(0) = 0$, $i = 1, \dots, r$. We will, again, look for a subgroup (or a union of several – shifted – subgroups) H of G^r such that $\pi(H) = \overline{\text{Orb}}_P(\Delta_{X^r})$. A problem is that now H does not have to contain the elements $\varphi(n) = \begin{pmatrix} a^{p_1(n)} \\ \vdots \\ a^{p_r(n)} \end{pmatrix}$, $n \in \mathbb{Z}^m$, of G^r (see subsection 1.14), and the argument used in Section 6 no longer works.

8.2. We do not, however, meet this problem if the group G is connected. Indeed, let this be the case, and let $\hat{V} = \text{Span}_{\mathbb{Z}} \begin{pmatrix} 1 & p_1 \\ \vdots & \vdots \\ 1 & p_r \end{pmatrix}$. Then $\varphi(n) \in G^{\hat{V}}$ for every n , so we may assume that $H \subseteq G^{[\hat{V}]}$. On the other hand, if $M_1 = G/G_2$ and X_1 is the torus $M_1/(\Gamma G_2) = G_2 \backslash X$, we have $\overline{\text{Orb}}_P(\Delta_{X_1^r}) = X_1^{[\hat{V}]}$ by Lemma 7.4. Thus $H \text{ mod } G_2^r$ contains the subgroup $M_1^{[\hat{V}]}$; by Proposition 5.13, $H = G^{[\hat{V}]}$.

8.3. We obtain:

Theorem. *If G is connected, then $\overline{\text{Orb}}_P(\Delta_{X^r}) = \pi(G^{[\hat{V}]})$.*

8.4. Let $V = \text{span}_{\mathbb{R}} \hat{V}$. We define the *C-complexity* of the system $P = \{p_1, \dots, p_r\}$ as the minimal d for which X_d is characteristic for the action (0.1) with respect to any ergodic nilsystem ($X = G/\Gamma$, $a \in G$) with a connected G . Copying Corollary 6.6, we get

Corollary. *The C-complexity of P is the minimal integer d for which $V^{*(d+1)} = \mathbb{R}^r$.*

8.5. Recall that if $p_j = c_{j,1}u_1 + \dots + c_{j,l}u_l$, $j = 1, \dots, r$, where u_1, \dots, u_l are distinct monomials (with trivial coefficient) and $c_{j,i} \in \mathbb{Q}$, then $V = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & c_{1,1} & \dots & c_{1,m} \\ \vdots & \vdots & & \vdots \\ 1 & c_{r,1} & \dots & c_{r,m} \end{pmatrix}$.

Examples. (1) For the system

$$P = \{n + n^2, 2n + 4n^2, 3n + 9n^2, 4n + 16n^2, 5n + 25n^2, 6n + 36n^2\}$$

we have $\hat{V} = \text{span}_{\mathbb{Z}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{pmatrix} \neq \mathbb{R}^6$. Thus, $\overline{\text{Orb}}_P(\Delta_{X^6})$ with respect to this action is the same as for the linear system

$$\{n_1 + n_2, 2n_1 + 4n_2, 3n_1 + 9n_2, 4n_1 + 16n_2, 5n_1 + 25n_2, 6n_1 + 36n_2\}$$

considered in the Example 6.7(3), and the C-complexity of P equals 2.

(2) The same is true for, say, the system

$$\{n_1^5 + n_1^2 n_2^3, 2n_1^5 + 4n_1^2 n_2^3, 3n_1^5 + 9n_1^2 n_2^3, 4n_1^5 + 16n_1^2 n_2^3, 5n_1^5 + 25n_1^2 n_2^3, 6n_1^5 + 36n_1^2 n_2^3\}.$$

9. A system of polynomial actions – the case of a Weyl system

In this section we will describe $\overline{\text{Orb}}_P(\Delta_{X^r})$ in the case X is a torus and a is a skew-product transformation of X .

9.1. We will now investigate a situation that is, in some sense, opposite to one considered in Section 8: assume now that \mathcal{G} is commutative and $a \notin \mathcal{G}$. In this case X is a torus, on which a acts as an affine unipotent transformation. Indeed, let $\gamma \in \Gamma$ be such that γ^{-1} belongs to the same connected component as a , so that $a = c\gamma$ with $c \in \mathcal{G}$. For any $x \in X$, $x = \pi(b)$, we have $ax = \pi(ab) = \pi(ab\gamma) = c\pi(\gamma^{-1}b\gamma)$. The mapping $b \mapsto \gamma^{-1}b\gamma$, $b \in \mathcal{G}$, is a linear transformation of \mathcal{G} . For $b \in \mathcal{G}$ let $Ab = b^{-1}\varphi(b) = [b, \gamma]$; since \mathcal{G} is commutative, A is also a linear transformation of \mathcal{G} . Let s be the nilpotency class of \mathcal{G} ; then $A^s = 0$, so that A is nilpotent.

9.2. We start using additive notation for the group \mathcal{G} and for the torus X . The action of A on \mathcal{G} factorizes to an action on X , $Ax = \gamma x - x$, and we have $ax = x + Ax + \alpha$, $x \in X$, where $\alpha = \pi(c) \in X$ and A is a nilpotent linear transformation of X . We call such a system a *Weyl system*. The orbit of the diagonal of a power of a Weyl system under a system of polynomial \mathbb{Z} -actions was studied in [BLLe1]; in this section we partly repeat the argument from [BLLe1].

9.3. For $x \in X$ we have $ax = (A + 1)x + \alpha$, and thus for $n \in \mathbb{Z}$,

$$\begin{aligned} a^n x &= (A + 1)^n x + ((A + 1)^{n-1} + \dots + (A + 1) + 1)\alpha \\ &= (A + 1)^n x + A^{-1}((A + 1)^n - 1)\alpha \\ &= \sum_{d=0}^{s-1} \binom{n}{d} A^d x + \sum_{d=1}^{s-1} \binom{n}{d} A^{d-1} \alpha = x + \sum_{d=1}^s \binom{n}{d} A^{d-1} (Ax + \alpha), \end{aligned} \tag{9.1}$$

where $\binom{n}{d} = \frac{n(n-1)\dots(n-d+1)}{d!}$.

9.4. Let $P = \{p_1, \dots, p_r\}$ be a system of polynomials $\mathbb{Z}^m \rightarrow \mathbb{Z}$ with zero constant term. For a polynomial p and an integer d let us write $p^{[d]}$ for the polynomial $\binom{p}{d} = \frac{1}{d!} p(p-1) \dots (p-d+1)$. The orbit $\text{Orb}_P(\bar{x})$ of a point $\bar{x} = \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix} \in \Delta_{X^r}$ is

$$\text{Orb}_P(\bar{x}) = \left\{ \begin{pmatrix} x + \sum_{d=1}^s p_1^{[d]}(n) A^{d-1}(Ax + \alpha) \\ \vdots \\ x + \sum_{d=1}^s p_r^{[d]}(n) A^{d-1}(Ax + \alpha) \end{pmatrix} \right\}_{n \in \mathbb{Z}^m} = \bar{x} + \left\{ \begin{pmatrix} \sum_{d=1}^s p_1^{[d]}(n) A^{d-1}(Ax + \alpha) \\ \vdots \\ \sum_{d=1}^s p_r^{[d]}(n) A^{d-1}(Ax + \alpha) \end{pmatrix} \right\}_{n \in \mathbb{Z}^m}.$$

The action of a is ergodic on X , thus it is ergodic on X/AX . Since $ax = x + \alpha \pmod{AX}$, α is irrational modulo AX . Thus, for almost all $x \in X$ the element $\alpha + Ax$ is irrational in X . When x is such that $Ax + \alpha \in X$ is irrational in X , by Lemma 7.5, $\overline{\text{Orb}}_P \bar{x}$ is the torus

$$\overline{\text{Orb}}_P(\bar{x}) = \bar{x} + \text{Span}_X \begin{pmatrix} \sum_{d=1}^s p_1^{[d]} A^{d-1} \\ \vdots \\ \sum_{d=1}^s p_r^{[d]} A^{d-1} \end{pmatrix},$$

(and is contained in this torus if $Ax + \alpha$ is not irrational). We obtain:

Theorem. *In the case of a Weyl system,*

$$\overline{\text{Orb}}_P(\Delta_{X^r}) = \Delta_{X^r} + \text{Span}_X \begin{pmatrix} \sum_{d=1}^s p_1^{[d]} A^{d-1} \\ \vdots \\ \sum_{d=1}^s p_r^{[d]} A^{d-1} \end{pmatrix}. \quad (9.2)$$

In particular, $\overline{\text{Orb}}_P(\Delta_{X^r})$ is connected.

9.5. Remark. Notice that $\overline{\text{Orb}}_P(\Delta_{X^r})$ depends not only on the system P but on the transformation A , that is, on the element $a \in G$ that defines the nil-rotation of X . This is in contrast to the “linear” and the “connected polynomial” cases, considered in Sections 6 and 8, where $\overline{\text{Orb}}_P(\Delta_{X^r})$ only depended on P .

9.6. The W -complexity of a system $P = \{p_1, \dots, p_r\}$ is the minimal d for which the d^{th} natural factor X_d is characteristic with respect to the action (0.1) for any ergodic Weyl system (X, a) . In the case of a Weyl system $X_d = X/A^{d+1}X$, and X_d is characteristic iff the torus $H = \overline{\text{Orb}}_P(\Delta_{X^r})$ contains $(A^{d+1}X)^r$.

To compute the W -complexity of P we replace (X, A) by a more convenient system, for which (X, A) is a factor. Put $\tilde{X} = X^s$ and define $\tilde{A}: \tilde{X} \rightarrow \tilde{X}$ by $\tilde{A}(x_1, \dots, x_s) = (0, x_1, \dots, x_{s-1})$, $x_1, \dots, x_s \in X$. Define $\sigma: \tilde{X} \rightarrow X$ by $\sigma(x_1, x_2, \dots, x_s) = x_1 + Ax_2 + \dots + A^{s-1}x_s$, $x_1, \dots, x_s \in X$. Then $\sigma \circ \tilde{A} = A \circ \sigma$, so that the system (X, A) is a factor of the system (\tilde{X}, \tilde{A}) .

Consider the torus \tilde{X}^r . For the point $\begin{pmatrix} \tilde{x} \\ \vdots \\ \tilde{x} \end{pmatrix} \in \tilde{X}^r$, $\tilde{x} = (x_1, \dots, x_s) \in \tilde{X}$, and $n \in \mathbb{Z}^m$ we have

$$\begin{pmatrix} \sum_{d=1}^s p_1^{[d]}(n) A^{d-1} \tilde{x} \\ \vdots \\ \sum_{d=1}^s p_r^{[d]}(n) A^{d-1} \tilde{x} \end{pmatrix} = \begin{pmatrix} p_1(n)x_1, p_1(n)x_2 + p_1^{[2]}(n)x_1, \dots, \sum_{d=1}^s p_1^{[d]}(n)x_{s-d+1} \\ \vdots \\ p_r(n)x_1, p_r(n)x_2 + p_r^{[2]}(n)x_1, \dots, \sum_{d=1}^s p_r^{[d]}(n)x_{s-d+1} \end{pmatrix},$$

and so,

$$\tilde{H} = \Delta_{\tilde{X}^r} + \text{Span}_{\tilde{X}} \left(\begin{array}{c} \sum_{d=1}^s p_1^{[d]} A^{d-1} \\ \vdots \\ \sum_{d=1}^s p_r^{[d]} A^{d-1} \end{array} \right) = \text{Span}_X \left(\begin{array}{cccccccc} 1 & 0 & \dots & 0 & 0 & p_1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & p_1^{[2]} & p_1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & p_1^{[s-1]} & p_1^{[s-2]} & \dots & p_1 & 0 \\ 0 & 0 & \dots & 0 & 1 & p_1^{[s]} & p_1^{[s-1]} & \dots & p_1^{[2]} & p_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \dots & 0 & 0 & p_r & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & p_r^{[2]} & p_r & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & p_r^{[s-1]} & p_r^{[s-2]} & \dots & p_r & 0 \\ 0 & 0 & \dots & 0 & 1 & p_r^{[s]} & p_r^{[s-1]} & \dots & p_r^{[2]} & p_r \end{array} \right) \subseteq \tilde{X}^r.$$

For convenience, let us now rearrange the coordinates of \tilde{X}^r so that it is identified with $(X^r)^s$ instead of $(X^s)^r$. Then $\tilde{H} = \text{span}_X W$ where

$$W = \text{Span}_{\mathbb{R}} \left(\begin{array}{cccccccccccc} 1 & 0 & \dots & 0 & 0 & p_1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \dots & 0 & 0 & p_r & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & p_1^{[2]} & p_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 & p_r^{[2]} & p_r & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & p_1^{[s-1]} & p_1^{[s-2]} & p_1^{[s-3]} & \dots & p_1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & p_r^{[s-1]} & p_r^{[s-2]} & p_r^{[s-3]} & \dots & p_r & 0 \\ 0 & 0 & \dots & 0 & 1 & p_1^{[s]} & p_1^{[s-1]} & p_1^{[s-2]} & \dots & p_1^{[2]} & p_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & p_r^{[s]} & p_r^{[s-1]} & p_r^{[s-2]} & \dots & p_r^{[2]} & p_r \end{array} \right) \subseteq (\mathbb{R}^r)^s. \quad (9.3)$$

The d^{th} natural factor \tilde{X}_d , $d = 0, \dots, s$, of the system (\tilde{X}, \tilde{A}) is the torus X^d , with the projection $\pi_d: \tilde{X} \rightarrow \tilde{X}_d$ given by $\pi_d(x_1, \dots, x_s) = (x_1, \dots, x_d)$, $x_i \in X$; the fiber of π_d is the subtorus $\tilde{D}_{d+1} = \tilde{A}^d \tilde{X} = \{0\}^d \times X^{s-d}$ of \tilde{X} . For $0 \leq d \leq s$, let $R_d = \{0\}^{d-1} \times \mathbb{R}^{s-d+1}$; then $\tilde{D}_{d+1} = \text{span}_X R_{d+1}$ and $\tilde{D}_{d+1}^r \cap \tilde{H} = \text{span}_X (R_{d+1}^r \cap W)$, so that $\tilde{D}_{d+1}^r \subseteq \tilde{H}$ iff $R_{d+1}^r \subseteq W$. Let $W_d = (W \cap R_d^r) \text{ mod } R_{d+1}^r$, viewed as a subspace of $R_d^r / R_{d+1}^r \simeq \mathbb{R}^r$. It is seen from formula (9.3) that the spaces $W_d \subseteq \mathbb{R}^r$ are nested, $W_0 \subseteq W_1 \subseteq \dots \subseteq W_s$, thus $W \supseteq R_{d+1}^r$ iff $W_{d+1} = \mathbb{R}^r$. We obtain:

9.7. Theorem. *The W-complexity of P equals the minimal d for which the subspace*

$$W \bmod R_{d+2}^r = \text{Span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & p_1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \dots & 0 & 0 & p_r & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & p_1^{[2]} & p_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 & p_r^{[2]} & p_r & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & p_1^{[d]} & p_1^{[d-1]} & p_1^{[d-2]} & \dots & p_1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & p_r^{[d]} & p_r^{[d-1]} & p_r^{[d-2]} & \dots & p_r & 0 \\ 0 & 0 & \dots & 0 & 1 & p_1^{[d+1]} & p_1^{[d]} & p_1^{[d-1]} & \dots & p_1^{[2]} & p_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & p_r^{[d+1]} & p_r^{[d]} & p_r^{[d-1]} & \dots & p_r^{[2]} & p_r \end{pmatrix}$$

of $(\mathbb{R}^r)^{d+1}$ contains the r -dimensional subspace $\{0\}^{rd} \times \mathbb{R}^r$.

When the W-complexity of P is d , formula (9.2) can be rewritten as

$$\overline{\text{Orb}}_P(\Delta_{X^r}) = \Delta_{X^r} + \text{Span}_X \begin{pmatrix} \sum_{k=1}^d p_1^{[k]} A^{k-1} \\ \vdots \\ \sum_{k=1}^d p_r^{[k]} A^{k-1} \end{pmatrix} + D_{d+1}^r,$$

where $D_{d+1}^r = A^d(X)$.

9.8. Examples. Consider the systems $P_1 = \{0, n, 2n, n^3\}$ and $P_2 = \{0, n, 2n, n^2\}$. If G is connected, then, for both systems, $\overline{\text{Orb}}_P(\Delta_{X^4}) = G^{[V]}$ where $V = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$; we thus

have $V^{*2} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 2 & 0 & 4 \\ 1 & 0 & 1 & 0 \end{pmatrix} = \mathbb{R}^4$ and the C-complexity of both P_1 and P_2 is equal to 1.

(1) Now let (X, a) be a Weyl system. For the space W from (9.3) corresponding to the system P_1 we have

$$W_1 = W \bmod R_2^4 = \text{Span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 \\ 1 & n \\ 1 & 2n \\ 1 & n^3 \end{pmatrix} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

and

$$W \bmod R_3^4 = \text{Span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 \\ 1 & n & 0 \\ 1 & 2n & 0 \\ 1 & n^3 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & \frac{n(n-1)}{2} & n \\ 0 & 1 & \frac{2n(2n-1)}{2} & 2n \\ 0 & 1 & \frac{n^3(n^3-1)}{2} & n^3 \end{pmatrix} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & \frac{1}{2} & 0 & 0 & 1 \\ 0 & 1 & -1 & 2 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & \frac{-1}{2} & \frac{1}{2} & 0 \end{pmatrix},$$

thus $W_2 = (W \cap R_2^4) \bmod R_3^4 = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & \frac{1}{2} & 1 & 0 \\ 1 & 2 & 2 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} = \mathbb{R}^4$, the W-complexity of P_1 is 1 and

$$\overline{\text{Orb}}_P(\Delta_{X^4}) = \text{span}_X \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} + D_2^4.$$

(2) For the system P_2 we have

$$W_1 = W \bmod R_2^4 = \text{Span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 \\ 1 & n \\ 1 & 2n \\ 1 & n^2 \end{pmatrix} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

and

$$W \bmod R_3^4 = \text{Span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 \\ 1 & n & 0 \\ 1 & 2n & 0 \\ 1 & n^2 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & \frac{n(n-1)}{2} & n \\ 0 & 1 & \frac{2n(2n-1)}{2} & 2n \\ 0 & 1 & \frac{n^2(n^2-1)}{2} & n^2 \end{pmatrix} = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & \frac{1}{2} & 0 & 1 \\ 0 & 1 & -1 & 2 & 0 & 2 \\ 0 & 1 & 0 & \frac{-1}{2} & \frac{1}{2} & 0 & 1 \end{pmatrix},$$

thus $W_3 = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \neq \mathbb{R}^4$ yet. $W \bmod R_4^4$ is spanned by a 12×6 polynomial or by a 12×13 numerical matrix (which we do not want to write out) from which one sees that $W_3 = (W \cap R_3^4) \bmod R_4^4 = \mathbb{R}^4$. (See [BLLe1] for more detail.) Hence the W -complexity of P_2 is 2 and

$$\overline{\text{Orb}}_P(\Delta_{X^4}) = \Delta_{X^4} + \text{Span}_X \begin{pmatrix} 0 \\ n + \frac{n(n-1)}{2}A \\ 2n + \frac{2n(2n-1)}{2}A \\ n^2 + \frac{n^2(n^2-1)}{2}A \end{pmatrix} + D_3^4 = \text{span}_X \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2-A & A \\ 1 & 4-2A & 4A \\ 1 & 0 & 2-A \end{pmatrix} + D_3^4.$$

9.9. We see from formula (9.3) that for any d ,

$$\tau_d(W) = \text{Span}_{\mathbb{R}} \begin{pmatrix} 1 & p_1^{[d]} & \dots & p_1^{[2]} & p_1 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & p_r^{[d]} & \dots & p_r^{[2]} & p_r \end{pmatrix} = \text{Span}_{\mathbb{R}} \begin{pmatrix} 1 & p_1 & p_1^2 & \dots & p_1^d \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & p_r & p_r^2 & \dots & p_r^d \end{pmatrix} \subseteq V^{*d},$$

where $V = \text{Span}_{\mathbb{R}} \begin{pmatrix} 1 & p_1 \\ \vdots & \vdots \\ 1 & p_r \end{pmatrix}$. Thus, if $W_d = \mathbb{R}^r$, then all the more $V^{*d} = \mathbb{R}^r$. Recalling Corollary 8.4, we obtain:

Proposition. *For any system of polynomials, the C -complexity does not exceed the W -complexity.*

10. Construction of a polynomial orbit of a point

We now return to the general situation, that is, we no longer assume that G is connected or that \mathcal{G} is commutative. In this section we describe a process which allows one to construct $\overline{\text{Orb}}_{\varphi}(x) = \{\varphi(n)x\}_{n \in \mathbb{Z}^m}$ for a polynomial action φ of \mathbb{Z}^m on X and a point $x \in X$.

10.1. We need the following fact:

Proposition. ([L4]) *Let T be the “maximal factor-torus of X ”, $T = \mathcal{G}_2 \backslash X$, and let $\eta: X \rightarrow T$ be the natural projection. If a polynomial mapping $\varphi: \mathbb{Z}^m \rightarrow G$ is such that $\overline{\{\eta(\varphi(n))\}}_{n \in \mathbb{Z}^m} = T$, then $\overline{\{\varphi(n)\}}_{n \in \mathbb{Z}^m} = X$.*

10.2. Let φ be a polynomial mapping $\mathbb{Z}^m \rightarrow G$ and let $x = \pi(b) \in X$. Define $\tilde{\varphi}(n) = b^{-1}\varphi(0)^{-1}\varphi(n)b$; then $\tilde{\varphi}$ is a polynomial mapping $\mathbb{Z}^m \rightarrow G$ with $\tilde{\varphi}(0) = \mathbf{1}_G$, and $\overline{\text{Orb}}_{\tilde{\varphi}}(x) = b\varphi(0)\overline{\text{Orb}}_{\tilde{\varphi}}(1_X)$. We will therefore look for the fu-subnilmanifold $\overline{\text{Orb}}_{\tilde{\varphi}}(1_X) = \overline{\{\pi(\tilde{\varphi}(n))\}}_{n \in \mathbb{Z}^m}$.

Next, we rewrite $\tilde{\varphi}$ in terms of \mathcal{G} . Namely, if $\tilde{\varphi}(n) = \prod_{i=1}^l b_i^{p_i(n)}$, choose $\beta_i \in \Gamma$ such that $b_i\beta_i^{-1} \in \mathcal{G}$, $i = 1, \dots, l$, and put $\varphi_1(n) = \tilde{\varphi}(n)(\prod_{i=1}^l \beta_i^{p_i(n)})^{-1}$; then $\pi(\tilde{\varphi}(n)) = \pi(\varphi_1(n))$ for all $n \in \mathbb{Z}^m$ and φ_1 takes values in \mathcal{G} .

We now start with $K_1 = \mathcal{G}$, $Z_1 = \mathbb{Z}^m$, and $\varphi_1: Z_1 \rightarrow K_1$. Let M_1 be the commutative group $K_1/[K_1, K_1]$, and let $\eta_1: K_1 \rightarrow M_1$ be the natural projection. Let T_1 be the torus $M_1/\eta_1(\Gamma)$ and let $\tau_1: M_1 \rightarrow T_1$ be the natural projection. Let $\psi_1 = \eta_1 \circ \varphi_1$ and $\xi_1 = \tau_1 \circ \psi_1$, then ψ_1 and ξ_1 are polynomial mappings from Z_1 to M_1 and to T_1 respectively. The orbit $S_1 = \overline{\{\xi_1(n)\}}_{n \in \mathbb{Z}_1}$ in the torus T_1 can be easily determined, as described in subsection 7.2. If $S_1 = T_1$, then by Proposition 10.1 the set $\{\pi(\varphi_1(n))\}_{n \in \mathbb{Z}_1}$ is dense in X , and we are done with $\overline{\text{Orb}}_{\varphi_1}(1_X) = X$. Assume that $S_1 \neq T_1$. If S_1 is connected, we put $Z_2 = Z_1$. If S_1 is disconnected, we can find a subgroup of finite index Z_2 in Z_1 such that $\overline{\{\xi_1(n)\}}_{n \in Z_2}$ is connected for any coset of Z_2 in Z_1 (see subsection 7.2). We may now deal with distinct components of S_1 individually, replacing Z_1 by the corresponding cosets of Z_2 . Let us consider the component $S'_1 = \overline{\{\xi_1(n)\}}_{n \in Z_2}$ only.

S'_1 is a proper subtorus of T_1 . Let $L_1 \subseteq M_1$ be the identity component of $\tau_1^{-1}(S'_1)$. There exists a polynomial mapping $\psi'_1: Z_2 \rightarrow L_1$ such that $\tau_1(\psi'_1) = \xi_1$. Then the polynomial mapping $\delta_1(n) = \psi_1(n)^{-1}\psi'_1(n)$, $n \in Z_2$, takes values in $\eta_1(\Gamma)$. Let γ_1 be any polynomial mapping $Z_1 \rightarrow \Gamma$ such that $\eta_1 \circ \gamma_1 = \delta_1$.

Now let $K_2 = \eta_1^{-1}(L_1)$; then K_2 is a proper connected subgroup of K_1 , and the mapping $\varphi_2(n) = \varphi_1(n)\gamma_1(n)$, $n \in Z_2$, takes values in K_2 and satisfies $\pi(\varphi_2(n)) = \pi(\varphi_1(n))$, $n \in Z_2$. We replace the group K_1 and the polynomial mapping φ_1 by the group K_2 and the polynomial mapping φ_2 .

If needed, we repeat this procedure, until after, say, k repetitions we get $S_k = T_k$; then by Proposition 10.1, $\overline{\text{Orb}}_{\varphi'_1|_{Z_k}}(1_X) = \overline{\text{Orb}}_{\varphi_k}(1_X) = \pi(K_k)$. The other components of $\overline{\text{Orb}}_{\varphi'}(1_X)$ can be found similarly.

10.3. In the process of the construction above we obtained the following result:

Proposition. *Let $\varphi: \mathbb{Z}^m \rightarrow G$ be a polynomial mapping with $\varphi(0) = \mathbf{1}_G$, let $x \in X$, and assume that $Y = \overline{\text{Orb}}_{\varphi}(x)$ is connected. Let K be the closed connected subgroup of G such that $Kx = Y$. Then there exists a polynomial mapping $\varphi': \mathbb{Z}^m \rightarrow K$ such that $\varphi'(n)x = \varphi(n)x$ for all $n \in \mathbb{Z}^m$.*

11. The general case – an algorithm

We will now consider a general nilsystem (X, a) (assuming, as before, for simplicity that X is connected and that G is generated by \mathcal{G} and a). In this case we are unable to obtain simple formulas for $\overline{\text{Orb}}_P(\Delta_{X^r})$; we will only describe a procedure which, in principle, allows one to find $\overline{\text{Orb}}_P(\Delta_{X^r})$. This procedure, however, involves too much computation, and can only be really applied in the simplest situations.

11.1. The case of the Weyl system hints that we have to represent the action of a on X as a unipotent affine transformation of X . Let $a = \alpha\gamma^{-1}$ where $\alpha \in \mathcal{G}$ and $\gamma \in \Gamma$, then G is generated by \mathcal{G} and γ . Let us determine the orbit of 1_X under the action of a . We have $a1_X = \pi(\alpha\gamma^{-1}) = \pi(\alpha)$, $a^21_X = \pi(a\alpha) = \pi(\alpha\gamma^{-1}\alpha) = \pi(\alpha\gamma^{-1}\alpha\gamma) = \pi(\alpha B\alpha)$, where $B(b) = \gamma^{-1}b\gamma$, $b \in \mathcal{G}$. By induction, $a^n1_X = \pi(\alpha B\alpha B^2\alpha \dots B^{n-1}\alpha)$, $n \in \mathbb{N}$.

11.2. For $b \in \mathcal{G}$, define $g_b(n) = (b\gamma^{-1})^n\gamma^n$, $n \in \mathbb{Z}$; g_b is a polynomial sequence in \mathcal{G} , with $g_b \in \text{Pol}^*(\mathbb{Z}, \mathcal{G})$, and for $n \geq 1$ one has $g_b(n) = bBbB^2b \dots B^{n-1}b$. Let $Ab = b^{-1}Bb = [b, \gamma]$, $b \in \mathcal{G}$; we would like to write g_b as a polynomial sequence in terms of the elements $A^k b$, $k \geq 0$, and their commutators, that is, in the form $\prod_{j=1}^{\infty} b_j^{q_j(n)}$, where b_j are commutator expressions of the elements $A^k b$, $k \geq 0$, and q_j are polynomials $\mathbb{Z}^m \rightarrow \mathbb{Z}$. We cannot find simple formulas for the polynomials q_j , but can only suggest a recurrent process that allows one to compute them one by one. We have $Dg_b(n) = g_b(n)^{-1}g_b(n+1) = B^n b$ for $n \in \mathbb{N}$ and hence, for all $n \in \mathbb{Z}$. On the other hand,

$$\begin{aligned} g_{Ab}(n) &= (b^{-1}Bb)B(b^{-1}Bb)B^2(b^{-1}Bb) \dots B^{n-1}(b^{-1}Bb) \\ &= (b^{-1}Bb)((Bb)^{-1}B^2b)((B^2b)^{-1}B^3b) \dots ((B^{n-1}b)^{-1}B^n b) = b^{-1}B^n b, \end{aligned}$$

$n \in \mathbb{N}$, thus

$$Dg_b = bg_{Ab}. \quad (11.1)$$

This equation, with the ‘‘initial condition’’ $g_b(0) = 1$, defines g uniquely.

Let us write g_b in the form

$$g_b(n) = \prod_{1 \leq k < s} (A^{k-1}b)^{q_k(n)} \prod_{1 \leq l < k < s} [A^{k-1}b, A^{l-1}b]^{q_{k,l}(n)} \dots, \quad (11.2)$$

where $q_k, q_{k,l}, \dots$ are polynomials $\mathbb{Z} \rightarrow \mathbb{Z}$ with $\deg q_k \leq k$, $\deg q_{k,l} \leq k+l, \dots$. Then

$$\begin{aligned} Dg_b(n) &= g_b(n)^{-1}g_b(n+1) = \left(\dots \prod_{1 \leq l < k < s} [A^{k-1}b, A^{l-1}b]^{q_{k,l}(n)} \prod_{1 \leq k < s} (A^{k-1}b)^{q_k(n)} \right)^{-1} \\ &\quad \cdot \left(\prod_{1 \leq k < s} (A^{k-1}b)^{q_k(n+1)} \prod_{1 \leq l < k < s} [A^{k-1}b, A^{l-1}b]^{q_{k,l}(n+1)} \dots \right) \\ &= \prod_{1 \leq k < s} (A^{k-1}b)^{Dq_k(n)} \prod_{1 \leq l < k < s} [A^{k-1}b, A^{l-1}b]^{Dq_{k,l}(n) - q_{k,l}(n)} \dots, \end{aligned}$$

$n \in \mathbb{Z}$, where we put $Dq(n) = q(n+1) - q(n)$. By (11.1), this is equal to

$$bg_{Ab}(n) = b \prod_{1 \leq k < s} (A^k b)^{q_k(n)} \prod_{1 \leq l < k < s} [A^k b, A^l b]^{q_{k,l}(n)} \dots$$

Comparing similar terms, we get

$$\begin{aligned} Dq_1 &\equiv b; & Dq_k &= q_{k-1} \text{ for } k \geq 2; \\ Dq_{k,1} &= q_k \text{ for } k \geq 2; & Dq_{k,l} &= q_{k-1,l-1} + q_k Dq_l \text{ for } k > l \geq 2; \quad \text{etc.} \end{aligned}$$

From this equations, we obtain $q_k(n) = \binom{n}{k}$ for $k \geq 1$ (which agrees with formula (9.1) in Section 9); $q_{k,1}(n) = \binom{n}{k+1}$ for $k \geq 2$; $q_{k,2}(n) = (k+1)\binom{n+1}{k+2}$ for $k \geq 3$; etc. Thus, the beginning of (11.2) is

$$g_b(n) = b^n (Ab)^{\binom{n}{2}} (A^2b)^{\binom{n}{3}} \dots [Ab, b]^{\binom{n}{3}} [A^2b, b]^{\binom{n}{4}} \dots [A^2b, Ab]^{4\binom{n+1}{5}} [A^3b, Ab]^{5\binom{n+1}{6}} \dots \quad (11.3)$$

11.3. We have $\pi(a^n 1_X) = \pi(g_\alpha(n))$, $n \in \mathbb{Z}$. Let $x \in X$, $x = \pi(b)$ with $b \in \mathcal{G}$. Then

$$a^n b = b(b^{-1}ab)^n = b(b^{-1}\alpha b b^{-1}\gamma^{-1}b)^n = b(\alpha[\alpha, b][b, \gamma]\gamma^{-1})^n = b(\alpha_b \gamma^{-1})^n$$

where $\alpha_b = \alpha[\alpha, b]Ab$. Hence, $a^n x = b\pi(g_{\alpha_b}(n))$.

Now, given a system $P = \{p_1, \dots, p_r\}$ of polynomials $\mathbb{Z}^m \rightarrow \mathbb{Z}$ with zero constant term, the orbit $\text{Orb}_P(\bar{x})$ of the point $\bar{x} = \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix} \in \Delta_{X^r}$ is $\{\bar{b}\pi^{\times r}(\varphi_x(n)) : n \in \mathbb{Z}^m\}$ where $\bar{b} = \begin{pmatrix} b \\ \vdots \\ b \end{pmatrix} \in \mathcal{G}^r$ and

$$\varphi_x(n) = \begin{pmatrix} g_{\alpha_b}(p_1(n)) \\ \vdots \\ g_{\alpha_b}(p_r(n)) \end{pmatrix}, \quad n \in \mathbb{Z}^m.$$

Assuming that $\overline{\text{Orb}}_P(\Delta_{X^r})$ may be disconnected, we will, for simplicity, confine ourselves to its identity component. For $x \in X$, $x = \pi(b)$ with $b \in \mathcal{G}$, let Y_x be the identity component of $\overline{\{\pi^{\times r}(\varphi_x(n))\}_{n \in \mathbb{Z}^m}}$ in X^r . Then the identity component of $\overline{\text{Orb}}_P(x)$ is $\bar{b}Y_x$. By subsection 1.15, Y_x is the same, up to translation, for almost all x , that is, there exists a subnilmanifold Y of X^r such that $Y_x \subseteq Y$ for all $x \in X$ and $Y_x = Y$ for almost all $x \in X$.

Let $H \subseteq \mathcal{G}^r$ be the identity component of $(\pi^{\times r})^{-1}(\overline{\text{Orb}}_P(\Delta_{X^r}))$. By 1.15, $\pi^{\times r}(H) = \Delta_{\mathcal{G}^r} Y$.

11.4. Consider the orbits $\overline{\text{Orb}}_P(\Delta_{X^r})$ corresponding to different elements $\alpha \in \mathcal{G}$. It follows from subsection 1.15 that these orbits are equal for almost all α , but, in principle, they may be smaller for some α . This cannot happen if the group

$$\widehat{H} = \left\langle \Delta_{\mathcal{G}^r}, \begin{pmatrix} g_\alpha(p_1(n)) \\ \vdots \\ g_\alpha(p_r(n)) \end{pmatrix}, n \in \mathbb{Z}^m, \alpha \in \mathcal{G} \right\rangle$$

is a minimal subgroup of \mathcal{G}^r with the property that

$$\widehat{H} = \Delta_{\mathcal{G}^r} \cdot \left\langle \begin{pmatrix} \prod_{k=1}^s (A^{k-1} \alpha')^{\binom{p_1(n)}{k}} \\ \vdots \\ \prod_{k=1}^s (A^{k-1} \alpha')^{\binom{p_r(n)}{k}} \end{pmatrix}, n \in \mathbb{Z}^m \right\rangle \text{ mod } \mathcal{G}_2^r,$$

where α' is any irrational element of $\mathcal{G}/\mathcal{G}_2$. Indeed, in this case \widehat{H} both contains H and is contained in H , and thus $H = \widehat{H}$ independently of the choice of α . This minimality will follow if the following conjecture is true:

Conjecture. For any $d \in \mathbb{N}$, $\widehat{H} \cap \mathcal{G}_d = \widehat{H}_d$.

All our efforts to either prove or disprove this proposition failed. If it is true, then $\overline{\text{Orb}}_P(\Delta_{X^r})$ is connected, does not depend on α , and is equal to $\pi^{\times r}(\widehat{H})$. If it is not true, $\overline{\text{Orb}}_P(\Delta_{X^r})$ depends on α , and, though this is not automatic, there is a good chance that it may be disconnected.

11.5. We will now get some evident properties of H . Let $\bar{a} = \begin{pmatrix} a \\ \vdots \\ a \end{pmatrix} \in G^r$; then for any $b \in \mathcal{G}$ and any $n \in \mathbb{Z}^m$ we have

$$\bar{a}^{-1} \begin{pmatrix} a^{p_1(n)} b \\ \vdots \\ a^{p_r(n)} b \end{pmatrix} \bar{a} = \begin{pmatrix} a^{p_1(n)} (a^{-1} b a) \\ \vdots \\ a^{p_r(n)} (a^{-1} b a) \end{pmatrix},$$

and so, $\bar{a}^{-1} \text{Orb}_P(\bar{x}) \bar{a} = \text{Orb}_P(\bar{x}')$ where $\bar{x} = \pi^{\times r} \begin{pmatrix} b \\ \vdots \\ b \end{pmatrix}$ and $\bar{x}' = \pi^{\times r} \begin{pmatrix} a^{-1} b a \\ \vdots \\ a^{-1} b a \end{pmatrix}$. So, $H = \bar{a}^{-1} H \bar{a}$. Since $H \supseteq \Delta_{\mathcal{G}^r}$, we obtain:

Proposition. $[H, \Delta_{\mathcal{G}^r}] \subseteq H$; in particular, $AH \subseteq H$.

11.6. Corollary. *If for some $d \in \mathbb{N}$ and a subgroup $\hat{V} \subseteq \mathbb{Z}^r$ one has $H \supseteq G_d^{[\hat{V}]}$ mod G_{d+1}^r , then $H \supseteq G_k^{[\hat{V}]}$ mod G_{k+1}^r for all $k \geq d$.*

Proof. We have

$$H \supseteq [H, \Delta_{\mathcal{G}^r}] \supseteq [G_d^{[\hat{V}]}, \Delta_{\mathcal{G}^r}] \text{ mod } G_{d+2}^r = G_{d+1}^{[\hat{V}]} \text{ mod } G_{d+2}^r$$

by Lemma 5.14. ■

11.7. Corollary. *If, for some $d \geq 2$, $H \supseteq G_d^r$ mod G_{d+1}^r , then $H \supseteq G_d^r$.*

Proof. By Corollary 11.6, $H \supseteq G_k^r$ mod G_{k+1}^r for all $k \geq d$, so $H \supseteq G_d^r$. ■

11.8. Lemma. *Let the C-complexity of P with respect to (X, a) be equal to l . Then $\mathcal{G}_{l+1} \backslash X$ is a characteristic factor of X for P .*

Proof. Let $V = \text{Span}_{\mathbb{R}} \begin{pmatrix} 1 & p_1 \\ \vdots & \vdots \\ 1 & p_r \end{pmatrix}$, then $V^{*(l+1)} = \mathbb{R}^r$. Applying Theorem 9.4 to the Weyl system $\mathcal{G}_2 \backslash X$, we see that $H \supseteq \mathcal{G}^{[V]}$ mod $\mathcal{G}_2^{[V]}$. By Lemma 5.12, $H \supseteq \mathcal{G}_{l+1}^{[V^{*(l+1)}]} = \mathcal{G}_{l+1}^r$. ■

Hence, when computing H , we may replace G by G/\mathcal{G}_{l+1} .

11.9. Example. If the C-complexity of P is 1, then $H \supseteq \mathcal{G}_2$, so we may factorize G by \mathcal{G}_2 and assume that \mathcal{G} is commutative. The system $(\mathcal{G} \backslash X, a)$ is a Weyl system, so $\overline{\text{Orb}}_P(\Delta_{X^r}) = \pi(H)$ is connected and $H = \mathcal{G}^{[W]} \mathcal{G}_2^r$, where W is given by (9.3). In particular, it follows that the complexity of P equals its Weyl complexity. We thus obtain from 9.8 that the complexity of $\{0, n, 2n, n^3\}$ is 1 and the complexity of $\{0, n, 2n, n^2\}$ is 2.

11.10. To find H , we may simplify the general procedure described in subsection 10.2 as follows. First, we replace G by G/\mathcal{G}_{l+1} , where l is the C-complexity of P . Then we replace α by $\alpha_b = \alpha[\alpha, b] A b$ where b is a “generic” element of \mathcal{G} . (Practically, this simply means that the coordinates of b are algebraically independent over $\mathbb{Q}(\alpha)$.) Then we define

$$\varphi(n) = \begin{pmatrix} g_\alpha(p_1(n)) \\ \vdots \\ g_\alpha(p_1(n)) \end{pmatrix}, \quad n \in \mathbb{Z}^m.$$

We start with $H_1 = \mathcal{G}^r$, $Z_1 = \mathbb{Z}^m$, $M_1 = H_1/\mathcal{G}_2^r$, $\eta_1: H_1 \rightarrow M_1$ and $\varphi_1 = \varphi$, and, as in subsection 10.2, determine a subgroup $Z_2 \subseteq Z_1$, a connected subgroup $L_1 \subseteq M_1$, and a polynomial mapping $\varphi_2: Z_2 \rightarrow \eta_1^{-1}(L_1)$ such that $\pi^{\times r}(\varphi_2(n)) = \pi^{\times r}(\varphi_1(n))$, $n \in Z_2$, and the image of $\{\varphi_2(n)\}_{n \in Z_2}$ is dense in the torus $L_1/(\eta_1(\Gamma)^r \cap L_1)$. Then we replace H_1 by the group $H_2 = \eta_1^{-1}(L_1)\Delta_{\mathcal{G}^r}$ and the mapping φ_1 by φ_2 . In the next step, we define M_2 to be the factor $H_2/([H_2, H_2]\mathcal{G}_3^r)$, and repeat the procedure.

Now assume that after d steps we obtain a connected group $H_d \subseteq \mathcal{G}$ and a polynomial mapping $\varphi_d: Z_d \rightarrow H_d$ such that $H_d \supseteq \mathcal{G}_d^r$, $\pi^{\times r} \circ \varphi_d = \pi^{\times r} \circ \varphi_1$ and the image of $\{\varphi_d(n)\}_{n \in Z_d}$ is dense in the torus $H_d/([H_d, H_d]\mathcal{G}_{d+1}^r(\Gamma^r \cap H_d))$. Then, by Proposition 10.1, the image of $\{\varphi_d(n)\}_{n \in Z_d}$ is dense in the nilmanifold $H_d/(\mathcal{G}_{d+1}^r(\Gamma^r \cap H_d))$. Since $H_d \supseteq \mathcal{G}_d^r$, we have $H \supseteq \mathcal{G}_d^r \bmod \mathcal{G}_{d+1}^r$. By Corollary 11.7, $H \supseteq \mathcal{G}_d^r$, so $H = H_d$.

12. A filtration of \mathcal{G}

For the needs of the next section, in which we will give an estimate of the complexity of a general polynomial system, we will now introduce a filtration of \mathcal{G} .

12.1. Assume that the group G is generated by \mathcal{G} and an element $\gamma \in \Gamma$. We define a mapping $A: \mathcal{G} \rightarrow \mathcal{G}$ by $Ab = [b, \gamma]$, $b \in \mathcal{G}$. For $k_1, k_2, \dots, k_l \in \mathbb{N}$ we define

$$\mathcal{G}_{(k_1, k_2, \dots, k_l)} = [A^{k_1-1}\mathcal{G}, A^{k_2-1}\mathcal{G}, \dots, A^{k_l-1}\mathcal{G}].$$

For $k, l \in \mathbb{N}$ with $k \geq l$ we put

$$\mathcal{G}_{k,l} = \langle \mathcal{G}_{(k_1, \dots, k_i)} : i \geq l, k_1 + \dots + k_i \geq k \rangle.$$

Clearly, $\mathcal{G}_{l,l} = \mathcal{G}_l$ for all l . We will also assume $\mathcal{G}_{k,l} = \mathcal{G}_l$ for $k < l$.

12.2. Using, when needed, the standard commutator identities

$$\begin{aligned} [b_1, b_2 b_3] &= [b_1, b_2][b_1, b_2, b_3][b_1, b_3], & [b_1 b_2, b_3] &= [b_1, b_3][b_1, b_3, b_2][b_2, b_3], \\ [b_1^{-1}, b_2] &= [b_1, b_2]^{-1}[b_2, b_1, b_1^{-1}] & \text{and} & [b_1, b_2^{-1}] &= [b_1, b_2]^{-1}[b_2, b_1, b_2^{-1}], \end{aligned} \tag{12.1}$$

one gets:

Lemma. For any $k, l \in \mathbb{N}$ with $k \geq l$,

- (i) $\mathcal{G}_{k,l}$ is normal in G ;
- (ii) $[\mathcal{G}_{k,l}, \mathcal{G}_{j,i}] \subseteq \mathcal{G}_{k+j, l+i}$ for any $j, i \in \mathbb{N}$, $j \geq i$;
- (iii) $A^j \mathcal{G}_{k,l} \subseteq \mathcal{G}_{k+j, l}$ for any $j \in \mathbb{N}$;
- (iv) the factor groups $\mathcal{G}_{k,l}/\mathcal{G}_{k, l+1}$ and $\mathcal{G}_{k,l}/\mathcal{G}_{k+1, l}$ are commutative.

12.3. Lemma. For all $k \geq 2$, $G_k = \mathcal{G}_{k,1}$.

Proof. Clearly, $\mathcal{G}_{k,1} \subseteq G_k$ for all k . For $k = 2$, using Lemma 12.2, we have

$$G_2 = [G, G] = [\langle \mathcal{G}, \gamma \rangle, \langle \mathcal{G}, \gamma \rangle] = \langle [\mathcal{G}, \gamma], [\mathcal{G}, \mathcal{G}], \mathcal{G}_{3,1} \rangle \subseteq \mathcal{G}_{2,1},$$

so $G_2 = \mathcal{G}_{2,1}$. If, by induction, $G_k = \mathcal{G}_{k,1}$ for some $k \geq 2$, then

$$G_{k+1} = [G_k, G] = [\mathcal{G}_{k,1}, \langle \mathcal{G}, \gamma \rangle] = \langle [\mathcal{G}_{k,1}, \gamma], [\mathcal{G}_{k,1}, \mathcal{G}], \mathcal{G}_{k+2,1} \rangle \subseteq \mathcal{G}_{k+1,1},$$

so $G_{k+1} = \mathcal{G}_{k+1,1}$. ■

12.4. Thus, we have the following filtration of \mathcal{G} :

$$\begin{array}{ccccccc}
& & & & & & \mathcal{G}_1 \\
& & & & & & \parallel \\
\mathcal{G} & = & \mathcal{G}_{1,1} & & \mathcal{G}_2 & & \\
& & \cup & & \parallel & & \\
\mathcal{G}_2 & = & \mathcal{G}_{2,1} \supseteq \mathcal{G}_{2,2} & & \mathcal{G}_3 & & \\
& & \cup & & \cup & & \parallel \\
\mathcal{G}_3 & = & \mathcal{G}_{3,1} \supseteq \mathcal{G}_{3,2} \supseteq \mathcal{G}_{3,3} & & & & \\
& & \vdots & & \vdots & & \ddots & & \mathcal{G}_s \\
& & \cup & & \cup & & \cup & & \parallel \\
\mathcal{G}_s & = & \mathcal{G}_{s,1} \supseteq \mathcal{G}_{s,2} \supseteq \mathcal{G}_{s,3} \supseteq \dots \supseteq \mathcal{G}_{s,s} & & & & \\
& & \cup & & & & \\
\{\mathbf{1}_G\} & = & \mathcal{G}_{s+1} = \mathcal{G}_{s+1,1} & & & &
\end{array}$$

12.5. Lemma. For any $k, l \in \mathbb{N}$ with $k \geq l \geq 2$ one has $\mathcal{G}_{k,l} = \langle A\mathcal{G}_{k-1,l}, [\mathcal{G}_{k-1,l-1}, \mathcal{G}] \rangle$.

Proof. Denote $\langle A\mathcal{G}_{k-1,l}, [\mathcal{G}_{k-1,l-1}, \mathcal{G}] \rangle$ by Q . By Lemma 12.2, $\mathcal{G}_{k,l} \supseteq Q$.

Let $k_1, \dots, k_l \in \mathbb{N}$ with $k_1 + \dots + k_l = k$. If $k_l = 1$, then

$$\mathcal{G}_{(k_1, \dots, k_{l-1}, k_l)} = [\mathcal{G}_{(k_1, \dots, k_{l-1})}, \mathcal{G}] \subseteq [\mathcal{G}_{k-1, l-1}, \mathcal{G}] \subseteq Q.$$

Let $k_l \geq 2$. For any $b_1, b_2, \dots, b_{l-1}, b_l \in \mathcal{G}$ one gets from formulas (12.1) that

$$\begin{aligned}
& A[A^{k_1-1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-2}b_l] \\
& = [A^{k_1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-2}b_l] \\
& \quad [A^{k_1-1}b_1, A^{k_2}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-2}b_l] \dots \\
& \quad [A^{k_1-1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}}b_{l-1}, A^{k_l-2}b_l] \\
& \quad [A^{k_1-1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-1}b_l] \text{ mod } \mathcal{G}_{k+1, l+1},
\end{aligned}$$

and thus

$$\begin{aligned}
& [A^{k_1-1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-1}b_l] \\
& = [A^{k_1-1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}}b_{l-1}, A^{k_l-2}b_l]^{-1} \dots \\
& \quad [A^{k_1-1}b_1, A^{k_2}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-2}b_l]^{-1} \\
& \quad [A^{k_1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-2}b_l]^{-1} \\
& \quad A[A^{k_1-1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-2}b_l] \text{ mod } \mathcal{G}_{k+1, l+1}.
\end{aligned}$$

We have $A[A^{k_1-1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-2}b_l] \in A\mathcal{G}_{k-1, l}$, and by induction on k_l (with fixed k),

$$\begin{aligned}
& [A^{k_1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-2}b_l], [A^{k_1-1}b_1, A^{k_2}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-2}b_l], \dots, \\
& [A^{k_1-1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}}b_{l-1}, A^{k_l-2}b_l] \in Q \text{ mod } \mathcal{G}_{k+1, l+1},
\end{aligned}$$

thus $[A^{k_1-1}b_1, A^{k_2-1}b_2, \dots, A^{k_{l-1}-1}b_{l-1}, A^{k_l-1}b_l] \in Q \text{ mod } \mathcal{G}_{k+1, l+1}$.

We have obtained that for any k_1, \dots, k_l with $k_1 + \dots + k_l = k$, $\mathcal{G}_{(k_1, k_2, \dots, k_{l-1}, k_l)} \subseteq Q \text{ mod } \mathcal{G}_{k+1, l+1}$. Hence, $\mathcal{G}_{k,l} \subseteq Q \text{ mod } (\mathcal{G}_{k+1, l} \mathcal{G}_{k, l+1})$. By induction on decreasing k , $\mathcal{G}_{k+1, l} \subseteq \langle A\mathcal{G}_{k,l}, [\mathcal{G}_{k, l-1}, \mathcal{G}] \rangle \subseteq Q$, and by induction on decreasing l , $\mathcal{G}_{k, l+1} \subseteq \langle A\mathcal{G}_{k-1, l+1}, [\mathcal{G}_{k-1, l}, \mathcal{G}] \rangle \subseteq Q$, thus $\mathcal{G}_{k,l} \subseteq Q$. ■

12.6. Corollary. For any $r, k, l \in \mathbb{N}$ with $k \geq l \geq 2$ one has $\mathcal{G}_{k,l}^r = \langle A^{\times r} \mathcal{G}_{k-1,l}, [\mathcal{G}_{k-1,l-1}^r, \Delta_{\mathcal{G}^r}] \rangle$.

Proof. It follows from Lemma 12.2 that $\mathcal{G}_{k,l}^r \supseteq \langle A^{\times r} \mathcal{G}_{k-1,l}, [\mathcal{G}_{k-1,l-1}^r, \Delta_{\mathcal{G}^r}] \rangle$. For any $i \in \{1, \dots, r\}$ we have

$$\begin{aligned} \text{row } i \left(\begin{array}{c} \mathbf{1}_G \\ \vdots \\ \mathbf{1}_G \\ \mathcal{G}_{k,l}^r \\ \mathbf{1}_G \\ \vdots \\ \mathbf{1}_G \end{array} \right) &= \left(\langle A \mathcal{G}_{k-1,l}, [\mathcal{G}_{k-1,l-1}, \mathcal{G}] \rangle \right) = \left\langle \left(\begin{array}{c} \mathbf{1}_G \\ \vdots \\ \mathbf{1}_G \\ A \mathcal{G}_{k-1,l} \\ \mathbf{1}_G \\ \vdots \\ \mathbf{1}_G \end{array} \right), \left(\begin{array}{c} \mathbf{1}_G \\ \vdots \\ \mathbf{1}_G \\ [\mathcal{G}_{k-1,l-1}, \mathcal{G}] \\ \mathbf{1}_G \\ \vdots \\ \mathbf{1}_G \end{array} \right) \right\rangle \\ &= \left\langle A^{\times r} \left(\begin{array}{c} \mathbf{1}_G \\ \vdots \\ \mathbf{1}_G \\ \mathcal{G}_{k-1,l} \\ \mathbf{1}_G \\ \vdots \\ \mathbf{1}_G \end{array} \right), \left[\left(\begin{array}{c} \mathbf{1}_G \\ \vdots \\ \mathbf{1}_G \\ \mathcal{G}_{k-1,l-1} \\ \mathbf{1}_G \\ \vdots \\ \mathbf{1}_G \end{array} \right), \Delta_{\mathcal{G}^r} \right] \right\rangle \subseteq \langle A^{\times r} \mathcal{G}_{k-1,l}, [\mathcal{G}_{k-1,l-1}^r, \Delta_{\mathcal{G}^r}] \rangle. \end{aligned}$$

■

13. The general case – estimation of complexity

The algorithm described in subsection 11.10 is too complicated for practical usage. We are therefore going to obtain (in Theorem 13.13 and Corollary 13.14) some estimates of the complexity of P .

13.1. We will introduce more linear algebra notation. Let $r \in \mathbb{N}$. We put $R = \mathbb{R}^s$, and sometimes interpret R^r as $(\mathbb{R}^r)^s$. For $w_1, \dots, w_l \in R^r = (\mathbb{R}^r)^s$, where $w_i = \begin{pmatrix} v_{i,1} \\ \vdots \\ v_{i,s} \end{pmatrix}$ with $v_{i,j} = \begin{pmatrix} c_{i,j,1} \\ \vdots \\ c_{i,j,r} \end{pmatrix} \in \mathbb{R}^r$, $i = 1, \dots, l$, $j = 1, \dots, r$, let $w_1 \otimes \dots \otimes w_l$ be the element

$$\left(\begin{array}{c} \left(\begin{array}{c} c_{1,1,1} \\ \vdots \\ c_{1,s,1} \end{array} \right) \otimes \dots \otimes \left(\begin{array}{c} c_{l,1,1} \\ \vdots \\ c_{l,s,1} \end{array} \right) \\ \vdots \\ \left(\begin{array}{c} c_{1,1,r} \\ \vdots \\ c_{1,s,r} \end{array} \right) \otimes \dots \otimes \left(\begin{array}{c} c_{l,1,r} \\ \vdots \\ c_{l,s,r} \end{array} \right) \end{array} \right) \in (R^{\otimes l})^r.$$

For subspaces W_1, \dots, W_l of R^r we define

$$W_1 \otimes \dots \otimes W_l = \text{span}_{\mathbb{R}} \{ w_1 \otimes \dots \otimes w_l : w_1 \in W_1, \dots, w_l \in W_l \} \subseteq (R^{\otimes l})^r,$$

and for a subspace W of R^r and $l \in \mathbb{N}$ we define $W^{\otimes l} = W \otimes \dots \otimes W$.

13.2. Consider the filtration $R = R_1 \subset R_2 \subset \dots \subset R_s \subset R_{s+1} = R_{s+2} = \dots = \{0\}$ of R , where $R_k = \{0\}^{k-1} \times (\mathbb{R})^{s-k+1}$ for $k \in \{1, \dots, s\}$. For $l, k_1, \dots, k_l \in \mathbb{N}$ we define $R_{(k_1, \dots, k_l)} = R_{k_1} \otimes \dots \otimes R_{k_l}$, and for $k, l \in \mathbb{N}$, $k \geq l$, we define $R_{k,l} = \sum_{k_1 + \dots + k_l = k} R_{(k_1, \dots, k_l)}$.

13.3. Let W be a subspace of R^r ; it will be important for us to know for what $k, l \in \mathbb{N}$ one has $W^{\otimes l} \supseteq R_{k,l}^r$. The following lemma helps determine such k, l :

Lemma. (i) For any $l, k_1, \dots, k_l \in \mathbb{N}$, $W^{\otimes l} \supseteq R_{(k_1, \dots, k_l)}^r$ implies $W^{\otimes l} \supseteq R_{(k_{\sigma(1)}, \dots, k_{\sigma(l)})}^r$ for any permutation σ of $\{1, \dots, l\}$.

(ii) Assume that $W \supseteq \Delta_{R^r}$. Then for any $l, k_1, \dots, k_l \in \mathbb{N}$, $W^{\otimes l} \supseteq R_{(k_1, \dots, k_l)}^r$ implies $W^{\otimes(l+1)} \supseteq R_{(k_1, \dots, k_l, k_{l+1})}^r$ for all $k_{l+1} \in \mathbb{N}$.

(iii) Assume that W is such that for all $k \in \mathbb{N}$ one has $(W \cap R_{k+1}^r) \bmod R_{k+2}^r \supseteq (W \cap R_k^r) \bmod R_{k+1}^r$ (both interpreted as subspaces of \mathbb{R}^r). Then for any $k, l \in \mathbb{N}$ with $k \geq l$, $W^{\otimes l} \supseteq R_{k,l}^r \bmod R_{k+1,l}^r$ implies $W^{\otimes l} \supseteq R_{k,l}^r$.

Proof. (i) follows from the fact that the product $W^{\otimes l}$ is invariant with respect to permutations of its factors. (ii) follows from the fact that $R_{(k_1, \dots, k_l)}^r \otimes \Delta_{R_{k_{l+1}}^r} = R_{(k_1, \dots, k_l, k_{l+1})}^r$.

(iii) Let $S: R \rightarrow R$ be the coordinate shift, $S(c_1, c_2, \dots, c_s) = (0, c_1, \dots, c_{s-1})$; by the assumption, $S^{\times r}(W) \subseteq W$. Assume that $W^{\otimes l} \supseteq R_{k,l}^r \bmod R_{k+1,l}^r$ for some k, l . Then for any $k_1, \dots, k_l \in \mathbb{N}$ with $k_1 + \dots + k_l = k$,

$$W^{\otimes l} \supseteq (S^{\times r} \otimes \text{Id}_{R^r} \otimes \dots \otimes \text{Id}_{R^r}) W^{\otimes l} \supseteq (S^{\times r} \otimes \text{Id}_{R^r} \otimes \dots \otimes \text{Id}_{R^r})(R_{(k_1, \dots, k_l)}^r \bmod R_{k+1,l}^r) \\ = R_{(k_1+1, \dots, k_l)}^r \bmod R_{k+2,l}^r.$$

Similarly, $W^{\otimes l} \supseteq R_{(k_1, \dots, k_{i-1}, k_i+1, k_{i+1}, \dots, k_l)}^r \bmod R_{k+2,l}^r$ for any $i \in \{1, \dots, l\}$. Hence, $W^{\otimes l} \supseteq R_{k+1,l}^r \bmod R_{k+2,l}^r$. By induction on j , $W^{\otimes l} \supseteq R_{j,l}^r \bmod R_{j+1,l}^r$ for all $j \geq k$, and thus $W^{\otimes l} \supseteq R_{k,l}^r$. ■

13.4. Now let $a \in G$, and let $\gamma \in \Gamma$ is such that $a\gamma^{-1} \in \mathcal{G}$; then $G = \langle \mathcal{G}, \gamma \rangle$. We start using the filtration of \mathcal{G} introduced in Section 12. We define $M = \mathcal{G}/\mathcal{G}_2$ and $M_l = \mathcal{G}_l/\mathcal{G}_{l+1}$, $l \in \mathbb{N}$. For $l, k_1, \dots, k_l \in \mathbb{N}$ we define $M_{(k_1, \dots, k_l)} = \mathcal{G}_{(k_1, \dots, k_l)} \bmod \mathcal{G}_{l+1}$, and for $k, l \in \mathbb{N}$ with $k \geq l$ we put $M_{k,l} = \mathcal{G}_{k,l} \bmod \mathcal{G}_{l+1}$. We then have $M_{k,l} = \langle M_{(k_1, \dots, k_l)} : k_1 + \dots + k_l = k \rangle$.

Since $A\mathcal{G}_2 \subseteq \mathcal{G}_2$, the mapping $A: M \rightarrow M$ is defined. For $b \in M$ and $w = \begin{pmatrix} v_1 \\ \vdots \\ v_s \end{pmatrix} \in R^r = (\mathbb{R}^s)^r$, we define $b^w = b^{v_1}(Ab)^{v_2} \dots (A^{s-1}b)^{v_s} \in M^r$. If W is a subspace of R^r , we define $M^{[W]} = \langle b^w : b \in M, w \in W \rangle \subseteq M^r$.

Now let $l \in \mathbb{N}$ and $b_1, \dots, b_l \in M$; we then have a mapping $(R^r)^l \rightarrow M_l^r$, $(w_1, \dots, w_l) \mapsto [b_1^{w_1}, \dots, b_l^{w_l}]$. Since this mapping is multilinear, it extends to a homomorphism $(R^{\otimes l})^r \rightarrow M_l^r$, $u \mapsto [b_1, \dots, b_l]^u$, with $[b_1, \dots, b_l]^{w_1 \otimes \dots \otimes w_l} = [b_1^{w_1}, \dots, b_l^{w_l}]$. For $U \subseteq \mathbb{R}^{\otimes l}$ we define $M^{[U]} = \langle [b_1, \dots, b_l]^u : b_1, \dots, b_l \in M, u \in U \rangle$.

Clearly, for any $l, k_1, \dots, k_l \in \mathbb{N}$, $M^{[R_{(k_1, \dots, k_l)}^r]} = M_{(k_1, \dots, k_l)}^r$, and for any $k, l \in \mathbb{N}$, $k \geq l$, $M^{[R_{k,l}^r]} = M_{k,l}^r$.

13.5. We now return to our nilsystem (X, a) and a system $P = \{p_1, \dots, p_r\}$ of polynomials $\mathbb{Z}^m \rightarrow \mathbb{Z}$ with zero constant term. Let $H \subseteq \mathcal{G}^r$ be the identity component of $\pi^{-1}(\overline{\text{Orb}_P(\Delta_{X^r})})$; our goal is to find an estimate of the complexity of P , that is, an integer d such that $H \supseteq G_{d+1}^r$.

13.6. By Theorem 9.4, $H \bmod \mathcal{G}_2^r = M^{[W]}$, where $W \subseteq R^r$ is given by formula (9.3). (Notice that W satisfies the assumptions of items (ii) and (iii) of Lemma 13.3.) Our further argument will be based on the following proposition:

Proposition. *For any $l \in \mathbb{N}$, $(H \cap \mathcal{G}_l^r) \bmod \mathcal{G}_{l+1}^r \supseteq M^{[W^{\otimes l}]}$.*

Proof. For any $b_1, \dots, b_l \in \mathcal{G}$ and any $w_1, \dots, w_l \in W$, $b_i^{w_i} \in H \bmod \mathcal{G}_2^r$. Hence, $[b_1^{w_1}, \dots, b_l^{w_l}] \in (H \cap \mathcal{G}_l^r) \bmod \mathcal{G}_{l+1}^r$. ■

13.7. For each $l = 1, \dots, s$ let d_l be the minimal integer such that $R_{d_l, l}^r \subseteq W^{\otimes l}$. Then for each l ,

$$\mathcal{G}_{d_l, l}^r \bmod \mathcal{G}_{l+1}^r = M_{d_l, l}^r = M^{[R_{d_l, l}^r]} \subseteq M^{[W^{\otimes l}]} \subseteq H \bmod \mathcal{G}_{l+1}^r,$$

and so

$$\mathcal{G}_{d_l, l}^r \subseteq H \bmod \mathcal{G}_{l+1}^r. \quad (13.1)$$

Note that $d_1 = c_W + 1$, where c_W is the W -complexity of P . We will assume that $d_1 \geq 2$. (If $d_1 = 1$ then $c_W = 0$, the polynomials from P are linearly independent, and the complexity of P is equal to 0 (or is ≤ 1 if we take into account disconnected nilsystems).)

13.8. Lemma. *For any l , $d_l \leq ld_1$.*

Proof. If $k_1 \geq d_1$ then $R_{k_1}^r \subseteq W$, thus by Lemma 13.3(ii), $R_{(k_1, k_2, \dots, k_l)}^r \subseteq W^{\otimes l}$ for any k_2, \dots, k_l . If $k_1 + \dots + k_l = k \geq ld_1$, then $k_i \geq d_1$ for some i , and thus $R_{(k_1, k_2, \dots, k_l)}^r \subseteq W^{\otimes l}$ by Lemma 13.3 (i). So, $R_{ld_1, l}^r \subseteq W^{\otimes l}$. ■

13.9. Lemma. *Assume that $\mathcal{G}_{k, l}^r \subseteq H \bmod \mathcal{G}_{j, l+1}^r$ for some $k, j, l \in \mathbb{N}$. Then*

- (i) $[\mathcal{G}_{k, l}^r, \Delta_{\mathcal{G}^r}] \subseteq H \bmod \mathcal{G}_{j+1, l+2}^r$;
- (ii) $A^{\times r} \mathcal{G}_{k, l}^r \subseteq H \bmod \mathcal{G}_{j+1, l+1}^r$.

Proof. (i) By the second formula in (12.1),

$$[\mathcal{G}_{k, l}^r, \Delta_{\mathcal{G}^r}] \subseteq [HG_{j, l+1}^r, \Delta_{\mathcal{G}^r}] \subseteq [H, \Delta_{\mathcal{G}^r}][H, \Delta_{\mathcal{G}^r}, \mathcal{G}_{j, l+1}^r][\mathcal{G}_{j, l+1}^r, \Delta_{\mathcal{G}^r}].$$

By Proposition 11.5, $[H, \Delta_{\mathcal{G}^r}] \subseteq H$, and by Lemma 12.2(ii), $[H, \Delta_{\mathcal{G}^r}, \mathcal{G}_{j, l+1}^r], [\mathcal{G}_{j, l+1}^r, \Delta_{\mathcal{G}^r}] \subseteq \mathcal{G}_{j+1, l+2}^r$. Hence, $[\mathcal{G}_{k, l}^r, \Delta_{\mathcal{G}^r}] \subseteq H \bmod \mathcal{G}_{j+1, l+2}^r$.

(ii) By the second formula in (12.1),

$$A^{\times r} \mathcal{G}_{k, l}^r \subseteq A^{\times r} (HG_{j, l+1}^r) \subseteq A^{\times r} H[A^{\times r} H, \mathcal{G}_{j, l+1}^r] A^{\times r} \mathcal{G}_{j, l+1}^r.$$

By Proposition 11.5, $A^{\times r} H \subseteq H$, and by Lemma 12.2(iii), $[A^{\times r} H, \mathcal{G}_{j, l+1}^r], A^{\times r} \mathcal{G}_{j, l+1}^r \subseteq \mathcal{G}_{j+1, l+1}^r$. Hence, $A^{\times r} \mathcal{G}_{k, l}^r \subseteq H \bmod \mathcal{G}_{j+1, l+1}^r$. ■

13.10. As a corollary of Lemma 13.9 and Corollary 12.6, we get:

Proposition. *Assume that for certain $k, l, j \in \mathbb{N}$, $l \geq 2$, one has $\mathcal{G}_{k, l-1}^r \subseteq H \bmod \mathcal{G}_{j, l}^r$ and $\mathcal{G}_{k, l}^r \subseteq H \bmod \mathcal{G}_{j, l+1}^r$. Then $\mathcal{G}_{k+1, l}^r \subseteq H \bmod \mathcal{G}_{j+1, l+1}^r$.*

13.11. For each $l = 1, \dots, s$ put $k_l = \max\{d_l, d_{l-1} + 1, \dots, d_1 + l - 1\}$.

Proposition. For any l and any $j \in \mathbb{N}$, $\mathcal{G}_{k_l+j,l}^r \subseteq H \bmod \mathcal{G}_{j+l+1,l+1}^r$.

Proof. We will use induction on l . For $l = 1$ we have $k_1 = d_1$, $\mathcal{G}_{d_1,1}^r \subseteq H \bmod \mathcal{G}_{2,2}^r$ by (13.1), and thus for any $j \in \mathbb{N}$,

$$\mathcal{G}_{k_1+j,1}^r = (A^{\times r})^j \mathcal{G}_{k_1,1}^r \bmod \mathcal{G}_{k_1+j,2}^r \subseteq H \bmod \mathcal{G}_{2+j,2}^r$$

by Lemma 13.9(ii).

Assume by induction that for certain $l \geq 2$ one has

$$\mathcal{G}_{k_{l-1}+j,l-1}^r \subseteq H \bmod \mathcal{G}_{l+j,l}^r \quad (13.2)$$

for all $j \in \mathbb{N}$. We have $k_l = \max\{d_l, k_{l-1} + 1\}$. Thus, by (13.1), $\mathcal{G}_{k_l,l}^r \subseteq H \bmod \mathcal{G}_{l+1,l+1}^r$, and since $\mathcal{G}_{k_l,l-1}^r \subseteq \mathcal{G}_{k_{l-1}+1,l-1}^r \subseteq H \bmod \mathcal{G}_{l+1,l}^r$ by (13.2), we get

$$\mathcal{G}_{k_l+1,l}^r \subseteq H \bmod \mathcal{G}_{l+2,l+1}^r \quad (13.3)$$

by Proposition 13.10. Next, since (13.3) holds and since $\mathcal{G}_{k_l+1,l-1}^r \subseteq \mathcal{G}_{k_{l-1}+2,l-1}^r \subseteq H \bmod \mathcal{G}_{l+2,l}^r$ by (13.2), we get

$$\mathcal{G}_{k_l+2,l}^r \subseteq H \bmod \mathcal{G}_{l+3,l+1}^r$$

by Proposition 13.10. Using induction on j , we obtain $\mathcal{G}_{k_l+j,l}^r \subseteq H \bmod \mathcal{G}_{l+j+1,l+1}^r$ for all $j \in \mathbb{N}$. ■

13.12. Now let c be the C-complexity of P , then $\mathcal{G}_{c+1}^r \subseteq H$ by Lemma 11.8. By (13.1),

$$\mathcal{G}_{d_c,c}^r \subseteq H \mathcal{G}_{c+1}^r = H. \quad (13.4)$$

By Proposition 13.11 and by (13.4),

$$\mathcal{G}_{k_{c-1}+d_c-c,c-1}^r \subseteq H \mathcal{G}_{c+d_c-c,c}^r = H \mathcal{G}_{d_c,c}^r = H. \quad (13.5)$$

By Proposition 13.11 and by (13.5),

$$\mathcal{G}_{k_{c-2}+k_{c-1}+d_c-c-(c-1),c-2}^r \subseteq H \mathcal{G}_{c-1+k_{c-1}+d_c-c-(c-1),c-1}^r = H \mathcal{G}_{k_{c-1}+d_c-c,c-1}^r = H.$$

After repeating this $c - 1$ times, we get $\mathcal{G}_d^r = \mathcal{G}_{d,1}^r \subseteq H$ for $d = k_1 + \dots + k_{c-1} + d_c - c - (c - 1) - \dots - 2$. Hence, the complexity of P does not exceed $d - 1$, and we obtain:

13.13. Theorem. Let c be the C-complexity of P , let W be the subspace of R^s defined by formula (9.3), and for each $l = 1, \dots, s$ let d_l be the minimal integer such that $R_{d_l,l}^r \subseteq W^{\otimes l}$ and $k_l = \max\{d_l, d_{l-1} + 1, \dots, d_1 + l - 1\}$. The complexity of P does not exceed $k_1 + \dots + k_{c-1} + d_c - c(c + 1)/2$.

13.14. Using Lemma 13.8 and the fact that $d_1 = c_w + 1$, we get

$W^{\otimes 2} \supseteq R_{(2,1)}^5 \bmod R_{4,2}^5$, so $W^{\otimes 2} \supseteq R_{(2,1)}^5 \bmod R_{4,2}^5$, and by Lemma 13.3(iii), that $W^{\otimes 2} \supseteq R_{3,2}^5$. Hence, $d_2 = 3$. By Theorem 13.13, the complexity of P_1 is $\leq d_1 + d_2 - 3 = 3$. (A more careful analysis, based on the method from subsection 11.10, shows that P_1 has complexity 2.)

The system P_2 has W-complexity 3 with

$$W \bmod R_3^5 = \text{span}_{\mathbb{R}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 4 & 0 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 9 & 0 & 0 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 8 & 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & 0 & 27 & 0 & 0 & 0 & 9 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

Again, one checks that $W^{\otimes 2} \not\supseteq R_{2,1}^5$ and that $W^{\otimes 2} \supseteq R_{3,2}^5$, from which $d_2 = 3$ and the complexity of P_2 is ≤ 4 . (Using the method from subsection 11.10 one can see that P_2 has complexity 3.)

Bibliography

- [B] V. Bergelson, The multifarious Poincaré recurrence theorem, *Descriptive set theory and dynamical systems*, 31–57, London Math. Soc. Lecture Note Ser. **277**, Cambridge Univ. Press, Cambridge, 2000.
- [BHK] V. Bergelson, B. Host and B. Kra, Multiple recurrence and nilsequences. With an appendix by I. Ruzsa, *Inventiones Mathematicae* **160** (2005), 261–303.
- [BLLe1] V. Bergelson, A. Leibman and E. Lesigne, Complexities of finite families of polynomials, Weyl systems, and constructions in combinatorial number theory, *J. D’Analyse Math.* **103** (2007), 47–92.
- [BLLe2] V. Bergelson, A. Leibman and E. Lesigne, Intersective polynomials and the polynomial Szemerédi theorem, *Advances in Math.* **219** (2008), no. 1, 369–388.
- [F] N. Frantzikinakis, Multiple ergodic averages for three polynomials and applications, *Trans. of AMS* **360** (2008), no. 10, 5435–5475.
- [FK] N. Frantzikinakis and B. Kra, Polynomial averages converge to the product of integrals, *Israel J. of Math.* **148** (2005), 267–276.
- [Fu] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. d’Analyse Math.* **31** (1977), 204–256.
- [GW] T. Gowers and J. Wolf, The true complexity of a system of linear equations, preprint.
- [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, Polynomial sequences in groups, *J. of Algebra* **201** (1998), 189-206 (erratum at <http://www.math.ohio-state.edu/~leibman/preprints>).
- [L2] A. Leibman, Polynomial mappings of groups, *Israel J. of Math.* **129** (2002), 29-60 (erratum at <http://www.math.ohio-state.edu/~leibman/preprints>).
- [L3] A. Leibman, Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold, *Ergodic Theory and Dynam. Syst.* **25** (2005), 201-213.
- [L4] A. Leibman, Pointwise convergence of ergodic averages for polynomial actions of \mathbb{Z}^d by translations on a nilmanifold, *Ergodic Theory and Dynam. Syst.* **25** (2005), 215-225.
- [L5] A. Leibman, Convergence of multiple ergodic averages along polynomials of several variables, *Israel J. of Math.* **146** (2005), 303-315.
- [L6] A. Leibman, Host-Kra and Ziegler factors, and convergence of multiple averages (appendix to V. Bergelson, Combinatorial and Diophantine Applications of Ergodic Theory), Handbook of Dynamical Systems vol. 1B, B. Hasselblatt and A. Katok, eds., Elsevier, 2005, 841-853.
- [L7] A. Leibman, Rational sub-nilmanifolds of a compact nilmanifold, *Ergodic Theory and Dynam. Syst.* **26** (2006), 787-798.
- [L8] A. Leibman, Orbits on a nilmanifold under the action of a polynomial sequences of translations, *Ergodic Theory and Dynam. Syst.* **27** (2007), 1239-1252.
- [Le] E. Lesigne, Sur une nil-variété, les parties minimales associées à une translation sont uniquement ergodiques, *Ergodic Theory and Dynam. Syst.* **11** (1991), 379-391.
- [Sh] N. Shah, Invariant measures and orbit closures on homogeneous spaces for actions of subgroups generated by unipotent elements, *Lie groups and ergodic theory (Mumbai, 1996)*, 229-271, Tata Institute of Fundamental Research, Bombay, 1998.
- [M] A. Malcev, On a class of homogeneous spaces, *Amer. Math. Soc. Transl.* **9** (1962), 276-307.
- [Z1] T. Ziegler, A non-conventional ergodic theorem for a nilsystem, *Ergodic Theory and Dynam. Syst.* **25** (2005), 1357-1370.
- [Z2] T. Ziegler, Universal characteristic factors and Furstenberg averages, *J. Amer. Math. Soc.* **20** (2007), no. 1, 53-97.