$\mathbf{IP}^{*}_{\mathbf{r}}$-recurrence and nilsystems

V. Bergelson and A. Leibman

V. Bergelson, Ohio State University, Department of Mathematics, OH 43210, USA; e-mail: bergelson.1@osu.edu
A. Leibman, Ohio State University, Department of Mathematics, OH 43210, USA; e-mail: leibman.1@osu.edu

Abstract

By a result due to Furstenberg, a homeomorphism T of a compact space is distal if and only if it possesses the property of IP*-recurrence, meaning that for any $x_0 \in X$, for any open neighborhood U of x_0, and for any sequence (n_i) in $\mathbb{Z}$, the set $R_U(x_0) = \{n \in \mathbb{Z} : T^n x_0 \in U\}$ has a non-trivial intersection with the set of finite sums $\{n_{i_1} + n_{i_2} + \cdots + n_{i_s} : i_1 < i_2 < \ldots < i_s,\ s \in \mathbb{N}\}$. We show that translations on compact nilmanifolds (which are known to be distal) are characterized by a stronger property of IP$^*_r$-recurrence, which asserts that for any $x_0 \in X$ and any neighborhood U of x_0 there exists $r \in \mathbb{N}$ such that for any r-element sequence $n_1, \ldots, n_r$ in $\mathbb{Z}$ the set $R_U(x_0)$ has a non-trivial intersection with the set $\{n_{i_1} + n_{i_2} + \cdots + n_{i_s} : i_1 < i_2 < \ldots < i_s,\ s \le r\}$. We also show that the property of IP$^*_r$-recurrence is equivalent to an ostensibly much stronger property of polynomial IP$^*_r$-recurrence. (This should be juxtaposed with the fact that for general distal transformations the polynomial IP*-recurrence is strictly stronger than the IP*-recurrence.)

0. Introduction

Let (X, T) be a topological dynamical system, meaning that X is a compact metric space and T is a self-homeomorphism of X. Given a point $x_0 \in X$ and an open neighborhood U of x_0, define $R_U(x_0) = \bigl\{n \in \mathbb{Z} : T^n x_0 \in U\bigr\}$, *the set of returns* of x_0 into U. Sets of returns reflect the properties of topological system, and it is of interest to characterize (and/or distinguish between) dynamical systems by arithmetic properties of these sets. An example of this kind is provided by a theorem of Furstenberg on sets of returns in distal systems. A system (X, T) is said to be *distal* if for any distinct $x, y \in X$, $\inf_{n\in\mathbb{Z}} \mathrm{dist}(T^n x, T^n y) > 0$. Given a sequence $n_1, n_2, \ldots$ in $\mathbb{Z}$, the set $\bigl\{n_{i_1} + \cdots + n_{i_s} : s \in \mathbb{N},\ i_1 < \cdots < i_s\bigr\}$ of finite sums of distinct elements of this sequence is called an *IP-set*. A subset E of $\mathbb{Z}$ is called an *IP*-set* if it intersects every IP-set$^{(1)}$. Furstenberg's theorem says that distal systems are characterized by *the IP*-recurrence property*:

Theorem 0.1. ([F], Theorem 9.11) *A system (X, T) is distal if and only if for any $x_0 \in X$ and any open neighborhood U of x_0 the set of returns $R_U(x_0)$ is an IP*-set.*

Another relevant example involves translations on compact abelian groups. *A set of differences* is a set of the form $\bigl\{n_i - n_j,\ j < i\bigr\}$, where (n_i) is an infinite sequence in $\mathbb{Z}$; a subset E of $\mathbb{Z}$ is said to be *a Δ^*-set* if it has a nonempty intersection with every set of differences in $\mathbb{Z}$. A point x in a system (X, T) is said to be *almost automorphic* if for any sequence (n_i) in $\mathbb{Z}$, $T^{n_i} x \longrightarrow y$ implies $T^{-n_i} y \longrightarrow x$. It is shown in [F], Theorem 9.13, that a system has the Δ^*-recurrence property (that is, that every set of returns in the system is a Δ^*-set) if and only if every point in the system is almost automorphic. Next, by a theorem of Veech (see [V], Theorem 1.2; see also [AGN]) every point of a minimal$^{(2)}$ system (X, T) is almost automorphic if and only if the family $\{T^n,\ n \in \mathbb{Z}\}$ of powers of T is equicontinuous. Now, it is not hard to see that for a minimal T the family $\{T^n,\ n \in \mathbb{Z}\}$ is equicontinuous if and only if (X, T) is isomorphic to a translation on a compact abelian group$^{(3)}$. Thus, the recurrence property characterizing minimal group translations is that

---

Partially supported by NSF grants DMS-1162073 and DMS-1500575.

(1) IP*-sets can be defined, in total analogy, in any general abelian group G. They possess the nice properties of "largeness" and "regularity": every IP*-set $E \subseteq G$ is syndetic (which means that G is covered by finitely many shifts of E), and the intersection of any two IP*-sets in G is an IP*-set as well.

(2) A system (X, T) is *minimal* if it has no proper closed subsystems, or, equivalently, if the orbit of every point of X is dense in X.

(3) The "only if" implication follows from the fact that for any $x_0 \in X$ one can define an additive group structure on the orbit $\{T^n x_0,\ n \in \mathbb{Z}\}$ by $T^n x_0 + T^m x_0 = T^{n+m} x_0$, $n, m \in \mathbb{Z}$, and then extend it, with the

of Δ^*.

Our goal in this paper is to provide a characterization, in terms of recurrence properties, of *nilsystems*, namely, systems of the form (X,T) where X is a *nilmanifold* (a compact homogeneous space of a nilpotent Lie group G) and T is a *niltranslation* (a translation on X defined by an element of G). The motivation for this study comes from the fact that nilsystems are intrinsically related to various problems arising in ergodic theory of multiple recurrence, combinatorics, and number theory, and better understanding of the recurrence properties of niltranslations may lead to interesting applications in these areas. (See, for example, [HK], [Z], [BLLe], [GT], [L].) It is well known that nilsystems are distal (see [AGH], [Ke1], [Ke2]), and thus are IP*-recurrent; however, not every minimal distal system is a nilsystem or a pre-nilsystem$^{(4)}$, and thus there must be a stronger than IP* property of recurrence that characterizes nilsystems.

For an integer $r \in \mathbb{N}$ and an r-element sequence $n_1, \ldots, n_r$ in $\mathbb{Z}$, we call the set $\big\{n_{i_1} + \cdots + n_{i_s} : 1 \leq s \leq r,\ i_1 < \cdots < i_s\big\}$ of sums of distinct elements of this sequence an *IP$_r$-set*. A set $E \subseteq \mathbb{Z}$ is called an *IP$^*_r$-set* if it has a nonempty intersection with every IP$_r$-set in $\mathbb{Z}$. We say that a set is an *IP$^*_0$-set* if it is an IP$^*_r$-set for some $r \in \mathbb{N}$. IP$^*_0$-sets form a proper subfamily of the family of IP*-sets: clearly, every IP$^*_0$-set is IP*, but not vice versa$^{(5)}$.

A special class of nilsystems is provided by affine skew product transformations of tori$^{(6)}$; it follows from [B], Theorem 7.7, that every such system has the IP$^*_0$-recurrence property: for every $x_0 \in \mathbb{T}^k$ and any open neighborhood U of x_0 the set of returns $R_U(x_0)$ is an IP$^*_0$-set. On the other hand, one can show that not every minimal distal system is IP$^*_0$-recurrent (see [BL3], Section 1). It is tempting to conjecture that it is the IP$^*_0$-recurrence property that characterizes the nilsystems. This, however, cannot be exactly so: any recurrence property must be stable under passing to inverse limits whereas inverse limits of nilsystems do not have to be nilsystems. Let us define *a pre-nilsystem* as the inverse limit of a sequence of nilsystems. (Notice that, in contrast with the definition of the so-called *pro-nilsystems*, in the definition of pre-nilsystems we don't require the nilpotency class of the nilsystems in the sequence to be bounded.) The following result provides a characterization of pre-nilsystems in terms of IP$^*_0$-recurrence:

Theorem 0.2. *Any pre-nilsystem (and so, any nilsystem) is IP$^*_0$-recurrent. Any IP$^*_0$-recurrent system is a disjoint union of pre-nilsystems. In particular, a minimal system is IP$^*_0$-recurrent iff it is a pre-nilsystem.*

Remarks 0.3. (i) The IP$^*_0$-sets also appear in [HSY], Theorem 8.17, where they are used for characterization of the so-called "∞-step AA" systems.
(ii) In analogy with IP$^*_0$-sets, one can define Δ^*_0-sets as those having a nonempty intersection with every large enough finite set of differences. In contrast with IP*/IP$^*_0$-recurrence, the classes of minimal Δ^*- and Δ^*_0-recurrent systems coincide. Indeed, as it was remarked above, any minimal Δ^*-recurrent system is a translation on a compact abelian group, and it is easy to see that any such translation is Δ^*_0-recurrent as well.
(iii) A different, and somewhat more technical, but similar in spirit characterization of pre-nilsystems via "absence of pairs with arbitrarily long finite IP-independence sets" was obtained in [DDMSY] (where pre-nilsystems are called *infinite-step nilsystems*).

The second statement of Theorem 0.2 is an easy corollary of the results from [HKM]. To prove the first statement, we use a coordinate approach. On any nilmanifold X one has natural coordinates such that under the action of a niltranslation T the sequence of coordinates of the image $T^n x_0$ of any point $x_0 \in X$ is given by *generalized polynomials* (see [BL2], Theorem A). We therefore need to deal with images of IP-sets under generalized polynomial mappings; these images form a subclass of *generalized polynomial*

---

help of equicontinuity, to all of X. This makes X a compact abelian group on which T acts as a minimal translation. (For details see, for example, [Ku], Theorem 2.42.)

(4) See, for example, [Ko], Theorem 7, [Iw], Corollary 3, or [DDMSY], Section 5.2.

(5) To see this, it is enough to exhibit an IP$^*_0$-set S which is not an IP*-set. One can take, for example $S = \bigcup_{r=1}^{\infty} S_r$, where $S_r = \{2^{2^r}, 2 \cdot 2^{2^r}, 3 \cdot 2^{2^r}, \ldots, r \cdot 2^{2^r}\}$, $r \in \mathbb{N}$. Since for each r, S_r is a dilation of the set $\{1, 2, ..., r\}$, S contains arbitrarily large IP$_r$-sets, but it contains no IP-sets since the distances between consecutive elements of S form a non-decreasing sequence which tends to infinity.

(6) An affine skew product transformation of the k-dimensional torus $\mathbb{T}^k = \mathbb{R}^k/\mathbb{Z}^k$ is defined by the formula $T(x_1, \ldots, x_k) = (x_1 + \alpha_1, x_2 + a_{2,1}x_1 + \alpha_2, \ldots, x_k + a_{k,k-1}x_{k-1} + \cdots + a_{k,1}x_1 + \alpha_k)$ with $\alpha_i \in \mathbb{T}$ and $a_{i,j} \in \mathbb{Z}$.

IP-sets. Conventional IP- and IP$_r$-sets in $\mathbb{Z}$ can be viewed as the images of mappings $\varphi\colon \mathcal{F}(A) \longrightarrow \mathbb{Z}$ from the semigroup $\mathcal{F}(A)$ of finite subsets of A, for $A = \mathbb{N}$ and, respectively, for $A = \{1, \ldots, r\}$, defined by $\varphi(\alpha) = \sum_{i\in\alpha} a_i$. Such a mapping φ is "linear" in the following sense: $\varphi(\alpha \cup \beta) = \varphi(\alpha) + \varphi(\beta)$ whenever $\alpha, \beta \in \mathcal{F}(A)$ are disjoint. Let H be an additive abelian group; one can introduce the notion of polynomial mappings $\mathcal{F}(A) \longrightarrow H$ as follows. For a mapping $\varphi\colon \mathcal{F}(A) \longrightarrow H$ and a set $\beta \in \mathcal{F}(A)$ let *the β-derivative* $D_\beta\varphi$ be the mapping $\mathcal{F}(A \setminus \beta) \longrightarrow H$ defined by $D_\beta\varphi(\alpha) = \varphi(\alpha + \beta) - \varphi(\alpha)$. Then we say that a mapping $\varphi\colon \mathcal{F}(\{1, \ldots, r\}) \longrightarrow H$ is *polynomial of degree* $\leq d$ if for any disjoint $\beta_0, \beta_1, \ldots, \beta_d \in \mathcal{F}(\{1, \ldots, r\})$, $D_{\beta_0} D_{\beta_1} \cdots D_{\beta_d}\varphi = 0$. (See [BL1], Section 8.1.) Examples of quadratic (that is, of degree ≤ 2) polynomial mappings are, in increasing generality, $\varphi(\alpha) = \bigl(\sum_{i\in\alpha} a_i\bigr)^2$, $\varphi(\alpha) = \bigl(\sum_{i\in\alpha} a_i\bigr)\bigl(\sum_{i\in\alpha} b_i\bigr) = \sum_{i,j\in\alpha} a_i b_j$, and $\varphi(\alpha) = \sum_{i,j\in\alpha} c_{i,j}$, where $a_i, b_j, c_{i,j} \in H$. *Generalized polynomial mappings* are the mappings built from (conventional) polynomial mappings using the operations of addition, multiplication, and taking the integer part. (An example is $\varphi = \bigl[[\varphi_1]\varphi_2 + \varphi_3\bigr]\varphi_4 + [\varphi_5][\varphi_6]\varphi_7$, which is comprised of the polynomial mappings $\varphi_1, \ldots, \varphi_7$.) Let us say that a generalized polynomial mapping φ has *total degree* $\leq D$ if the sum $\sum_i \deg\varphi_i$ of the degrees of all the "conventional" polynomial mappings φ_i of which φ is comprised does not exceed D, and let us say that a generalized polynomial mapping is *constant free* if all the φ_i vanish at $\emptyset$: $\varphi_i(\emptyset) = 0$. Let us also say that a generalized polynomial mapping is *open* if it is contained in the ideal generated by the conventional constant-free polynomial mappings of the ring of constant-free generalized polynomial mappings. (In other words, a generalized polynomial mapping is open if it contains no "closed" summands of the form $[\varphi_1]\cdots[\varphi_k]$, where φ_i are generalized polynomial mappings.) For $x \in \mathbb{R}$ let $\|x\| = \operatorname{dist}(x, \mathbb{Z})$. The following result, which will be used in the proof of Theorem 0.2, is of independent interest:

Theorem 0.4. (Cf. Theorem 1.11 below.) *For any $D \in \mathbb{N}$ and $\varepsilon > 0$ there exists $r = r(D, \varepsilon) \in \mathbb{N}$ such that for any open constant-free generalized polynomial mapping $\varphi\colon \mathcal{F}(\{1, \ldots, r\}) \longrightarrow \mathbb{R}$ of total degree $\leq D$ there exists a nonempty $\alpha \subseteq \{1, \ldots, r\}$ for which $\|\varphi(\alpha)\| < \varepsilon$.*

The *VIP-sets* in $\mathbb{Z}^l$ are defined as the images $\bigl\{\varphi(\alpha) : \alpha \in \mathcal{F}(\mathbb{N}),\ \alpha \neq \emptyset\bigr\}$ of polynomial mappings $\varphi\colon \mathcal{F}(\mathbb{N}) \longrightarrow \mathbb{Z}^l$ with $\varphi(\emptyset) = 0$, and we say that a set $E \subseteq \mathbb{Z}^l$ is *a VIP*-set* if E has a nonempty intersection with every VIP-set in $\mathbb{Z}^l$. Similarly, for all $d, r \in \mathbb{N}$, we define *VIP$_{d,r}$-sets* as the images $\bigl\{\varphi(\alpha) : \alpha \subseteq \{1, \ldots, r\},\ \alpha \neq \emptyset\bigr\}$ of polynomial mappings $\varphi\colon \mathcal{F}(\{1, \ldots, r\}) \longrightarrow \mathbb{Z}^l$ of degree $\leq d$ and with $\varphi(\emptyset) = 0$, and say that a set $E \subseteq \mathbb{Z}^l$ is a *VIP$^*_{d,r}$-set* if it has a nonempty intersection with every VIP$_{d,r}$-set. We will also say that a set $E \subseteq \mathbb{Z}^l$ is a *VIP$^*_0$-set* if for any $d \in \mathbb{N}$, E is an VIP$^*_{d,r}$-set for some $r \in \mathbb{N}$.

We are now going to formulate a corollary of Theorem 0.4 that provides a strong enhancement and generalization of some classical Diophantine results. *Generalized polynomials* $\mathbb{Z}^l \longrightarrow \mathbb{R}$ are functions obtained from conventional polynomials $\mathbb{Z}^l \longrightarrow \mathbb{R}$ using the operations of addition, multiplication, and taking the integer part. (Note that unlike generalized polynomial mappings, whose domain is $\mathcal{F}(A)$, the generalized polynomials are defined on $\mathbb{Z}^l$.) In analogy with generalized polynomial mappings, we say that a generalized polynomial is *constant-free* if it is comprised of conventional polynomials with zero constant term. The total degree of a generalized polynomial is also defined as the sum of the degrees of the polynomials it is comprised of.

Since the composition $\psi\circ\varphi$ of a polynomial mapping $\varphi\colon \mathcal{F}(A) \longrightarrow \mathbb{Z}^l$ of dergree $\leq d$ and a generalized polynomial $\psi\colon \mathbb{Z}^l \longrightarrow \mathbb{R}$ of total degree $\leq D$ is a generalized polynomial mapping of total degree $\leq dD$, Theorem 0.4 now implies the following result:

Theorem 0.5. (Cf. [F], Theorem 2.19, and [B], Theorem 7.7) *For any $D, d \in \mathbb{N}$ and $\varepsilon > 0$ there exists $r = r(D, d, \varepsilon) \in \mathbb{N}$ such that for any $l \in \mathbb{N}$ and any open constant-free generalized polynomial $\psi\colon \mathbb{Z}^l \longrightarrow \mathbb{R}$ of total degree $\leq D$ the set $\bigl\{n \in \mathbb{Z}^l : \|\psi(n)\| < \varepsilon\bigr\}$ is a VIP$^*_{d,r}$-set.*

Let G be a nilpotent Lie group; an *l-parameter polynomial sequence* in G is a mapping $g\colon \mathbb{Z}^l \longrightarrow G$ of the form $T_1^{p_1(n)} \cdots T_b^{p_b(n)}$, $n \in \mathbb{Z}^l$, where $T_i \in G$ and p_i are polynomials $\mathbb{Z}^l \longrightarrow \mathbb{Z}$; the *naive degree* of g is defined as $\max_i \deg p_i$.(7) Using the fact that the coordinates of a point of a nilmanifold under the action of a polynomial sequence of niltranslations are generalized polynomials, we obtain as a corollary of Theorem 0.5

---

(7) A more fundamental notion of *degree* of a polynomial sequence in a nilpotent group can be defined as the number of "differentiations" which it takes in order to reduce the polynomial sequence to a constant. For our purposes, however, the "naive" degree is quite sufficent.

the following strengthening of the first part of Theorem 0.2:

Theorem 0.6. (Cf. Theorem 1.12 below.) *Let X be a nilmanifold with metric ρ (compatible with the homogeneous space structure on X). For any $a,d\in\mathbb{N}$ and $\varepsilon>0$ there exists $r=r(a,d,\varepsilon)\in\mathbb{N}$ such that for any $x_0\in X$, any $l\in\mathbb{N}$, and any l-parameter polynomial sequence g of niltranslations on X of naive degree $\leq a$ and with $g(0)=\mathrm{Id}_X$, the set $\bigl\{n\in\mathbb{Z}^l:\rho(g(n)x_0,x_0)<\varepsilon\bigr\}$ is a VIP$^*_{d,r}$-set.*

A shifted VIP$^*_{d,r}$-set, that is, a set of the form $E+m$ where $E\subseteq\mathbb{Z}^l$ is a VIP$^*_{d,r}$-set and $m\in\mathbb{Z}^l$, is called an *VIP$^*_{d,r,+}$-set*. From Theorem 0.9 we get the following corollary:

Corollary 0.7. *Let X be a nilmanifold with metric ρ. For any $a,d\in\mathbb{N}$ and $\varepsilon>0$ there exists $r=r(a,d,\varepsilon)\in\mathbb{N}$ such that for any $x_0\in X$, any $l\in\mathbb{N}$, any l-parameter polynomial sequence g of niltranslations on X of naive degree $\leq a$, and any point y in the closure of the orbit $\{g(n)x_0\}_{n\in\mathbb{Z}^l}$ of x_0 under g, the set $\bigl\{n\in\mathbb{Z}^l:\rho(g(n)x_0,y)<\varepsilon\bigr\}$ is a VIP$^*_{d,r,+}$-set.*

(Indeed, choose $n_0\in\mathbb{Z}^l$ such that the point $y_0=g(n_0)x_0$ satisfies $\rho(y_0,y)<\varepsilon/2$, and let $h(n)=g(n+n_0)g(n_0)^{-1}$, $n\in\mathbb{Z}^l$. Then $h(n)y_0=g(n+n_0)x_0$, and for any n such that $\rho(h(n)y_0,y_0)<\varepsilon/2$ we have $\rho(g(n+n_0)x_0,y)<\varepsilon$. So, the set $\bigl\{n\in\mathbb{Z}^l:\rho(g(n)x_0,y)<\varepsilon\bigr\}-n_0$ contains the set $\bigl\{n\in\mathbb{Z}^l:\rho(h(n)y_0,y_0)<\varepsilon\bigr\}$, which is a VIP$^*_{d,r}$-set by Theorem 0.6.)

Remark 0.8. Corollary 0.7 can be viewed as a generalization and strengthening of a classical theorem of Weyl ([W]) which says that if x_0 is a point of the torus $\mathbb{T}^b=\mathbb{R}^b/\mathbb{Z}^b$, p is a polynomial $\mathbb{Z}^l\longrightarrow\mathbb{T}^b$, and y is a point in the closure in $\mathbb{T}^b$ of the orbit $\{x_0+p(n)\}_{n\in\mathbb{Z}^l}$ of x_0 under the shifts by the values of p, then for any $\varepsilon>0$ the set $\bigl\{n\in\mathbb{Z}^l:\rho(x_0+p(n),y)<\varepsilon\bigr\}$ is syndetic.

We say that a dynamical system (X,T) is *VIP*-recurrent* if for any $x_0\in X$ and any open neighborhood U of x_0 the set of returns $R_U(x_0)=\bigl\{n\in\mathbb{Z}:T^nx_0\in U\bigr\}$ is a VIP*-set, and is *VIP$^*_0$-recurrent* if for any $x_0\in X$ and any open neighborhood U of x_0 the set $R_U(x_0)$ is a VIP$^*_0$-set. The VIP*-recurrence property turns out to be strictly stronger than that of the IP*-recurrence: there exist distal but not VIP*-recurrent systems.(8) As for the VIP$^*_0$-recurrence, we get, as a corollary of Theorem 0.6, that, via Theorem 0.2, VIP$^*_0$-recurrence is equivalent to IP$^*_0$-recurrence:

Theorem 0.9. *Any pre-nilsystem is VIP$^*_0$-recurrent, and any VIP$^*_0$-recurrent system is a disjoint union of pre-nilsystems.*

In Section 1 of the paper we prove (a more precise version of) Theorems 0.4 and 0.5 and deduce Theorem 0.6 from them. In Section 2 we obtain the second statement of Theorem 0.2.

1. Sets of visits of open bounded generalized polynomials with no constant term to a neighborhood of zero

Let A be a set and $(H,+)$ be an abelian group. For $r\in\mathbb{N}$ we will denote by $[1,r]$ the interval $\{1,\ldots,r\}$ in $\mathbb{N}$. We denote by $\mathcal{F}(A)$ the set of finite subsets of A, by $A^{(d)}$, $d\in\mathbb{N}$, the set of subsets of A of cardinality d, and by $A^{(\leq d)}$, $d\in\mathbb{N}$, the set of nonempty subsets of A of cardinality $\leq d$, $A^{(\leq d)}=\bigcup_{l=1}^d A^{(l)}$.

We start with discussing polynomial mappings on $\mathcal{F}(A)$. We say that a mapping $\varphi\colon\mathcal{F}(A)\longrightarrow H$ is *linear* if it satisfies the identity $\varphi(\alpha\cup\beta)=\varphi(\alpha)+\varphi(\beta)$ whenever $\alpha,\beta\in\mathcal{F}(A)$ are disjoint, and will denote the set of linear mappings $\mathcal{F}(A)\longrightarrow H$ by $\mathrm{Lin}(A,H)$. A mapping $\varphi\in\mathrm{Lin}(A,H)$ is uniquely defined by its values at singletons: for any $\alpha\in\mathcal{F}(A)$, $\varphi_\alpha=\sum_{a\in\alpha}\widehat{\varphi}(\{a\})$. We will call the mapping $\widehat{\varphi}\colon A\longrightarrow H$ defined by $\widehat{\varphi}(a)=\varphi(\{a\})$ *the producing function* for φ; we then have $\varphi(\alpha)=\sum_{a\in\alpha}\widehat{\varphi}(a)$, $\alpha\in\mathcal{F}(A)$.

For a mapping $\varphi\colon\mathcal{F}(A)\longrightarrow H$ and $\beta\in\mathcal{F}(A)$ we define *the β-derivative* $D_\beta\varphi$ of φ by $D_\beta\varphi(\alpha)=\varphi(\alpha\cup\beta)-\varphi(\alpha)$, $\alpha\in\mathcal{F}(A\setminus\beta)$. We say that a mapping φ is *polynomial of degree* $\leq d$ if for any $d+1$ pairwise disjoint sets $\beta_0,\ldots,\beta_d\in\mathcal{F}(A)$ one has $D_{\beta_d}\cdots D_{\beta_0}\varphi=0$.

We will denote by $\mathrm{Pol}_d(A,H)$ the group of polynomial mappings $\mathcal{F}(A)\longrightarrow H$ of degree $\leq d$. We will mainly deal with polynomial mappings "having zero constant term"; let us denote by $\mathrm{Pol}^0_d(A,H)$ the

---

(8) See [P], Corollary 5.1, where it is shown that for any nonlinear polynomial $p\colon\mathbb{Z}\longrightarrow\mathbb{Z}$ there exists an affine skew product transformation T such that $\mathrm{liminf}_n\mathrm{dist}(T^{p(n)}0,0)>0$.

subgroup $\big\{\varphi\in\mathrm{Pol}_d(A,H):\varphi(\emptyset)=0\big\}$ of $\mathrm{Pol}_d(A,H)$. Notice that $\mathrm{Lin}(A,H)=\mathrm{Pol}_1^0(A,H)$.

One can show (see [BL1], sections 8.3-8.5) that any polynomial mapping $\varphi\in\mathrm{Pol}_d^0(A,H)$ can be represented in the form $\varphi(\alpha)=\Phi(\alpha^d)$, $\alpha\in\mathcal{F}(A)$, for some mapping $\Phi\in\mathrm{Lin}(A^d,H)$, so that

$$\varphi(\alpha)=\sum_{v\in\alpha^d}\widehat{\Phi}(v),\ \ \alpha\in\mathcal{F}(A),$$

where $\widehat{\Phi}\colon A^d\longrightarrow H$ is the producing function for Φ. We will call $\widehat{\Phi}$ *a q-producing function for* φ.

The q-producing function for a polynomial mapping $\varphi\in\mathrm{Pol}_d^0(A,H)$ is not canonically defined. A more natural is *the t-producing function* for φ, a function $\widetilde{\Phi}\colon A^{(\leq d)}\longrightarrow H$ such that for any $\alpha\in\mathcal{F}(A)$,

$$\varphi(\alpha)=\sum_{u\in\alpha^{(\leq d)}}\widetilde{\Phi}(u).$$

The t-producing function $\widetilde{\Phi}$ for φ is defined uniquely (and provides a natural approach to the definition of polynomial mappings in the case H is a commutative semigroup). In terms of $\widetilde{\Phi}$, φ is the sum of its *homogeneous components*, $\varphi=\varphi_1+\cdots+\varphi_d$, where for each i, $\varphi_i(\alpha)=\sum_{\delta\in\alpha^{(d)}}\widetilde{\Phi}(\delta)$. To obtain the t-producing function $\widetilde{\Phi}$ for φ from a q-producing function $\widehat{\Phi}$ one simply sums up the values of $\widehat{\Phi}$ at the elements of $A^{(d)}$ corresponding to the same element of $A^{(\leq d)}$: for any $u\in A^{(\leq d)}$,

$$\widetilde{\Phi}(u)=\sum_{\substack{v=(a_1,\ldots,a_d)\in\alpha^d\\ \{a_1,\ldots,a_d\}=u}}\widehat{\Phi}(v).\tag{1.1}$$

Let B be a collection of pairwise disjoint finite subsets of A; we will call B *a disjoint subcollection* in A; if $|B|=s$ we will say that B is *a disjoint s-subcollection*. Given a disjoint subcollection B in A, we have an injection $\mathcal{F}(B)\longrightarrow\mathcal{F}(A)$ defined by $\gamma\mapsto\bigcup\gamma$, and we will identify $\mathcal{F}(B)$ with its image in $\mathcal{F}(A)$. Given a polynomial mapping $\varphi\colon\mathcal{F}(A)\longrightarrow H$, we call the polynomial mapping $\varphi|_{\mathcal{F}(B)}$ *a subpolynomial* of φ corresponding to the disjoint subcollection B and denote it by $\varphi_{\downarrow B}$. Any disjoint subcollection B of a disjoint subcollection in A induces the disjoint subcollection $B'=\big\{\bigcup C:C\in B\big\}$ in A; abusing notation, we will denote the subpolynomial $\varphi_{\downarrow B'}$ of φ by $\varphi_{\downarrow B}$.

Let $\widehat{\Phi}\colon A^d\longrightarrow H$ be a q-producing function for a polynomial mapping $\varphi\colon\mathcal{F}(A)\longrightarrow H$ of degree $\leq d$ and let $\Phi\in\mathrm{Lin}(A^d,H)$ be the linear mapping produced by $\widehat{\Phi}$. Given a disjoint s-subcollection $B=\{B_1,\ldots,B_s\}$ in A, one finds a q-producing function for the subpolynomial $\varphi_{\downarrow B}$ as follows. For any $\beta\subseteq B$ we have

$$\begin{aligned}\varphi_{\downarrow B}(\beta)&=\varphi\Big(\bigcup_{C\in\beta}C\Big)=\Phi\Big(\Big(\bigcup_{C\in\beta}C\Big)^d\Big)=\sum_{v\in(\bigcup_{C\in\beta}C)^d}\widehat{\Phi}(v)\\&=\sum_{C_1,\ldots,C_d\in\beta}\ \sum_{v\in C_1\times\cdots\times C_d}\widehat{\Phi}(v)=\sum_{(C_1,\ldots,C_d)\in\beta^d}\Phi(C_1,\ldots,C_d);\end{aligned}\tag{1.2}$$

thus, the mapping $\Phi|_{B^d}$ is a q-producing function for $\varphi_{\downarrow B}$.

The following proposition establishes the IP_r^*-recurrence property of polynomial mappings with values in the torus $\mathbb{T}=\mathbb{R}/\mathbb{Z}$.

Proposition 1.1. (Cf. [B], Theorem 7.7) *For any $k,d\in\mathbb{N}$ and $\varepsilon>0$ there exists $r=r(k,d,\varepsilon)\in\mathbb{N}$ such that for any $\varphi_1,\ldots,\varphi_k\in\mathrm{Pol}_d^0([1,r],\mathbb{T})$ there exists a nonempty $\alpha\in\mathcal{F}([1,r])$ such that $\mathrm{dist}(\varphi_i(\alpha),0)<\varepsilon$ for all $i\in\{1,\ldots,k\}$ (where "dist" is the distance on $\mathbb{T}$).*

Proof. Put $c = \lceil 1/\varepsilon \rceil$ and partition the torus $\mathbb{T}$ into c intervals of length $\le 1/\varepsilon$. By the Polynomial Hales-Jewett theorem (see [BL1], Theorem 0.10), there exists $r \in \mathbb{N}$ such that for any partition of $\mathcal{F}([1,r]^d \times [k])$ into c subsets there exist $\gamma \subset [1,r]^d \times [k]$ and a nonempty $\alpha \subseteq [1,r]$ such that $\gamma \cap (\alpha^d \times [k]) = \emptyset$ and the sets $\gamma, \gamma \cup (\alpha^d \times \{1\}), \ldots, \gamma \cup (\alpha^d \times \{k\})$ belong to the same element of the partition. Let $\varphi_1, \ldots, \varphi_k \in \mathrm{Pol}^0_d([1,r],\mathbb{T})$. For each i let $\widehat{\Phi}_i\colon [1,r]^d \longrightarrow \mathbb{T}$ be a q-producing function for φ_i. Define a mapping $\widehat{\Phi}\colon [1,r]^d \times [k] \longrightarrow \mathbb{T}^k$ by $\widehat{\Phi}(v,i) = \widehat{\Phi}_i(v)$, $v \in [1,r]^d$, $i \in [k]$, and let $\Phi \in \mathrm{Lin}([1,r]^d \times [k], \mathbb{T})$ be the linear mapping produced by $\widehat{\Phi}$. Then, via Φ, the partition of $\mathbb{T}$ defines a partition of $\mathcal{F}([1,r]^d \times [k])$ into c subsets. Applying the Polynomial Hales-Jewett theorem, we can find $\gamma \subset [1,r]^d \times [k]$ and a nonempty $\alpha \subseteq [1,r]$ such that $\gamma \cap (\alpha^d \times [k]) = \emptyset$ and the sets $\gamma, \gamma \cup (\alpha^d \times \{1\}), \ldots, \gamma \cup (\alpha^d \times \{k\})$ belong to the same element of the partition; then for any i, $\Phi(\gamma)$ and $\Phi(\gamma \cup (\alpha^d \times \{i\}))$ belong to the same partition of $\mathbb{T}$, and so, $\mathrm{dist}\big(\Phi(\gamma), \Phi(\gamma \cup (\alpha^d \times \{i\}))\big) < \varepsilon$. Since $\Phi(\gamma \cup (\alpha^d \times \{i\})) = \Phi(\gamma) + \Phi(\alpha^d \times \{i\}) = \Phi(\gamma) + \varphi_i(\alpha)$, this implies that $\mathrm{dist}(0, \varphi_i(\alpha)) < \varepsilon$. ∎

Recall that by $\|x\|$ we denote the distance from $x \in \mathbb{R}$ to $\mathbb{Z}$. We may then reformulate Proposition 1.1 as follows:

Corollary 1.2. *For any $k, d \in \mathbb{N}$ and $\varepsilon > 0$ there exists $r = r(k,d,\varepsilon) \in \mathbb{N}$ such that for any $\varphi_1, \ldots, \varphi_k \in \mathrm{Pol}^0_d([1,r],\mathbb{R})$ there exists a nonempty $\alpha \in \mathcal{F}([1,r])$ such that $\|\varphi_i(\alpha)\| < \varepsilon$ for all $i \in \{1, \ldots, k\}$.*

Next we show that if r is large enough, any polynomial mapping $\varphi \in \mathrm{Pol}^0_d([1,r],\mathbb{T})$ has a subpolynomial whose q-producing function is arbitrarily small:

Proposition 1.3. *For any $d, s \in \mathbb{N}$ and $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that for any $\varphi \in \mathrm{Pol}^0_d([1,r],\mathbb{T})$ there exists a disjoint s-subcollection B in $[1,r]$ such that a q-producing function $\widehat{\Phi}_B$ for $\varphi_{\downarrow_B}$ satisfies $\mathrm{dist}(\widehat{\Phi}_B, 0) < \varepsilon$.*

Proof. Take $r_0 = r(s^d, d, \varepsilon)$ as in Corollary 1.2, and put $A = [1,s] \times [1,r_0]$ (and $r = |A| = sr_0$). Let $\varphi \in \mathrm{Pol}^0_d([1,r],\mathbb{T})$. Let $\widehat{\Phi}\colon [1,r] \longrightarrow \mathbb{T}$ be a q-producing function for φ and let $\Phi \in \mathrm{Lin}([1,r]^d, \mathbb{T})$ be the linear mapping produced by $\widehat{\Phi}$. For each $i = (i_1, \ldots, i_d) \in [1,s]^d$ define a polynomial mapping $\varphi_I \in \mathrm{Pol}^0_d([1,r_0],\mathbb{T})$ by $\varphi_I(\alpha) = \Phi((\{i_1\} \times \alpha) \times \cdots \times (\{i_d \times \alpha)\})$. By Corollary 1.2 there exists $\alpha \subseteq [1,r_0]$ such that $\mathrm{dist}(\varphi_I(\alpha), 0) < \varepsilon$ for all $I \in [1,s]^d$. Take the disjoint s-subcollection $B = \big\{\{i\} \times \alpha : i \in [1,s]\big\}$ in A. By the choice of α, for any $w \in B^d$ we have $\mathrm{dist}(\Phi(w), 0) < \varepsilon$. Since, by (1.2), $\Phi|_{B^d}$ is a q-producing function for $\varphi_{\downarrow_B}$, we are done. ∎

Replacing in Proposition 1.3 ε by ε/s^d, we obtain:

Corollary 1.4. *For any $d, s \in \mathbb{N}$ and $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that for any $\varphi \in \mathrm{Pol}^0_d([1,r],\mathbb{T})$ there exists a disjoint s-subcollection B in $[1,r]$ such that $\mathrm{dist}(\varphi_{\downarrow_B}, 0) < \varepsilon$.*

By formula (1.1), any value of the t-producing function for $\varphi \in \mathrm{Pol}^0_d(A,\mathbb{R})$ is a sum of less than d^d values of the q-producing function for φ. Hence, Proposition 1.3 implies the following corollary:

Proposition 1.5. *For any $d, s \in \mathbb{N}$ and $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that for any $\varphi \in \mathrm{Pol}^0_d([1,r],\mathbb{T})$ there exists a disjoint s-subcollection B in $[1,r]$ such that the t-producing function $\widetilde{\Phi}_B$ for $\varphi_{\downarrow_B}$ satisfies $\mathrm{dist}(\widetilde{\Phi}_B, 0) < \varepsilon$.*

In terms of polynomial mappings with values in $\mathbb{R}$, Proposition 1.5 takes the following form:

Corollary 1.6. *For any $d, s \in \mathbb{N}$ and $\varepsilon > 0$ there exists $r \in \mathbb{N}$ such that for any $\varphi \in \mathrm{Pol}^0_d([1,r],\mathbb{R})$ there exists a disjoint s-subcollection B in $[1,r]$ such that the t-producing function $\widetilde{\Phi}_B$ for $\varphi_{\downarrow_B}$ satisfies $\|\widetilde{\Phi}_B\| < \varepsilon$.*

Now let $\varphi \in \mathrm{Pol}^0_d([1,r],\mathbb{R})$ be a polynomial mapping whose t-producing function $\widetilde{\Phi}$ satisfies $\|\widetilde{\Phi}\| < 1/r^d$. We will denote by $[x]$ the integer and by $\{x\}$ the fractional parts of $x \in \mathbb{R}$. If $x \in \mathbb{R}$ satisfies $\|x\| < \varepsilon$, then either $\{x\} < \varepsilon$ or $\{x\} > 1 - \varepsilon$. If $x_1, \ldots, x_n \in \mathbb{R}$ satisfy $\{x_i\} < 1/n$, $i = 1, \ldots, n$, then $\big[\sum_{i=1}^n x_i\big] = \sum_{i=1}^n [x_i]$. Thus, if $\widetilde{\Phi}$ satisfies $\{\widetilde{\Phi}\} < 1/r^d$, then for any $\alpha \subseteq [1,r]$,

$$[\varphi(\alpha)] = \Big[\sum_{u \in \alpha^{(\le d)}} \widetilde{\Phi}(u)\Big] = \sum_{u \in \alpha^{(\le d)}} [\widetilde{\Phi}(u)]$$

and so, $[\varphi]$ is also a polynomial mapping, $[\varphi] \in \mathrm{Pol}^0_d([1,r],\mathbb{Z})$, with the t-producing function $[\widetilde{\Phi}]$.

For any $x \in \mathbb{R} \setminus \mathbb{Z}$, $[x] = -[-x] - 1$ and $\{-x\} = 1 - \{x\}$, so, if $x_1, \dots, x_n \in \mathbb{R}$ satisfy $\{x_i\} > 1 - 1/n$, $i = 1, \dots, n$, then

$$\left[\textstyle\sum_{i=1}^n x_i\right] = -\left[-\textstyle\sum_{i=1}^n x_i\right] - 1 = -\left[\textstyle\sum_{i=1}^n (-x_i)\right] - 1 = -\textstyle\sum_{i=1}^n [-x_i] - 1 = \textstyle\sum_{i=1}^n (-[-x_i]) - 1.$$

Applying this to $\widetilde{\Phi}$, we see that if $\widetilde{\Phi}$ satisfies $\{\widetilde{\Phi}\} > 1 - 1/r^d$, then for any $\alpha \subseteq [1, r]$,

$$[\varphi(\alpha)] = \Big[\sum_{u \in \alpha^{(\le d)}} \widetilde{\Phi}(u)\Big] = \sum_{u \in \alpha^{(\le d)}} \Big(-[-\widetilde{\Phi}(u)]\Big) - 1.$$

So, $[\varphi] + 1$ is a polynomial mapping, $[\varphi] + 1 \in \mathrm{Pol}_d^0([1, r], \mathbb{Z})$, with the t-producing function $-[-\widetilde{\Phi}]$.

In the general case, when $\|\widetilde{\Phi}\| < 1/r^d$, we may have neither $\{\widetilde{\Phi}\} < 1/r^d$ nor $\{\widetilde{\Phi}\} > 1 - 1/r^d$. However, if φ is a homogeneous polynomial of degree $l \le d$ (which means that $\varphi(\alpha) = \sum_{u \in \alpha^{(l)}} \widetilde{\Phi}(u)$), then, given $s \in \mathbb{N}$, if r is large enough, by the classical Ramsey theorem we can choose an s-element subset B of $[1, r]$ such that either $\{\widetilde{\Phi}(u)\} < 1/r^d$ for all $u \in B^{(d)}$ or $\{\widetilde{\Phi}(u)\} > 1 - 1/r^d$ for all $u \in B^{(d)}$. Identifying B with the "singleton disjoint '-subcollection' $\{\{b\} : b \in B\}$ in $[1, r]$, we will therefore have $[\varphi_{\downarrow B}] \in \mathrm{Pol}_d^0(B, \mathbb{Z}) + e$ with $e \in \{0, -1\}$.

For a general $\varphi \in \mathrm{Pol}_d^0([1, r], \mathbb{R})$, applying this argument to all homogeneous components of φ and using a diagonal process, we arrive at the following lemma:

Lemma 1.7. *For any $d, s \in \mathbb{N}$ there exists $r \in \mathbb{N}$ such that for any $\varphi \in \mathrm{Pol}_d^0([1, r], \mathbb{R})$ whose t-producing function $\widetilde{\Phi}$ satisfies $\|\widetilde{\Phi}\| < 1/r^d$ there exists a (singleton) disjoint subcollection B in $[1, r]$ such that $[\varphi_{\downarrow B}] \in \mathrm{Pol}_d^0(B, \mathbb{Z}) + e$ with $e \in \{0, -1, \dots, -d\}$.*

Combining Lemma 1.7 with Corollary 1.6 we obtain:

Theorem 1.8. *For any $d, s \in \mathbb{N}$ there exists $r \in \mathbb{N}$ such that for any $\varphi \in \mathrm{Pol}_d^0([1, r], \mathbb{R})$ there exists a disjoint s-subcollection B in $[1, r]$ such that $[\varphi] \in \mathrm{Pol}_d^0(B, \mathbb{Z}) + e$ with $e \in \{0, -1, \dots, -d\}$.*

Using induction on k, one can extend Theorem 1.8 to the case of k polynomials:

Theorem 1.9. *For any $k, d_1, \dots, d_k, s \in \mathbb{N}$ there exists $r = r(k, (d_1, \dots, d_k), s) \in \mathbb{N}$ such that for any $\varphi_i \in \mathrm{Pol}_{d_i}^0([1, r], \mathbb{R})$, $i = 1, \dots, k$, there exists a disjoint s-subcollection B in $[1, r]$ such that for every $i \in \{1, \dots, k\}$, $[\varphi_i] \in \mathrm{Pol}_{d_i}^0(B, \mathbb{Z}) + e_i$, with $e_i \in \{0, -1, \dots, -d\}$.*

A generalized polynomial is a function obtained from conventional polynomials using the operations of taking the integer part, addition, and multiplication. We say that a generalized polynomial ψ is *constant free* if all polynomials involved in the expression of ψ have zero constant term. (More precisely, a generalized polynomial is constant free if it has *a representation* in which all polynomials have zero constant term. A similar convention applies to all the definitions below.) We say that a polynomial ψ is *open* if it is contained in the ideal, in the ring of constant free generalized polynomials, generated by the ordinary polynomials. This is equivalent to saying that ψ (or rather a representation of ψ) has no summand that is a product of "closed" generalized polynomials $[\psi_i]$. Any open constant-free generalized polynomial is representable in the form

$$\psi = \sum_{j=1}^m [\psi_{j,1}] \cdots [\psi_{j,l_j}] \psi_{j,0} \tag{1.3}$$

where for every j, $\psi_{j,1}, \dots, \psi_{j,l_j}$ are open constant-free generalized polynomials and $\psi_{j,0}$ are conventional polynomials with zero constant term.

We now introduce the notions of height, width, and degree for (a representation of) a generalized polynomial ψ:

The height $h(\psi)$ of ψ is the maximum length of sequences of nested brackets in ψ: we put $h(\psi) = 0$ if ψ is a conventional polynomial and we say that $h(\psi) \le h$ if ψ has a representation (1.3) where for all j and all $t \ge 1$, $h(\psi_{j,t}) \le h - 1$.

The width $w(\psi)$ is the maximum number of components in ψ itself and in all its components: we put

$w(\psi)=1$ if ψ is a conventional polynomial and we say that $w(\psi)\leq w$ if ψ has a representation (1.3) where $w(\psi_{j,t})\leq w$ for all j and all $t\geq 1$ and also $\sum_{j=1}^{m}(l_j+1)\leq w$.
The degree $d(\psi)$ of ψ is defined as usual under the assumption that $\deg[\psi]=\deg\psi$: we say that $d(\psi)\leq d$ if ψ has a representation (1.3) with $\max_{j=1}^{m}\bigl(\sum_{t=0}^{l_j}\deg\psi_{j,t}\bigr)\leq d$.
(For example, for $\psi(x)=[[x^2+1]x][x^3+2x]x+[x^2](x+1)+x^3$ we have $h(\psi)=2$, $w(\psi)=6$, and $d(\psi)=7$.)

We extend the above definitions to generalized polynomial mappings with domain $\mathcal{F}(A)$, and will denote by $\mathrm{GPol}^0_{d,h,w}(A,H)$ the set (the algebra) of open constant-free generalized polynomial mappings $\varphi\colon\mathcal{F}(A)\longrightarrow H$, where $H=\mathbb{R}$ or $\mathbb{Z}$, with $d(\varphi)\leq d$, $h(\varphi)\leq h$, and $w(\varphi)\leq w$. Given $\varphi\in\mathrm{GPol}^0_{d,h,w}(A,H)$ and a disjoint subcollection B in A, we define the generalized polynomial mapping $\varphi_{\downarrow B}\in\mathrm{GPol}^0_{d,h,w}(B,H)$ as the restriction of φ to the set $\mathcal{F}(B)$ considered as a subset of $\mathcal{F}(A)$.

The following theorem says that generalized polynomial mappings turn into ordinary polynomial mappings after being restricted to a suitable disjoint subcollection in their domain:

Theorem 1.10. *For any* $k,d_1,\ldots,d_k,h,w,s\in\mathbb{N}$ *there exists* $r=r(k,(d_1,\ldots,d_k),h,w,s)\in\mathbb{N}$ *such that for any* $\varphi_i\in\mathrm{GPol}^0_{d_i,h,w}([1,r],\mathbb{R})$, $i=1,\ldots,k$, *there exists a disjoint* s*-subcollection* B *in* $[1,r]$ *such that* $\varphi_{i\downarrow B}\in\mathrm{Pol}^0_{d_i}(B,\mathbb{R})$, $i=1,\ldots,k$.

Proof. We will use induction on h; when $h=0$ the statement is trivial. Take r_0 to be the maximum of the integers $r(l,(b_1,\ldots,b_l),s)$ in Theorem 1.9 over all integers $l\leq kw$ and all l-tuples $(b_1,\ldots,b_l)$ of nonnegative integers with $\sum_{j=1}^{l}b_j\leq w\sum_{i=1}^{k}d_i$. By induction on h, let r be the maximum of the integers $r(l,(d_1,\ldots,d_l),h-1,w,r_0)$ in the assertion of Theorem 1.10 over all integers $l\leq kw$ and all l-tuples $(b_1,\ldots,b_l)$ of nonnegative integers with $\sum_{j=1}^{l}b_j\leq w\sum_{i=1}^{k}d_i$. Let $\varphi_i\in\mathrm{GPol}^0_{d_i,h,w}([1,r],\mathbb{R})$, $i=1,\ldots,k$. For each i reprsent φ_i in the form

$$\varphi_i=\sum_{j=1}^{m_i}[\varphi_{i,j,1}]\cdots[\varphi_{i,j,l_{i,j}}]\varphi_{i,j,0},$$

where for every i,j we have $\varphi_{i,j,0}\in\mathrm{Pol}^0_{d_{i,j,0}}([1,r],\mathbb{R})$ and for every $t\geq 1$ we have $\varphi_{i,j,t}\in\mathrm{GPol}^0_{d_{i,j,t},h-1,w}([1,r],\mathbb{R})$ with

$$\textstyle\sum_{j=1}^{m_i}(l_{i,j}+1)\leq w \text{ for all } i \text{ and } \sum_{t=0}^{l_{i,j}}d_{i,j,t}\leq d_i \text{ for all } i,j,$$

so that

$$\textstyle\sum_{i=1}^{k}\sum_{j=1}^{m_i}(l_{i,j}+1)\leq kw \text{ and } \sum_{i=1}^{k}\sum_{j=1}^{m_i}\sum_{t=0}^{l_{i,j}}d_{i,j,t}\leq w\sum_{i=1}^{k}d_i.$$

By the choice of r there exists a disjoint r_0-subcollection $B_0\subset\mathcal{F}([1,r])$ such that $\varphi_{i,j,t\downarrow B_0}\in\mathrm{Pol}^0_{d_{i,j,t}}(B_0,\mathbb{R})$ for all i,j,t. Then by the choice of r_0 there exists a disjoint s-subcollection B in B_0 such that for all i,j,t, $\bigl[\varphi_{i,j,t\downarrow B}\bigr]\in\mathrm{Pol}^0_{d_{i,j,t}}(B,\mathbb{Z})$. Hence for every i,

$$\varphi_{i\downarrow B}=\sum_{j=1}^{m_i}\bigl[\varphi_{i,j,1\downarrow B}\bigr]\cdots\bigl[\varphi_{i,j,l_{i,j}\downarrow B}\bigr]\varphi_{i,j\downarrow B}\in\mathrm{Pol}^0_{d_i}(B,\mathbb{R}).$$

■

Combining Theorem 1.10 and Corollary 1.2, we obtain:

Theorem 1.11. *For any* $k,d,h,w\in\mathbb{N}$ *there exists* $r=r(k,d,h,w)\in\mathbb{N}$ *such that for any* $\varphi_1,\ldots,\varphi_k\in\mathrm{GPol}^0_{d,h,w}([1,r],\mathbb{R})$ *there exists a nonempty* $\alpha\in\mathcal{F}([1,r])$ *such that* $\|\varphi_i(\alpha)\|<\varepsilon$, $i=1,\ldots,k$.

Let $X = G/\Gamma$ be a k-dimensional compact nilmanifold; we may and will assume that X is connected. (Any nilmanifold is a subnilmanifold of a connected one.) Let ρ be a metric on X (induced by a metric on G compatible with the Lie group structure thereon). Fix a point $x_0 \in X$, and let $\tau = (\tau_1, \ldots, \tau_k): X \longrightarrow [0,1)^k$ be Maltsev's coordinates on X centered at x_0. The inverse mapping τ^{-1} is continuous, and the distance $\rho(x, x_0)$ from $x \in X$ to x_0 is continuous with respect to the distance from $\tau(x)$ to the set of vertices $\{0,1\}^k$ of the cube $[0,1]^k$. (See, for example, [BL2], Section 1.5.)

Let g be an *(l-parameter) polynomial sequence* in G, that is, a mapping $g: \mathbb{Z}^l \longrightarrow G$ of the form $g(n) = T_1^{p_1(n)} \cdots T_b^{p_b(n)}$, $n \in \mathbb{Z}^l$, where $T_1, \ldots, T_b \in G$, $p_1, \ldots, p_b$ are polynomials $\mathbb{Z}^l \longrightarrow \mathbb{Z}$; we define $\text{n-deg}\, g$, *the naive degree of* g, as $\max_{i=1}^b \deg p_i$. Then for each $i = 1, \ldots, k$, the sequences $\psi_i(n) = \tau_i(g(n)x_0)$, $n \in \mathbb{Z}^c$, of coordinates of x_0 under the action of g are open $[0,1)$-valued generalized polynomials, with parameters depending only on X and $\text{n-deg}\, g$ (see [BL2], Theorem A and Theorem A**), and if $g(0) = 1_G$, these polynomials can be assumed to be constant-free. For any polynomial mapping $\varphi \in \mathrm{Pol}_d([1,r]), \mathbb{Z}^c)$, the composition mappings $\psi_i \circ \varphi: \mathcal{F}([1,r]) \longrightarrow [0,1)$, $i = 1, \ldots, k$, are open constant-free generalized polynomial mappings, with parameters only depending on X, d, and $\text{n-deg}\, g$. From Theorem 1.11 we now obtain the following result:

Theorem 1.12. *Let $X = G/\Gamma$ be a nilmanifold with metric ρ. For any $a, d \in \mathbb{N}$ and $\varepsilon > 0$ there exists $r = r(a, d, \varepsilon) \in \mathbb{N}$ such that for any l, any l-parameter polynomial sequence g in G with* $\text{n-deg}\, g \leq a$ *and $g(0) = 1_G$, any $x_0 \in X$, and any $\varphi \in \mathrm{Pol}_d^0([1,r], \mathbb{Z}^l)$ there exists a nonempty $\alpha \in \mathcal{F}([1,r])$ such that $\rho\big(g(\varphi(\alpha))x_0, x_0\big) < \varepsilon$.*

Remark 1.13. Theorem 1.12 easily extends to *generalized polynomial sequences* in nilpotent groups, that is, to sequences of the form $g(n) = T_1^{p_1(n)} \cdots T_b^{p_b(n)}$ where p_i are generalized polynomials $\mathbb{Z}^l \longrightarrow \mathbb{Z}$.

2. IP$_0^*$-recurrence implies approximability by nilsystems

In this section we prove the second statement of Theorem 0.2. Let (X, ρ) be a compact metric space, T be a self homeomorphism of X, and assume that (X,T) is IP$_0^*$-recurrent. Then, in particular, (X,T) is IP*-recurrent, so by Theorem 0.1, (X,T) is distal, and thus is a disjoint union of minimal subsystems (see [F], corollary to Theorem 8.7). Hence, we may assume that (X,T) is minimal.

Now, by the way of contradiction, assume that a minimal system (X,T) is not a pre-nilsystem, that is, not an inverse limit of nilsystems; our goal is to show that there exists a point $x \in X$ and $\varepsilon > 0$ such that for every $r \in \mathbb{N}$ there exists a linear mapping $\varphi \in \mathrm{Lin}([1,r], \mathbb{Z})$ such that $\rho(T^{\varphi(\alpha))}x, x) > \varepsilon$ for every nonempty $\alpha \subseteq [1,r]$.

We will use the following result ([HKM] Theorem 1.3 and Corollary 4.2): for any r, the maximal r-step pro-nilfactor of (X,T) is defined by a closed T-invariant equivalence relation $\mathbf{RP}^{[r]} \subseteq X^2$ (called *the regionally proximal relation of order* r), with $(x_0, y_0) \in \mathbf{RP}^{[r]}$ if and only if for any $\delta > 0$ there exists a point $x \in X$ and a mapping $\varphi \in \mathrm{Lin}([1,r], \mathbb{Z})$ such that

$$\rho(x, x_0) < \delta \text{ and } \rho(T^{\varphi(\alpha)}x, y_0) < \delta \text{ for all nonempty } \alpha \subseteq [1,r]. \tag{2.1}$$

Our assumption that (X,T) is not a pre-nilsystem is equivalent to the assumption that $\bigcap_{r=1}^\infty \mathbf{RP}^{[r]} \neq \Delta$, where Δ is the diagonal of X^2. Fix $(x_0, y_0) \in \bigcap_{r=1}^\infty \mathbf{RP}^{[r]}$ with $x_0 \neq y_0$. Let $\varepsilon = \inf_{n \in \mathbb{Z}} \rho(T^n x_0, T^n y_0)$; since (X,T) is distal, we have $\varepsilon > 0$. Since (X,T) is minimal, the orbit $\{T^n x_0\}_{n \in \mathbb{Z}}$ of x_0 is dense in X. Let $r \in \mathbb{N}$ and let $U \subseteq X$ be an open set. Choose $n \in \mathbb{Z}$ such that $T^n x_0 \in U$ and choose $\delta > 0$ such that $\rho(T^n x, T^n y) < \varepsilon/3$ whenever $\rho(x,y) < \delta$. Find $x \in X$ such that (2.1) holds and $T^n x \in U$. Then $\rho(T^n x, T^n x_0) < \varepsilon/3$ and $\rho(T^{\varphi(\alpha)}T^n x, T^n y_0) < \varepsilon/3$ for all nonempty $\alpha \subseteq [1,r]$, and since $\rho(T^n x_0, T^n y_0) \geq \varepsilon$, we have that $\rho(T^{\varphi(\alpha)}T^n x, T^n x) > \varepsilon/3$ for all nonempty $\alpha \subseteq [1,r]$. This proves that for any $r \in \mathbb{N}$ the open set

$$R_r = \big\{x \in X : \text{there exists } \varphi \in \mathrm{Lin}([1,r], \mathbb{Z}) \text{ such that } \rho(T^{\varphi(\alpha)}x, x) > \varepsilon/3 \text{ for all nonempty } \alpha \in [1,r]\big\}$$

is dense in X. By Baire category theorem $\bigcap_{r=1}^\infty R_r$ is nonempty, which gives us what we wanted – a point $x \in X$ such that for every $r \in \mathbb{N}$ there exists a mapping $\varphi \in \mathrm{Lin}([1,r], \mathbb{Z})$ such that $\rho(T^{\varphi(\alpha))}x, x) > \varepsilon/3$ for every nonempty $\alpha \subseteq [1,r]$.

Bibliography

[AGH] L. Auslander, L. Green and F. Hahn, *Flows on homogeneous spaces*, Annals of Math. Studies, vol. 53, Princeton Univ. Press, 1963.

[AGN] L. Auslander, G. Greschonig and A. Nagar, Reflections on equicontinuity, *Proc. of AMS* **142** (2014), no. 9, 3129-3137.

[B] V. Bergelson, Ultrafilters, IP sets, Dynamics, and Combinatorial Number Theory, *Contemp. Math.* **530** (2010), 23-47.

[BL1] V. Bergelson and A. Leibman, Set-polynomials and polynomial extension of the Hales-Jewett Theorem, *Ann. of Math.* **150** (1999), no. 1, 33-75.

[BL2] V. Bergelson and A. Leibman, Distribution of values of bounded generalized polynomials, *Acta Math.* **198** (2007), 155-230.

[BL3] V. Bergelson and A. Leibman, Sets of large values of correlation functions for polynomial cubic configurations, *Ergodic Th. and Dynam. Sys.* **38** (2018), no. 8, 499-522.

[BLLe] V. Bergelson, A. Leibman and E. Lesigne, Intersective polynomials and the polynomial Szemerédi theorem, *Advances in Math.* **219** (2008), no. 1, 369-388.

[F] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton Univ. Press, 1981.

[DDMSY] P. Dong, S. Donoso, A. Maas, S. Shao, and X. Ye, Infinite-step nilsystems, independence and complexity, *Ergodic Th. And Dynam. Sys.* **33** (2011), 118-143.

[GT] B. Green and T. Tao, The Möbius function is strongly orthogonal to nilsequences, *Ann. of Math.* **175** (2012), no. 2, 541-566.

[HK] B. Host and B. Kra, Non-conventional ergodic averages and nilmanifolds, *Ann. of Math.* **161** (2005), no. 1, 397-488.

[HKM] B. Host, B. Kra, and A. Maass, Nilsequences and a structure theorem for topological dynamical systems, *Adv. in Math.* **224** (2010), no. 1, 103-129.

[HSY] W. Huang, S. Shao, and X.D. Ye, Nil Bohr-sets and almost automorphy of higher order, *Mem. Amer. Math. Soc.* **241** (2016), no. 1143.

[Iw] A. Iwanik, Generic smooth cocycles of degree zero over irrational rotations, *Studia Mathematica* **115** (1995), no. 3, 241-250.

[Ke1] H. B. Keynes, Topological dynamics in coset transformation groups, *Bull. Amer. Math. Soc.* **72** (1966), 1033-1035.

[Ke2] H. B. Keynes, A study of the proximal relation in coset transformation groups, *Trans. Amer. Math. Soc.* **128** (1967), 389-402.

[Ko] K. Kodaka, Anzai and Furstenberg transformations in the 2-torus and topologically quasi-discrete spectrum, *Canad. Math. Bull.* **38** (1995), no. 1, 87-92.

[Ku] P. Kůrka, *Topological and Symbolic Dynamics*, Société Mathématique de France, 2003.

[L] A. Leibman, Nilsequences, null-sequences, and multiple correlation sequences, *Ergodic Th. and Dynam. Sys.* **35** (2015), no. 1, 176-191.

[P] R. Pavlov, Some counterexamples in topological dynamics, *Ergodic Th. and Dynam. Sys.* **28** (2008), no. 4, 1291-1322.

[V] W. A. Veech, The equicontinuous structure relation for minimal abelian transformation groups, *Amer. J. Math.* **90** (1968), 723-732.

[W] H. Weyl, Über die Gleichverteilung von Zahlen mod Eins, *Math. Ann.* **77** (1916), 313-352.

[Z] T. Ziegler, Universal characteristic factors and Furstenberg averages, *J. Amer. Math. Soc.* **20** (2007), no. 1, 53-97.