ON SEVERAL NOTIONS OF COMPLEXITY OF POLYNOMIAL PROGRESSIONS

BORYS KUCA

ABSTRACT. For a polynomial progression

$$(x, x + P_1(y), ..., x + P_t(y)),$$

we define four notions of complexity: Host-Kra complexity, Weyl complexity, true complexity and algebraic complexity. The first two describe the smallest characteristic factor of the progression, the third one refers to the smallest-degree Gowers norm controlling the progression, and the fourth one concerns algebraic relations between terms of the progressions. We conjecture that these four notions are equivalent, which would give a purely algebraic criterion for determining the smallest Host-Kra factor or the smallest Gowers norm controlling a given progression. We prove this conjecture for all progressions whose terms only satisfy homogeneous algebraic relations and linear combinations thereof. This family of polynomial progressions includes, but is not limited to, arithmetic progressions, progressions with linearly independent polynomials $P_1, ..., P_t$ and progressions whose terms satisfy no quadratic relations. For progressions that satisfy only linear relations, such as

$$(x, x + y^2, x + 2y^2, x + y^3, x + 2y^3),$$

we derive several combinatorial and dynamical corollaries: (1) an estimate for the count of such progressions in subsets of $\mathbb{Z}/N\mathbb{Z}$ or totally ergodic dynamical systems; (2) a lower bound for multiple recurrence; (3) and a popular common difference result in $\mathbb{Z}/N\mathbb{Z}$. Lastly, we show that Weyl complexity and algebraic complexity always agree, which gives a straightforward algebraic description of Weyl complexity.

1. INTRODUCTION

A polynomial $P \in \mathbb{R}[y]$ is *integral* if $P(\mathbb{Z}) \subset \mathbb{Z}$ and $P(0) = 0$. For $t \in \mathbb{N}_+$, an *integral polynomial progression* of length $t + 1$ is a tuple $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ given by

$$\vec{P}(x, y) = (x, \ x + P_1(y), \ ..., \ x + P_t(y))$$

for distinct integral polynomials $P_1, ..., P_t$. We moreover say that a set $A \subset \mathbb{N}$ contains $\vec{P}(x, y)$ for some $x, y \in \mathbb{N}$ if $\vec{P}(x, y) \in A^{t+1}$. A major result on integral polynomial progressions is the polynomial Szemerédi theorem by Bergelson and Leibman, which extends the famous theorem of Szemerédi on arithmetic progressions.

Theorem 1.1 (Polynomial Szemerédi theorem). [BL96] *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ be an integral polynomial progression, and suppose that $A \subseteq \mathbb{N}$ is dense[1]. Then A contains $\vec{P}(x, y)$ for some $x, y \in \mathbb{N}_+$.*

Theorem 1.1 can be deduced from the following ergodic theoretic statement using the Furstenberg correspondence principle.

---

[1] Meaning that $\limsup\limits_{N \to \infty} \frac{|A \cap [N]|}{N} > 0$, where $[N] = \{1, ..., N\}$.

Theorem 1.2. [BL96, HK05a] *Let* $(X, \mathcal{X}, \mu, T)$ *be an invertible measure-preserving dynamical system,* $t \in \mathbb{N}_+$ *and* $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ *be an integral polynomial progression. If* $\mu(A) > 0$ *for* $A \in \mathcal{X}$, *then*

$$(1) \qquad \lim_{N \to \infty} \mathbb{E}_{n \in [N]} \mu(A \cap T^{P_1(n)} A \cap ... \cap T^{P_t(n)} A) > 0,$$

where $[N] = \{1, ..., N\}$ *and* $\mathbb{E}_{x \in X} = \frac{1}{|X|} \sum_{x \in X}$ *for any set* X.

To prove Theorem 1.1, one thus needs to understand limits of multiple ergodic averages of the form

$$(2) \qquad \mathbb{E}_{n \in [N]} T^{P_1(n)} f_1 \cdots T^{P_t(n)} f_t$$

for $f_1, ..., f_t \in L^\infty(\mu)$. By a remarkable result of Host and Kra [HK05a, HK05b], there exists a family of factors[2] $(\mathcal{Z}_s)_{s \in \mathbb{N}}$, called henceforth *Host-Kra factors*, with the property that weak or L^2 limits of expressions of the form (2) remain unchanged if we project any of the functions f_i onto one of the factors $\mathcal{Z}_s$ for some s dependent on $\vec{P}$ and i.

Definition 1.3 (Characteristic factors). *Let* $(X, \mathcal{X}, \mu, T)$ *be an invertible measure-preserving dynamical system,* $t \in \mathbb{N}_+$ *and* $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ *be an integral polynomial progression.*

Suppose that $1 \leqslant i \leqslant t$. *A factor* $\mathcal{Y}$ *of* $\mathcal{X}$ *is* characteristic *for the* L^2-*convergence of* $\vec{P}$ *at* i *if for all choices of* $f_1, ..., f_t \in L^\infty(\mu)$, *the* L^2-*limit of* (2) *is 0 whenever* $\mathbb{E}(f_i | \mathcal{Y}) = 0$.

Similarly, suppose that $0 \leqslant i \leqslant t$. *A factor* $\mathcal{Y}$ *of* $\mathcal{X}$ *is* characteristic *for the weak convergence of* $\vec{P}$ *at* i *if for all choices of* $f_0, ..., f_t \in L^\infty(\mu)$, *the weak limit of* (2), *i.e. the expression*

$$(3) \qquad \lim_{N \to \infty} \mathbb{E}_{n \in [N]} \int_X f_0 \cdot T^{P_1(n)} f_1 \cdots T^{P_t(n)} f_t d\mu,$$

is 0 whenever $\mathbb{E}(f_i | \mathcal{Y}) = 0$.

Theorem 1.4 ([HK05a, Lei05a]). *Let* $t \in \mathbb{N}_+$. *For each integral polynomial progression* $\vec{P} \in \mathbb{R}[x, y]^{t+1}$, *there is* $s \in \mathbb{N}$ *such that for all invertible ergodic systems* $(X, \mathcal{X}, \mu, T)$, *the factor* $\mathcal{Z}_s$ *is characteristic for the* L^2 *convergence of* $\vec{P}$ *at* i *for all* $0 \leqslant i \leqslant t$.

The utility of Host-Kra factors comes from the fact that they are inverse limits of *nilsystems*, and so understanding (2) for arbitrary systems comes down to proving certain equidistribution results on spaces called nilmanifolds that possess rich algebraic structure. Importantly, $\mathcal{Z}_s$ is a factor of $\mathcal{Z}_{s+1}$ for each $s \in \mathbb{N}$, hence it is natural to inquire about the smallest value of s for which the factor $\mathcal{Z}_s$ is characteristic for $\vec{P}$ at i.

Definition 1.5 (Host-Kra complexity). *Let* $t \in \mathbb{N}_+$ *and* $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ *be an integral polynomial progression. Fix* $0 \leqslant i \leqslant t$. *The progression* $\vec{P}$ *has* Host-Kra complexity s *at* i, *denoted* $\mathcal{HK}_i(\vec{P})$, *if* s *is the smallest natural number such that the factor* $\mathcal{Z}_s$ *is characteristic for the weak convergence of* $\vec{P}$ *at* i *for all invertible totally ergodic dynamical systems* $(X, \mathcal{X}, \mu, T)$. *We say* $\vec{P}$ *has* Host-Kra complexity s *if* $\max_i \mathcal{HK}_i(\vec{P}) = s$.

---

[2]The definitions of factors, Weyl systems, nilsystems, and other concepts from ergodic theory and higher order Fourier analysis used in the introduction will be provided in subsequent sections.

Investigating complexity has been of particular interest for a class of dynamical systems called Weyl systems, leading to another notion of complexity, a variant of which is given below.

Definition 1.6 (Weyl complexity). *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ be an integral polynomial progression. Fix $0 \leqslant i \leqslant t$. The progression $\vec{P}$ has Weyl complexity s at i, denoted $\mathcal{W}_i(\vec{P})$, if s is the smallest natural number such that the factor $\mathcal{Z}_s$ is characteristic for the weak convergence of $\vec{P}$ at i for all Weyl systems $(X, \mathcal{X}, \mu, T)$. We say $\vec{P}$ has Weyl complexity s if $\max_i \mathcal{W}_i(\vec{P}) = s$.*

In previous works [BLL07, Lei09, Fra08, Fra16], the aforementioned notions of complexity have been defined for a polynomial family $\mathcal{P} = \{P_1, ..., P_t\}$ rather than for a progression $\vec{P}$. However, we want to extend the definitions of complexity to "index 0", i.e. the x term in $\vec{P}$, which is why we prefer to define it for $\vec{P}$ rather than $\mathcal{P}$. Similarly, complexity has previously been defined for L^2 convergence rather than weak convergence. However, the existence of L^2 limit (Theorem 1.4) and a basic functional analysis imply that weak and L^2 limits are identical.

Host-Kra factors are deeply related to a family of seminorms called *Gowers-Host-Kra seminorms*. For $s \in \mathbb{N}_+$ and $f \in L^\infty(\mu)$, the Gowers-Host-Kra seminorm of f of degree s is denoted by $|||f|||_s$ and satisfies the property

$$(4) \qquad |||f|||_{s+1} = 0 \iff \mathbb{E}(f|\mathcal{Z}_s) = 0$$

as well as the monotonicity property

$$(5) \qquad |||f|||_1 \leqslant |||f|||_2 \leqslant |||f|||_3 \leqslant ...$$

Gowers-Host-Kra seminorms have natural finitary analogues. For the transformation $Tx = x + 1$ on $X = \mathbb{Z}/N\mathbb{Z}$ with N prime and the uniform probability measure μ, the weak limit (3) becomes

$$(6) \qquad \mathbb{E}_{x,y \in \mathbb{Z}/N\mathbb{Z}} f_0(x) f_1(x + P_1(y)) \cdots f_t(x + P_t(y)).$$

The Gowers-Host-Kra seminorm of any $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ is a norm (for $s > 1$) called the *Gowers norm* and denoted by U^s, and it takes the form

$$(7) \qquad ||f||_{U^s} = \left(\mathbb{E}_{x,h_1,...,h_s \in \mathbb{Z}/N\mathbb{Z}} \prod_{w \in \{0,1\}^s} \mathcal{C}^{|w|} f(x + w_1 h_1 + ... + w_s h_s) \right)^{\frac{1}{2^s}},$$

where $\mathcal{C} : z \mapsto \overline{z}$ is the conjugation operator and $|w| = w_1 + \cdots + w_s$. As a result, $||f||_{U^s} = 0$ for some $s > 1$ if and only if $||f||_{U^2} = 0$ if and only if $f = 0$, and so inquiring about the smallest characteristic factor of this system in the sense of Definition 1.3 makes little sense. We can however ask which Gowers norm "controls" $\vec{P}$ in a more finitary way, and this leads to another notion of complexity.

Definition 1.7 (True complexity). *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ be an integral polynomial progression. Fix $0 \leqslant i \leqslant t$. The progression $\vec{P}$ has true complexity s at i, denoted $\mathcal{T}_i(\vec{P})$, if s is the smallest natural number with the following property: for every $\epsilon > 0$, there exist*

$\delta > 0$ and $N_0 \in \mathbb{N}$ such that for all primes $N > N_0$ and all functions $f_0, ..., f_t : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ satisfying $\max_i ||f_i||_\infty \leqslant 1$, we have

$$|\mathbb{E}_{x,y \in \mathbb{Z}/N\mathbb{Z}} f_0(x) f_1(x + P_1(y)) \cdots f_t(x + P_t(y))| < \epsilon$$

whenever $||f_i||_{U^{s+1}} < \delta$. We say $\vec{P}$ has true complexity s if $\max_i \mathcal{T}_i(\vec{P}) = s$.

We have so far defined three notions of complexity, that of Host-Kra, Weyl and true complexity. They are all defined in terms of ergodic theory or higher order Fourier analysis and have to do with "controlling" expressions like (2) and (6) by characteristic factors, Gowers-Host-Kra seminorms and Gowers norms. We shall now introduce one more notion, defined purely in terms of algebraic properties of polynomial progressions, and conjecture that all four concepts of complexity are in fact the same.

Definition 1.8 (Algebraic relations and algebraic complexity). *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression. An* algebraic relation *of degree $(j_0, ..., j_t)$ satisfied by $\vec{P}$ is a tuple $(Q_0, ..., Q_t) \in \mathbb{R}[u]^{t+1}$ such that*

$$(8) \qquad Q_0(x) + Q_1(x + P_1(y)) + ... + Q_t(P_t(y)) = 0,$$

where $\deg Q_i = j_i$ for each $0 \leqslant i \leqslant t$. The progression $\vec{P}$ has algebraic complexity s at i *for some $0 \leqslant i \leqslant t$, denoted $\mathcal{A}_i(\vec{P})$, if s is the smallest natural number such that for any algebraic relation $(Q_0, ..., Q_t)$ satisfied by $\vec{P}$, the degree of Q_i is at most s. It has* algebraic complexity s *if $\max_i \mathcal{A}_i(\vec{P}) = s$.*

Conjecture 1.9 (Four notions of complexity are the same). *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression. Fix $0 \leqslant i \leqslant t$. Then*

$$\mathcal{HK}_i(\vec{P}) = \mathcal{W}_i(\vec{P}) = \mathcal{T}_i(\vec{P}) = \mathcal{A}_i(\vec{P}) \leqslant t - 1.$$

The heuristic for Conjecture 1.9 is as follows: evaluating expressions like (3) and (6) comes down to understanding the distribution of certain polynomial sequences on nilmanifolds, and the only obstructions to equidistribution come from algebraic relations of the form (8).

Several substatements of Conjecture 1.9, such as the equivalence of Weyl and Host-Kra complexity and the upper bound on complexities, have previously been conjectured in [BLL07, Lei09, Fra08, Fra16]. Similarly, the equivalence of true and algebraic complexity has been studied and proved for linear forms [GW10, GW11a, GW11b, GW11c] as well as certain subclasses of polynomial progressions [Pel19, Kuc19, Kuc20]. However, we have not seen the full statement of Conjecture 1.9 anywhere in the literature. In particular, we have not found a conjecture relating Host-Kra and Weyl complexity to algebraic complexity, even though the aforementioned papers researching the topic mention that algebraic relations form a source of obstructions preventing a progression from having a characteristic small-degree Host-Kra factor.

Before we state our main result, we have to distinguish between two large families of progressions.

Definition 1.10 (Homogeneous and inhomogeneous relations and progressions). *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression. An* algebraic relation $(Q_0, ..., Q_t) \in \mathbb{R}[u]^{t+1}$ *is* homogeneous *of degree d if it is of the form*

$$(Q_0(u), ..., Q_t(u)) = (a_0 u^d, ..., a_t u^d)$$

for some $a_0, ..., a_t \in \mathbb{R}$ (some but not all of which may be zero), and inhomogeneous *otherwise. The progression $\vec{P}$ is* homogeneous *if all the algebraic relations that it satisfies are linear combinations of its homogeneous algebraic relations, and it is called* inhomogeneous *otherwise.*

An example of a homogeneous progression is $(x,\ x+y,\ x+2y,\ x+y^3)$, which only satisfies a homogeneous relation

(9) $$x - 2(x + y) + (x + 2y) = 0.$$

Other examples include arithmetic progressions, progressions with $P_1, ..., P_t$ being linearly independent such as $(x,\ x + y,\ x + y^2)$, or progressions whose terms satisfy no quadratic relations, such as $(x,\ x + y^2,\ x + 2y^2,\ x + y^3,\ x + 2y^3)$. By contrast, the progression $(x,\ x+y,\ x+2y,\ x+y^2)$ is inhomogeneous because it satisfies both (9) and the inhomogeneous relation

(10) $$x^2 + 2x - 2(x + y)^2 + (x + 2y)^2 - 2(x + y^2) = 0$$

that cannot be broken down into a sum of homogeneous relations. These two progressions will accompany us as running examples throughout the paper.

Our main result is the following.

Theorem 1.11 (Conjecture 1.9 holds for homogeneous progressions)**.** *Let $t \in \mathbb{N}+$. If $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ is a homogeneous polynomial progression, then it satisfies Conjecture 1.9.*

Although we have defined Host-Kra complexity using totally ergodic systems, we can extend our results to ergodic systems. The main difference is that if a system has complexity 0, then the $\mathcal{Z}_0$ factor has to be replaced by the rational Kronecker factor $\mathcal{K}_{rat}$.

Corollary 1.12. *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ be a homogeneous polynomial progression, and suppose that $\mathcal{A}_i(\vec{P}) = s$ for some $0 \leqslant i \leqslant t$ and $s \in \mathbb{N}$. For all invertible ergodic dynamical systems $(X, \mathcal{X}, \mu, T)$, the factor $\mathcal{Z}_s$ is characteristic for the weak or L^2 convergence of $\vec{P}$ at i if $s > 0$, and $\mathcal{K}_{rat}$ is characteristic for the weak or L^2 convergence of $\vec{P}$ at i if $s = 0$.*

Theorem 1.11 and Corollary 1.12 can be viewed as extensions of [HK05a, HK05b, FK05, FK06, Fra08, BLL07, Lei09], which find characteristic factors for linear configurations, linearly independent polynomials, progressions of length 4, examine Weyl complexity for arbitrary integral polynomial progression, and give an upper bound for Host-Kra complexity for general integral progressions. Theorem 1.11 also partly extends [GW10, GW11a, GW11b, GW11c, GT10, Pel19, Kuc19, Kuc20], which among other things determine true complexity for certain families of linear forms and integral polynomial progressions.

In particular, we extend our earlier work from [Kuc20]. In that paper, we prove equidistribution results on nilmanifolds for progressions of the form $(x,\ x + Q(y),\ x + R(y),\ x + Q(y) + R(y))$ with $\deg Q < \deg R$, or $(x,\ x + Q(y),\ x + 2Q(y),\ x + R(y),\ x + 2R(y))$ with $\deg Q < (\deg R)/2$, both of which are homogeneous. That equidistribution results follow from inducting on the filtration of a certain nilmanifold associated with the progression; the induction scheme involved is quite sensitive to the progression in question. Here, we achieve a much more general equidistribution result (part (i) of Theorem 1.15) by obtaining a solid understanding of the algebra behind homogeneous progressions and introducing a more flexible induction scheme.

From the fact that all progressions of algebraic complexity 1 are homogeneous, we deduce the following counting result.

Corollary 1.13. *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression of algebraic complexity at most 1. Suppose that $Q_1, ..., Q_d \in \mathbb{R}[y]$ are integral polynomials such that $P_i(y) = \sum_{j=1}^{d} a_{ij} Q_j(y)$ for $a_{ij} \in \mathbb{Z}$ for each $0 \leqslant i \leqslant t$ and $1 \leqslant j \leqslant d$. Let $L_i(y_1, ...y_d) = \sum_{j=1}^{d} a_{ij} y_j$. Then the following is true.*

(i) For any $f_0, ..., f_t : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ with $\max_i ||f_i||_\infty \leqslant 1$, we have

$$\mathbb{E}_{x,y \in \mathbb{Z}/N\mathbb{Z}} \prod_{i=0}^{t} f_i(x + P_i(y)) = \mathbb{E}_{x,y_1,...,y_d \in \mathbb{Z}/N\mathbb{Z}} \prod_{i=0}^{t} f_i(x + L_i(y_1, ..., y_d)) + o(1),$$

where the error term $o(1)$ is taken as $N \to \infty$ over primes and does not depend on the choice of $f_0, ..., f_t$.

(ii) For any invertible totally ergodic dynamical system $(X, \mathcal{X}, \mu, T)$ and $f_0, ..., f_t \in L^\infty(\mu)$, we have

$$\lim_{N \to \infty} \mathbb{E}_{n \in [N]} \int_X \prod_{i=0}^{t} T^{P_i(n)} f_i d\mu = \lim_{N \to \infty} \mathbb{E}_{n_1,...,n_d \in [N]} \int_X \prod_{i=0}^{t} T^{L_i(n_1,...,n_d)} f_i d\mu.$$

We shall illustrate Corollary 1.13 for the specific example of

$$\vec{P}(x, y) = (x, \; x + y^2, \; x + 2y^2, \; x + y^3, \; x + 2y^3).$$

Taking $Q_1(y) = y^2$ and $Q_2(y) = y^3$ as in the statement of Corollary 1.13, we let $\vec{L}(x, y_1, y_2) = (x, \; x + y_1, \; x + 2y_1, \; x + y_2, \; x + 2y_2)$. For any $A \subset \mathbb{Z}/N\mathbb{Z}$, we then have

$$|\{(x, y) \in (\mathbb{Z}/N\mathbb{Z})^2 : (x, \; x + y^2, \; x + 2y^2, \; x + y^3, \; x + 2y^3) \in A^5\}|$$
$$= |\{(x, y_1, y_2) \in (\mathbb{Z}/N\mathbb{Z})^3 : (x, \; x + y_1, \; x + 2y_1, \; x + y_2, \; x + 2y_2) \in A^5\}|/N + o(N^2)$$

upon setting $f_0 = ... = f_t = 1_A$. If $(X, \mathcal{X}, \mu, T)$ is a totally ergodic system and $A \in \mathcal{X}$, then we similarly obtain that

$$\lim_{N \to \infty} \mathbb{E}_{n \in [N]} \mu(A \cap T^{n^2} A \cap T^{2n^2} A \cap T^{n^3} A \cap T^{2n^3} A)$$
$$= \lim_{N \to \infty} \mathbb{E}_{n,m \in [N]} \mu(A \cap T^n A \cap T^{2n} A \cap T^m A \cap T^{2m} A).$$

For progressions of algebraic complexity 1, we also prove the following result, which generalises Theorem C of [Fra08], Theorem 1.12 of [GT10], and results from [BHK05]. In additive combinatorics, problems of this type are known as finding popular common differences; in ergodic theory, one speaks of establishing lower bounds for multiple recurrence.

Theorem 1.14. *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression of algebraic complexity at most 1, with the following property: there exist linearly independent integral polynomials $Q_1, ..., Q_k$ such that*

$$(11) \qquad \{a_1 Q_1 + ... + a_k Q_k : \; a_1, ..., a_k \in \mathbb{Z}\} = \{b_1 P_1 + ... + b_t P_t : \; b_1, ..., b_t \in \mathbb{Z}\}.$$

Then the following is true.

(i) Let $(X, \mathcal{X}, \mu, T)$ be an ergodic invertible measure preserving system and $A \in \mathcal{X}$. Suppose that $\mu(A) > 0$. Then for every $\epsilon > 0$, the set

$$\{n \in \mathbb{N} : \mu(A \cap T^{P_1(n)} A \cap ... \cap T^{P_t(n)} A) > \mu(A)^{t+1} - \epsilon\}$$

is syndetic, i.e. it has bounded gaps.

(ii) Suppose that $A \subset \mathbb{N}$ has upper density $\alpha > 0$. Then for every $\epsilon > 0$, the set

$$\{n \in \mathbb{N} : \mu(A \cap (A + P_1(n)) \cap ... \cap (A + P_t(n))) > \alpha^{t+1} - \epsilon\}$$

is syndetic.

(iii) For any $\alpha, \epsilon > 0$ and prime N, and any subset $A \subset \mathbb{Z}/N\mathbb{Z}$ of size $|A| \geqslant \alpha N$, we have

$$|\{n \in \mathbb{Z}/N\mathbb{Z} : |A \cap (A + P_1(n)) \cap ... \cap (A + P_t(n))| > (\alpha^{t+1} - \epsilon)N\}| \gg_{\alpha, \epsilon} N.$$

The definition of homogeneity (Definition 1.10) is equivalent to a certain linear algebraic property that will be described in details in Section 4; this property makes it possible to explicitly describe closures of orbits of nilsequences evaluated at terms of homogeneous polynomial progressions, from which we deduce Theorem 1.11. Homogeneous polynomial progressions are moreover the largest family of integral polynomial progressions for which such an explicit description is possible, and even the simplest examples of inhomogeneous progressions lead to complications absent in the homogeneous case. The following result makes this precise. As with all other results in this section, all the concepts in Theorem 1.15 are explained in subsequent sections.

Theorem 1.15 (Dichotomy between homogeneous and inhomogeneous progressions). *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ be an integral polynomial progression. Suppose that G is a connected, simply-connected, nilpotent Lie group with a rational filtration $G_\bullet$ and Γ is a cocompact lattice. There exists a subnilmanifold G^P/Γ^P of G^{t+1}/Γ^{t+1} with the following property.*

(i) *If $\vec{P}$ is homogeneous, then for every irrational polynomial sequence $g : \mathbb{Z} \to G$ adapted to $G_\bullet$, the sequence*

$$g^P(x, y) = (g(x), \ g(x + P_1(y)), \ ..., \ g(x + P_t(y)))$$

is equidistributed on G^P/Γ^P.

(ii) *If $\vec{P}$ is inhomogeneous, then for every irrational polynomial sequence $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$, the closure of g^P is a union of finitely many translates of a subnilmanifold of G^P/Γ^P. For every $\vec{P}$, we can moreover find a filtered nilmanifold G/Γ and an irrational polynomial sequence $g : \mathbb{Z} \to G$ such that g^P is equidistributed on a proper subnilmanifold of G^P/Γ^P.*

While we have not been able to prove full Conjecture 1.9 for inhomogeneous progressions, we are able to say a bit more about the relationship between various notions of complexity in the general case.

Theorem 1.16. *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression. Fix $0 \leqslant i \leqslant t$. Then*

$$\mathcal{W}_i(\vec{P}) = \mathcal{A}_i(\vec{P}) \leqslant \min(\mathcal{T}_i(\vec{P}), \mathcal{HK}_i(\vec{P})).$$

Of the various statements made in Theorem 1.16, the fact that Host-Kra complexity bounds Weyl complexity is a simple consequence of definitions and shall be explained in Section 11. Similarly, the fact that algebraic complexity is bounded from above by true complexity has been shown in Theorem 1.13 of [Kuc20]. It is the equivalence of Weyl and algebraic complexities that is a new statement here.

Outline of the paper. We start the paper by introducing basic ergodic theoretic definitions and results concerning nilsystems in Section 2, and we explain why analyzing expressions like (3) comes down to answering equidistribution questions on nilmanifolds. We then show in Section 3 that in studying equidistribution on nilmanifolds, we can restrict ourselves to nilmanifolds that are quotients of connected groups at the expense of replacing a linear sequence by a polynomial one.

Section 4 explains key differences between homogeneous and inhomogeneous progressions, and in particular it shows the upper bound on algebraic complexity for homogeneous progressions in Theorem 1.11. Definitions introduced in this section allow us to state in the infinitary version of an equidistribution result for homogeneous polynomial progressions on nilmanifolds (Theorem 5.3) in Section 5, from which we deduce that for homogeneous progressions, Host-Kra complexity is bounded from above by algebraic complexity (Corollary 5.4). We further use Theorem 5.3 to deduce Corollaries 1.12 and 1.13(ii).

In Section 6, we introduce finitary analogues of tools from Section 2. These are needed in Section 7, in which we show that proving the equivalence of true and algebraic complexity for homogeneous progression comes down to proving Theorem 6.7, a finitary version of Theorem 5.3. We also explain in Section 7 how to prove Corollary 1.13(i). Theorem 6.7, the main technical part of this paper, is derived in Section 8. Unfortunately, Theorem 6.7 fails for inhomogeneous progressions, as explained in Section 9. In Section 10, we propose a method to handle inhomogeneous progressions. While we succeed in proving an analogue of Theorem 5.3 for the inhomogeneous progression $(x,\ x+y,\ x+2y,\ x+y^2)$ in Proposition 10.1, we have been unable to extend this construction to all inhomogeneous progressions. Subsequently, we show in 11 that Weyl and algebraic complexity are always equal, which is the main statement of Theorem 1.16. We conclude the paper by proving Theorem 1.14 in Section 12.

2. Infinitary nilmanifold theory

2.1. **Basic definitions from ergodic theory.** Let $(X, \mathcal{X}, \mu, T)$ be an invertible measure-preserving dynamical system (henceforth, we shall simply call it a *system*). The background in ergodic theory that we need can be found in [HK05b, HK18], among others; here, we only reiterate the most important definitions.

Definition 2.1. *A* factor *of a system* $(X, \mathcal{X}, \mu, T)$ *can be defined in three equivalent ways:*

(i) *it is a T-invariant sub-σ-algebra of $\mathcal{X}$;*

(ii) *it is a system $(Y, \mathcal{Y}, \nu, S)$ together with a* factor map *$\pi : X' \to Y'$, i.e. a measurable map defined for a measurable T-invariant set X' of full measure, satisfying $S \circ \pi = \pi \circ T$ on X' and $\mu \circ \pi^{-1} = \nu$;*

(iii) *it is a T-invariant subalgebra of $L^{\infty}(\mu)$.*

For $r \in \mathbb{N}$, we let $\mathcal{K}_r$ be the factor spanned by all T^r-invariant functions in $L^{\infty}(\mu)$. In particular, $\mathcal{K}_1 = \mathcal{I}$ is the factor spanned by T-invariant functions, and the *rational Kronecker factor* $\mathcal{K}_{rat} = \bigvee_{r \in \mathbb{N}} \mathcal{K}_r$ is the factor spanned by all the functions in in $L^{\infty}(\mu)$ that are T^r-invariant for some $r \in \mathbb{N}$. A system is *ergodic* if $\mathcal{K}_1 = \mathcal{I}$ is the trivial factor spanned by constant functions, and it is *totally ergodic* if $\mathcal{K}_{rat}$ is the trivial factor.

Of particular interest to us is a sequence of factors $(\mathcal{Z}_s)_{s \in \mathbb{N}}$ defined in [HK05b], which we refer to as *Host-Kra factors*. In accordance with Definition 2.1, we shall sometimes think of $\mathcal{Z}_s$ as a sub-σ-algebra of $\mathcal{X}$, and at other times we will consider a factor map $\pi_s : X \to Z_s$ and a factor $(Z_s, \mathcal{Z}_s, \lambda, S)$ of $(X, \mathcal{X}, \mu, T)$. If we concurrently talk about Host-Kra factors of two distinct spaces X and Y, we may write $Z_s(X)$ and $Z_s(Y)$ to mean Host-Kra factors of X and Y respectively. We do not explicitly use the definition of Host-Kra factors anywhere in the paper, and so we leave the interested reader to look it up in [HK05b, HK18]. Instead, we rely on two properties of this family of factors that concern their utility and structure respectively. First, these factors are characteristic for the convergence of polynomial progressions, as proved in Theorem 1.4. Rephrasing Theorem 1.4 in terms of Definition 1.5, we can say that each integral polynomial progression has a finite Host-Kra complexity. Second, each factor $\mathcal{Z}_s$ is an inverse limit[3] of s-step nilsystems, which are objects of primary importance to us.

2.2. **Nilsystems.** Let G be a Lie group with connected component G^0 and identity 1. A *filtration* on G of degree s is a chain of subgroups

$$G = G_0 = G_1 \geqslant G_2 \geqslant ... \geqslant G_s \geqslant G_{s+1} = G_{s+2} = ... = 1$$

satisfying $[G_i, G_j] \leqslant G_{i+j}$ for each $i, j \in \mathbb{N}$. We denote it as $G_{\bullet} = (G_i)_{i=0}^{\infty}$. A natural example of filtration is the lower central series, given by $G_{k+1} = [G, G_k]$ for each $k > 1$, where the *commutator* of two elements $a, b \in G$ is defined as $[a, b] = a^{-1}b^{-1}ab$, and $[A, B]$ is the subgroup of G generated by all the commutators $[a, b]$ with $a \in A, b \in B$. The group G is *s-step nilpotent* if $G_{s+1} = 1$, where G_{s+1} is the s-th element of the lower central series of G. The only 0-step nilpotent group is the trivial group, and 1-step nilpotent groups are precisely abelian groups.

For the rest of the paper, we let G be a nilpotent Lie group and $\Gamma \leqslant G$ be a cocompact lattice. We call the quotient $X = G/\Gamma$ a *nilmanifold*. The group G acts on X by left translation, and for each $a \in G$, we call the map $T_a(g\Gamma) = (ag)\Gamma$ a *nilrotation*. Setting $\mathcal{G}/\Gamma$ to be the Borel σ-algebra of X and ν to be the Haar measure with respect to left translation, we call the system $(G/\Gamma, \mathcal{G}/\Gamma, \nu, T_a)$ a *nilsystem*.

---

[3]The system $(X, \mathcal{X}, \mu, T)$ is an *inverse limit* of a sequence of factors $(X, \mathcal{X}_i, \mu, T)$ if $\mathcal{X}_i$ form an increasing sequence of factors of $\mathcal{X}$ such that $\mathcal{X} = \bigvee_{i \in \mathbb{N}} \mathcal{X}_i$ up to sets of measure zero.

A subgroup $H \leqslant G$ is *rational* if $H/(H \cap \Gamma)$ is closed in G/Γ. A filtration $G_\bullet$ is rational if G_i is a rational subgroup for each $i \in \mathbb{N}$. We shall assume throughout the paper that each filtration that we discuss is rational.

In the case when $(G/\Gamma, \mathcal{G}/\Gamma, \nu, T_a)$ is an ergodic nilsystem, which will always be our case anyway, we can make two simplifying assumptions about the group G. By passing to universal cover, we assume that G is simply connected. Replacing the nilsystem with several simpler nilsystems, we further assume that G is spanned by G^0 and a. These assumptions, justified in Chapter 11 of [HK18], hold for the rest of the paper.

We also denote $\Gamma_i = G_i \cap \Gamma$ and $\Gamma^0 = G^0 \cap \Gamma$. The rationality of G_i in G means that Γ_i is cocompact in G_i.

Proposition 2.2 (Conditions for total ergodicity of nilsystems, Corollary 7 and 8 of [HK18])**.** *Let $(G/\Gamma, \mathcal{G}/\Gamma, \nu, T_a)$ be an ergodic nilsystem. There exists $r \in \mathbb{N}_+$ such that $T_a^j(G^0/\Gamma^0)$ is totally ergodic with respect to T_a^r for all $0 \leqslant j < r$.*

Moreover, the following are equivalent:

(i) T_a is totally ergodic;

(ii) G/Γ is connected;

(iii) $G = G^0\Gamma$.

Nilsystems allow a particularly simple description of factors. If $G_\bullet$ is the lower central series filtration, then

$$(12) \qquad\qquad Z_s = \frac{G}{G_{s+1}\Gamma}$$

for all $s \in \mathbb{N}_+$ (see Chapter 11 of [HK18]). For $s = 0$, we have $Z_0 = G/(G^0\Gamma) \cong (\mathbb{Z}/r\mathbb{Z})$, where r is the smallest positive integer for which $a^r \in G^0$. It follows from Proposition 2.2 that Z_0 is trivial if and only if the nilsystem is totally ergodic.

Let $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression. By Theorem 1.4, there exists $s \in \mathbb{N}$ such that for every ergodic system $(X, \mathcal{X}, \mu, T)$ and all choices of $f_0, ..., f_t \in L^\infty(\mu)$, we have

$$\lim_{N \to \infty} \mathbb{E}_{n \in [N]} \int_X f_0 \cdot T^{P_1(n)} f_1 \cdot ... \cdot T^{P_t(n)} f_t d\mu$$

$$(13) \qquad = \lim_{N \to \infty} \mathbb{E}_{n \in [N]} \int_{Z_s} \mathbb{E}(f_0|\mathcal{Z}_s) \cdot S^{P_1(n)} \mathbb{E}(f_1|\mathcal{Z}_s) \cdot ... \cdot S^{P_t(n)} \mathbb{E}(f_t|\mathcal{Z}_s) d\lambda$$

Using the fact that Z_s is an inverse limit of ergodic s-step nilsystems, we can approximate the average (13) arbitrarily well by projections onto ergodic nilsystems. Hence we are left with understanding averages of the form

$$(14) \qquad \lim_{N \to \infty} \mathbb{E}_{n \in [N]} \int_{G/\Gamma} \tilde{f}_0(b\Gamma) \cdot \tilde{f}_1(a^{P_1(n)}b\Gamma) \cdot ... \cdot \tilde{f}_t(a^{P_t(n)}b\Gamma) d\nu(b\Gamma)$$

where $\tilde{f}_i$ is the projection of f_i onto an ergodic s-step nilsystem $(G/\Gamma, \mathcal{G}/\Gamma, \nu, T_a)$ for all $0 \leqslant i \leqslant t$. If T is totally ergodic, then so is the nilrotation T_a.

2.3. **Polynomial sequences.** Let $G_\bullet$ be a filtration on G of degree s. A *polynomial sequence* $g : \mathbb{Z} \to G$ adapted to $G_\bullet$ is a sequence

$$(15) \qquad g(n) = \prod_{i=0}^{s} g_i^{\binom{n}{i}}.$$

with the property that $g_i \in G_i$ for each i. Such sequences form a group denoted as $\mathrm{poly}(\mathbb{Z}, G_\bullet)$ by Proposition 6.2 of [GT12]. One may ask why we define polynomial sequence as (15) rather than in the seemingly more natural form

$$(16) \qquad g(n) = \prod_{i=0}^{s} g_i^{n^i}.$$

The reason is that if g is written in the form (15), then we have the following nice statement.

Lemma 2.3 (Lemma 2.8 of [CS12]). *Suppose that $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$. The sequence $g(n) = \prod_{i=0}^{s} g_i^{\binom{n}{i}}$ takes values in $H \leqslant G$ if and only if $g_0, ..., g_s \in H$.*

Proof. The converse direction is straightforward, and we prove the forward direction by induction on $0 \leqslant k \leqslant s$. For $k = 0$, we observe that $g_0 = g(0) \in H$. Suppose that the statement holds for k, i.e. $g_0, ..., g_k \in H$. Then $g(k+1) = \left(\prod_{i=0}^{k} g_i^{\binom{k}{i}} \right) g_{k+1}$. Since $g(k+1), g_0, ..., g_k$ are all in H, it follows that $g_{k+1} \in H$. $\qquad \square$

Lemma 2.3 is not true if g is written in the form (16); for instance, $g(n) = \binom{n}{2} = \frac{1}{2}n^2 - \frac{1}{2}n$ takes values in $\mathbb{Z}$ even though $\frac{1}{2}, -\frac{1}{2} \notin \mathbb{Z}$.

In a similar manner, we define for any $D \in \mathbb{N}_+$ the group $\mathrm{poly}(\mathbb{Z}^D, G_\bullet)$ of D-parameter polynomial sequence $g : \mathbb{Z}^D \to G$ adapted to $G_\bullet$, i.e. sequences of the form

$$g(n_1, ..., n_D) = \prod_{i=0}^{s} \prod_{i_1+...+i_D=i} g_{i_1,...,i_D}^{\binom{n_1}{i_1} \cdots \binom{n_D}{i_D}}$$

for $g_{i_1,...,i_D} \in G_{i_1+...i_D}$.

2.4. **Infinitary equidistribution theory on nilmanifolds.** For the rest of Section 2, we assume that G is connected. For $D \in \mathbb{N}_+$, a polynomial sequence $g \in \mathrm{poly}(\mathbb{Z}^D, G_\bullet)$ is *equidistributed* on G/Γ if

$$\mathbb{E}_{n \in [N]^D} F(g(n)\Gamma) \to \int_{G/\Gamma} F d\nu$$

for any continuous function $F : G/\Gamma \to \mathbb{C}$. The following notion is useful when discussing equidistribution.

Definition 2.4 (Horizontal characters). *A horizontal character on G is a continuous group homomorphism $\eta : G \to \mathbb{R}$ for which $\eta(\Gamma) \leqslant \mathbb{Z}$.*

In particular, each horizontal character vanishes on $[G, G]$.

Equidistribution on nilmanifolds was studied by Leibman, who provided a useful criterion for when a polynomial sequence is equidistributed on a nilmanifold. We only need the version of the statement in the case when G is connected, as we will be able to reduce to this case.

Theorem 2.5 (Leibman's equidistribution theorem, [Lei05b])**.** *Let $D \in \mathbb{N}_+$ and $g \in \mathrm{poly}(\mathbb{Z}^D, G_\bullet)$. The following are equivalent:*

(i) *g is equidistributed in G/Γ;*

(ii) *the projection of g onto $G/[G,G]$ is equidistributed in $G/[G,G]\Gamma$;*

(iii) *if $\eta : G \to \mathbb{R}$ is a horizontal character for which $\eta \circ g$ is constant, then η is trivial.*

We shall also need a stronger notion of equidistribution, that of irrational sequences.

Definition 2.6. *Suppose that $G_\bullet$ is a filtration on G and $i \in \mathbb{N}_+$, and let*

$$G_i^\nabla = \langle G_{i+1}, [G_j, G_{i-j}], 1 \leqslant j < i \rangle$$

An i-th level character is a continuous group homomorphism $\eta_i : G_i \to \mathbb{R}$ that vanishes on G_i^∇ and satisfies $\eta_i(\Gamma_i) \in \mathbb{Z}$. An element g_i of G_i is irrational if $\eta_i(g_i) \notin \mathbb{Z}$ for any nontrivial i-th level character η_i. A sequence $g(n) = \prod\limits_{i=0}^s g_i^{\binom{n}{i}}$ is irrational if g_i is irrational for all $i \in \mathbb{N}_+$.

All irrational sequences are equidistributed, but not vice versa. For instance, let $g(n) = a_1 n + ... + a_s n^s$ be a real-valued polynomial. It is a polynomial sequence in $\mathbb{R}$ adapted to the filtration $G_1 = ... = G_s = \mathbb{R}$, $G_{s+1} = 0$. Thus, g is irrational iff $a_s \notin \mathbb{Q}$, and g is equidistributed iff there exists $1 \leqslant i \leqslant s$ with $a_i \notin \mathbb{Q}$. It is clear in this case that irrational implies equidistributed, but not vice versa.

We want to emphasise that whether a sequence is irrational or not depends on what filtration we are using, whereas the notion of equidistribution does not depend on the filtration.

3. Reducing to the case of connected groups

The expression (14) indicates that to understand Host-Kra complexity of a polynomial progression $\vec{P}$, we have to understand the distribution of orbits

$$(17) \qquad\qquad (b\Gamma,\ a^{P_1(n)}b\Gamma,\ ...,\ a^{P_t(n)}b\Gamma)$$

inside a connected nilmanifold G^{t+1}/Γ^{t+1}. The point of this section is to show that we can replace linear orbits $(a^n b\Gamma)_{n\in\mathbb{N}}$ on G/Γ by polynomial orbits $(g_b(n)\Gamma^0)_{n\in\mathbb{N}}$ on G^0/Γ^0 for some irrational polynomial sequence $g_b : \mathbb{Z} \to G^0$ with respect to a certain naturally defined filtration $G_\bullet^0$ on G_0. This way, we want to reduce the question of finding the closure for (17) inside $(G/\Gamma)^{t+1}$ to finding the closure for

$$(18) \qquad\qquad (g_b(m)\Gamma^0,\ g_b(m + P_1(n))\Gamma^0,\ ...,\ g_b(m + P_t(n))\Gamma^0)$$

inside $(G^0/\Gamma^0)^{t+1}$. The connectedness of G^0 allows us to use tools from Section 2.4.

Lemma 3.1. *Let $(G/\Gamma, \mathcal{G}/\Gamma, \nu, T_a)$ be a totally ergodic nilsystem and $F : (G/\Gamma)^{t+1} \to \mathbb{R}$ be essentially bounded. Then*

$$\mathbb{E}_{n\in[N]} \int_{G/\Gamma} F(b\Gamma, a^{P_1(n)}b\Gamma, ..., a^{P_t(n)}b\Gamma)d\nu(b\Gamma)$$

$$= \mathbb{E}_{m,n\in[N]} \int_{G/\Gamma} F(a^m b\Gamma, a^{m+P_1(n)}b\Gamma, ..., a^{m+P_t(n)}b\Gamma)d\nu(b\Gamma).$$

Proof. Since T_a is measure preserving, we have

$$\int_{G/\Gamma} F(b\Gamma, a^{P_1(n)}b\Gamma, ..., a^{P_t(n)}b\Gamma)d\nu(b\Gamma) = \int_{G/\Gamma} F(a^m b\Gamma, a^{m+P_1(n)}b\Gamma, ..., a^{m+P_t(n)}b\Gamma)d\nu(b\Gamma)$$

for any $m, n \in \mathbb{N}$. Consequently,

$$\int_{G/\Gamma} F(b\Gamma, a^{P_1(n)}b\Gamma, ..., a^{P_t(n)}b\Gamma)d\nu(b\Gamma)$$
$$= \mathbb{E}_{m \in [N]} \int_{G/\Gamma} F(a^m b\Gamma, a^{m+P_1(n)}b\Gamma, ..., a^{m+P_t(n)}b\Gamma)d\nu(b\Gamma),$$

from which the lemma follows. $\qquad\square$

The main result of this section is the following.

Proposition 3.2. *Let $(G/\Gamma, \mathcal{G}/\Gamma, \nu, T_a)$ be a totally ergodic nilsystem and $b \in G^0$. Suppose that $G_\bullet$ is the lower central series filtration on G and $G_\bullet^0 = G_\bullet \cap G^0$. Then there exists an irrational sequence $g_b \in \mathrm{poly}(\mathbb{Z}, G_\bullet^0)$ such that $g_b(n)\Gamma = a^n b\Gamma$.*

We observe that with this filtration on G^0, we have $G_k^0 = G_k$ for $k \geqslant 2$. That follows from the fact that the groups G^k are connected for $k \geqslant 2$ (Lemma 5 of [HK18]), and hence are contained in G^0.

We lose no generality in assuming that $b \in G^0$; Proposition 2.2 and the connectedness of G/Γ imply that for all $b \in G$ there exists $b' \in G^0$ such that $b\Gamma = b'\Gamma$.

Proof. The connectedness of G/Γ implies that $G = G^0\Gamma$, and so there exist $\alpha \in G^0$ and $\gamma \in \Gamma$ such that $a = \alpha\gamma^{-1}$. Then

$$a^n b\Gamma = (\alpha\gamma^{-1})^n b\Gamma = (\alpha\gamma^{-1})^n b\gamma^n\Gamma.$$

It follows from normality of G^0 and the fact that α and b are elements of G^0 that the sequence $g_b(n) = (\alpha\gamma^{-1})^n b\gamma^n$ takes values in G^0. Since the sequences $h_1(n) = a^n b$ and $h_2(n) = \gamma^n$ are adapted to $G_\bullet$, and the set $\mathrm{poly}(\mathbb{Z}, G_\bullet)$ is a group, we deduce that $g_b = h_1 h_2$ is adapted to $G_\bullet^0 = G_\bullet \cap G^0$.

We want a more precise description of g_b, and for this we shall use some results from Sections 11-13 of [Lei09]. Let $g = g_b$ for the identity $b = 1$; that is, $g(n) = (\alpha\gamma^{-1})^n \gamma^n$. Leibman showed in Section 11.2 of [Lei09] that

(19)
$$g(n) = \prod_{1 \leqslant k_1 \leqslant s} (A^{k-1}\alpha)^{q_{k_1}(n)} \prod_{1 \leqslant k_2 < k_1 < s} [A^{k_1-1}\alpha, A^{k_2-1}\alpha]^{q_{k_1,k_2}(n)}...,$$

where $Ax = [x, \gamma]$ and $q_{k_1,...,k_r}$ are integral polynomials with $\deg q_{k_1,...,k_r} \leqslant k_1 + ... + k_r$. More explicitly, we have

(20)
$$g(n) = \alpha^n (A\alpha)^{\binom{n}{2}}(A^2\alpha)^{\binom{n}{3}} \cdots [A\alpha, \alpha]^{\binom{n}{3}}[A^2\alpha, \alpha]^{\binom{n}{4}} \cdots [A^2\alpha, A\alpha]^{4\binom{n+1}{5}}[A^3\alpha, A\alpha]^{5\binom{n+1}{6}} \cdots$$

The coefficients of g can be analysed using a family of subgroups of G^0 introduced in Section 12 of [Lei09]. For $k_1, ..., k_l \in \mathbb{N}_+$, we let $G_{(k_1,...,k_l)}^0$ be the subgroup of G^0 generated by all

l-fold commutators[4] of elements of the form $A^{k_1-1}h_1, ..., A^{k_l-1}h_l$ for $h_1, ..., h_l \in G^0$. We then define

$$G_{k,l}^0 = \langle G_{(k_1,...,k_i)}^0 : i \geqslant l, \ k_1 + ... + k_l \geqslant k \rangle$$

for integers $1 \leqslant l \leqslant k$ and set $G_{k,l}^0 = G_{l,l}^0$ whenever $l > k$.

The following lemma lists some basic properties of the groups $G_{k,l}^0$ that we shall use.

Lemma 3.3. *For any integers $1 \leqslant l \leqslant k$,*

(i) $G_{k,l}^0$ *is normal in G;*

(ii) $[G_{k,l}^0, G_{i,j}^0] \leqslant G_{k+i,l+j}^0$ *for any integers $1 \leqslant i \leqslant j$;*

(iii) $A^j G_{k,l}^0 \leqslant G_{k+j,l}^0$ *for any $j \in \mathbb{N}$;*

(iv) $G_{k+1,l}^0$ *and $G_{k,l+1}^0$ are subgroups of $G^{k,l}$, and the quotient groups $G_{k,l}^0/G_{k+1,l}^0$ and $G_{k,l}^0/G_{k,l+1}^0$ are abelian;*

(v) *for $k \geqslant 2$, $G_k = G_k^0 = G_{k,1}^0 = \langle A^{k-1}G^0, G_{k,2}^0 \rangle = \langle AG_{k-1}^0, G_{k,2}^0 \rangle$;*

(vi) $(G^0)_k^\nabla = \langle G_{k,2}^0, G_{k+1}^0 \rangle$

Proof. Properties (i)-(iv) are proved in Lemma 12.2 of [Lei09]. For $k \geqslant 2$, the statement $G_k = G_k^0$ in (v) is true by definition, and the statement $G_k = G_{k,1}^0$ is proved in Lemma 12.3 of [Lei09]. To finish the proof of (v), it remains to show that $G_{k,1}^0 = \langle A^{k-1}G^0, G_{k,2}^0 \rangle = \langle AG_{k-1}^0, G_{k,2}^0 \rangle$ for $k \geqslant 2$. For $k = 2$, this is true by definition of $G_{k,1}^0$ and the fact that $G_{k,2}^0 \geqslant G_{k,3}^0 \geqslant ...$, which follows from part (iv). We assume that the statement is true for some $k \geqslant 2$. That G_{k+1}^0 contains $\langle AG_k^0, G_{k+1,2}^0 \rangle$ follows from the fact that both AG_k^0 and $G_{k+1,2}^0$ are contained in the $(k+1)$-th element of the lower central series of G, which is precisely G_{k+1}^0. For the other direction, we observe that

$$G_{k+1}^0 = [G_k, G] = [G_k^0, \langle G^0, \gamma \rangle] \leqslant \langle [G_k^0, G^0], [G_k^0, \gamma] \rangle$$
$$\leqslant \langle [A^{k-1}G^0, G^0], [G_{k,2}^0, G^0], AG_k^0 \rangle \leqslant \langle G_{k+1,2}^0, AG_k^0 \rangle.$$

A similar argument shows that $G_{k+1}^0 = \langle A^k G^0, G_{k+1,2}^0 \rangle$.

Before we prove property (vi), we recall that $(G^0)_k^\nabla = \langle G_{k+1}, [G_j, G_{k-j}] : 1 \leqslant j < k \rangle$. That (vi) holds for $k = 1$ can be verified by inspection. For $k \geqslant 2$, we observe that $[A^{j-1}G^0, A^{k-j-1}G^0] \leqslant [G_j^0, G_{k-j}^0]$, and so

$$G_{k,2}^0 \leqslant \langle [G_j^0, G_{k-j}^0] : 1 \leqslant j < k \rangle;$$

when coupled with property (v), this implies that $(G^0)_k^\nabla \geqslant \langle G_{k,2}^0, G_{k+1}^0 \rangle$. For the converse, we have

$$[G_j^0, G_{k-j}^0] = [\langle A^{j-1}G^0, G_{j,2}^0 \rangle, \langle A^{k-j-1}G^0, G_{k-j,2}^0 \rangle] \leqslant \langle G_{k,2}^0, G_{k,3}^0, G_{k,4}^0 \rangle \leqslant G_{k,2}^0,$$

for each $1 \leqslant j < k$, from which it follows that $(G^0)_k^\nabla \leqslant \langle G_{k,2}^0, G_{k+1}^0 \rangle$. $\square$

---

[4]A 1-fold commutator is any element $h \in G$. For $l > 1$, an l-fold commutator is an element of the form $[h_i, h_j]$, where h_i is an i-fold commutator, h_j is an j-fold commutator and $i + j = l$.

Letting $g(n) = \prod_{i=1}^{s} g_i^{\binom{n}{i}}$, we observe from (19), (20) as well as parts (v) and (vi) of Lemma 3.3 that

$$(21) \qquad\qquad g_i = A^{i-1}\alpha \mod (G^0)_i^\nabla.$$

For an arbitrary $b \in G^0$, we have $g_b(n) = a^n b \gamma^n = b(\alpha_b \gamma^{-1})^n \gamma^n$, where $\alpha_b = \alpha[\alpha, b]Ab$, as observed in Section 11.3 of [Lei09]. Letting $g_b(n) = \prod_{i=0}^{s} g_{b,i}^{\binom{n}{i}}$, it is therefore true that

$$(22) \qquad\qquad g_{b,i} = A^{i-1}\alpha_b = A^{i-1}\alpha \mod (G^0)_i^\nabla$$

for all $i \in \mathbb{N}_+$.

For $i = 1$, we have $g_{b,1} = \alpha \mod G_2^0$, and we claim that $g_{b,i}$ is irrational. The ergodicity of a implies that for almost every b, the sequence $n \mapsto a^n b$ is equidistributed in G/Γ, and so the same is true for the sequence g_b in G^0/Γ^0. Consequently, the projection $\pi(g_b) : \mathbb{Z} \to G^0/(G_2^0\Gamma^0)$ is equidistributed as well. Since $\pi(g_b(n)) = \pi(b) + \pi(\alpha)n$, it follows that $\pi(\alpha)$ is an irrational element of G^0/G_2^0, and so $g_{b,1}$ is an irrational element of G^0.

Before proving that $g_{b,i}$ are irrational for $i > 1$, we discuss some properties of the map $A : G \to G$. From the definition of the filtration $G_\bullet^0$ we observe that $AG_i^0 \leqslant G_{i+1}^0$ for all $i \geqslant 1$ (this is also a consequence of parts (iv) and (v) of Lemma 3.3). Therefore the map $A_i := A|_{G_i^0}$ takes values in G_{i+1}^0, and moreover $A_i(\Gamma_i) \leqslant \Gamma_{i+1}$. We also observe that the projection $\overline{A}_i : G_i^0 \to G_{i+1}^0/(G^0)_{i+1}^\nabla$ is a (continuous) group homomorphism because

$$A(xy) = [xy, \gamma] = [x, \gamma][[x, \gamma], y][y, \gamma] = Ax[Ax, y]Ay = AxAy \mod G_{2i+1,2}^0$$

for any $x, y \in G_i^0$ and $G_{2i+1,2}^0 \leqslant G_{i+1,2}^0 \leqslant (G^0)_{i+1}^\nabla$ by parts (iv) and (vi) of Lemma 3.3. From part (v) of Lemma 3.3 it follows that $\overline{A}_i$ is surjective. Finally, we note using parts (iii) and (v) of Lemma 3.3 that $A_i((G^0)_i^\nabla) \leqslant (G^0)_{i+1}^\nabla$.

Suppose that $g_{b,i}$ is irrational but $g_{b,i+1}$ is not for some $1 \leqslant i < s$. Then there exists a nontrivial $(i+1)$-th level character $\eta_{i+1} : G_{i+1}^0 \to \mathbb{R}$ such that $\eta_{i+1}(g_{b,i+1}) \in \mathbb{Z}$. From (22) and the fact that η_{i+1} vanishes on $(G^0)_{i+1}^\nabla$, we deduce that $\eta_{i+1}(g_{b,i+1}) = \eta_{i+1}(A^i\alpha)$. We also let $\overline{\eta}_{i+1} : G_{i+1}^0/(G^0)_{i+1}^\nabla \to \mathbb{R}$ be the induced map.

Let $\eta_i := \eta_{i+1} \circ A_i : G_i^0 \to \mathbb{R}$. It is an i-th level character as a consequence of four facts: the vanishing of η_{i+1} on $(G^0)_{i+1}^\nabla$, the inclusion $(G_{i+1,2}^0) \leqslant (G^0)_{i+1}^\nabla$ (both of which imply that $\eta_i = \overline{\eta}_{i+1} \circ \overline{A}_i$ is a continuous group homomorphism), the inclusion $A_i((G^0)_i^\nabla) \leqslant (G^0)_{i+1}^\nabla$, and the fact that $\eta_i(\Gamma_i) \leqslant \mathbb{Z}$. It moreover satisfies

$$\eta_i(g_{b,i}) = \eta_i(A^{i-1}\alpha) = \eta_{i+1}(A^i\alpha) = \eta_{i+1}(g_{b,i+1}),$$

implying that $\eta_i(g_{b,i}) \in \mathbb{Z}$. The nontriviality of η_{i+1} implies that $\overline{\eta}_{i+1}$ and $\overline{A}_i$ are surjective maps onto nontrivial groups; hence η_i is nontrivial. This contradicts the irrationality of $g_{b,i}$. By induction, $g_{b,1}, ..., g_{b,s}$ are all irrational, implying that g_b is irrational. $\qquad\square$

Proposition 3.2 is vaguely reminiscent of Proposition 3.1 of [FK05] in that we replace a linear sequence by a polynomial object on a simpler space. These two results are not equivalent, however, in that in Proposition 3.2, we end up with a polynomial sequence on a nilmanifold of a connected group where in Proposition 3.1 of [FK05], one obtains a unipotent affine transformation on a torus.

Lemma 3.4. *Let $G_\bullet$ and $G_\bullet^0$ be as given in Proposition 3.2. Then $Z_i(G/\Gamma) = Z_i(G^0/\Gamma^0) = G^0/(G_{i+1}^0\Gamma^0)$ for each $i \in \mathbb{N}$.*

Proof. We do the cases $i = 0$ and $i > 0$ separately. For $i > 0$, we recall from (12) that $Z_i(G/\Gamma) = G/G_{i+1}\Gamma$. Since $G/\Gamma = G^0/\Gamma^0$ by connectedness of G/Γ, and $G_j = G_j^0$ for $j \geqslant 2$, it follows that

$$Z_i(G^0/\Gamma^0) = Z_i(G/\Gamma) = G/G_{i+1}\Gamma = G^0/G_{i+1}^0\Gamma^0.$$

For $i = 0$, we have $Z_i(G/\Gamma) = G/G^0\Gamma = 1 = G^0/G^0\Gamma^0 = Z_i(G^0/\Gamma^0)$. □

4. Homogeneous and inhomogeneous polynomial progressions

The central message of this paper is that homogeneous polynomial progressions satisfy certain linear algebraic properties that make them pliable for our analysis. In this section, we explicitly describe these properties.

Let $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ be an integral polynomial progression. Let V_k be the subspace of $\mathbb{R}[x, y]$ given by

$$V_k = \mathrm{Span}_\mathbb{R}\{(x + P_i(y))^j : 0 \leqslant i \leqslant t, \ 1 \leqslant j \leqslant k\}$$
$$= \mathrm{Span}_\mathbb{R}\left\{\binom{x + P_i(y)}{j} : 0 \leqslant i \leqslant t, \ 1 \leqslant j \leqslant k\right\},$$

and similarly let

$$W_k = \mathrm{Span}_\mathbb{R}\left\{\binom{x + P_i(y)}{k} : 0 \leqslant i \leqslant t\right\}.$$

We also set

$$V^* = \mathrm{Span}_\mathbb{R}\{(Q_0, ..., Q_t) \in \mathbb{R}[u]^{t+1} : \ Q_0(0) = ... = Q_t(0) = 0,$$
$$Q_0(x) + Q_1(x + P_1(y)) + ... + Q_t(x + P_t(y)) = 0\}$$

to be the space of all algebraic relations with zero constant terms satisfied by $\vec{P}$. We recall that an algebraic relation $(Q_0, ..., Q_t)$ is homogeneous if there exists $d \in \mathbb{N}_+$ and $a_0, ..., a_d \in \mathbb{R}$ not all zero such that $Q_i(u) = a_i u^d$ for each $0 \leqslant i \leqslant t$. We call $\vec{P}$ homogeneous if V^* is spanned by homogeneous algebraic relations, and inhomogeneous otherwise.

The concepts of integral polynomial progression and homogeneity, as well as our results in this paper, could likely be extended to multiparameter polynomial progressions of the form

$$(x, \ x + P_1(y_1, ..., y_r), \ ..., \ x + P_t(y_1, ..., y_r));$$

however, we do not pursue this generalisation so as not to obfuscate the notation.

Some important examples of homogeneous progressions include:

(i) linear progressions $(x, \ x + a_1 y, \ ..., \ x + a_t y)$ for distinct nonzero integers $a_1, ..., a_t$, as well as their multiparameter generalizations;

(ii) progressions of algebraic complexity 0, i.e. progressions where the polynomials $P_1, ..., P_t$ are integral and linearly independent;

(iii) progressions of algebraic complexity 1, such as $(x,\ x+y,\ x+y^2,\ x+y+y^2)$, which satisfy no quadratic or higher-order algebraic relation.

Another, less obvious example of a homogeneous progression is $(x,\ x+y,\ x+2y,\ x+y^3)$, already mentioned in the introduction, which only satisfies the homogeneous relation

$$(23) \qquad\qquad x - 2(x+y) + (x+2y) = 0.$$

This progression should be contrasted with $(x,\ x+y,\ x+2y,\ x+y^2)$, which is inhomogeneous because it satisfies both (23) and the inhomogeneous relation

$$(24) \qquad\qquad x^2 + 2x - 2(x+y)^2 + (x+2y)^2 - 2(x+y^2) = 0$$

that cannot be written down as a sum of homogeneous relations. More generally, progressions of the form

$$(x,\ x+y,\ ...,\ x+(t-1)y,\ x+P_t(y))$$

are all inhomogeneous whenever $1 < \deg P_t < t$.

For $k \in \mathbb{N}_+$, we define

$$W_k^c = W_k \cap \sum_{j \neq k} W_j \quad \text{and} \quad W^c = \sum_k W_k^c,$$

as well as the family of quotient spaces

$$W_k' = W_k / W_k^c = W_k \bigg/ \bigg(W_k \cap \sum_{j \neq k} W_j \bigg).$$

Proposition 4.1 (Equivalent conditions for homogeneity). *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression. The following are equivalent:*

(i) *$\vec{P}$ is homogeneous;*

(ii) *W_k^c is trivial for each $k \in \mathbb{N}_+$;*

(iii) *$W_k' = W_k$ for each $k \in \mathbb{N}_+$.*

Proof. The equivalence of (ii) and (iii) follows trivially from the definition of W_k', and we focus on showing the equivalence of (i) and (ii) instead. The inhomogeneity of $\vec{P}$ implies the existence of a nontrivial algebraic relation $(Q_0(u), ..., Q_t(u)) = (\sum_k a_{0k}u^k, ..., \sum_k a_{tk}u^k)$ that is not a sum of homogeneous algebraic relations. What this means is that there exists $k \in \mathbb{N}_+$ for which

$$R(x,y) = a_{0k}x^k + a_{1k}(x+P_1(y))^k + ... + a_{tk}(x+P_t(y))^k \neq 0.$$

Since

$$Q_0(x) + Q_1(x+P_1(y)) + \cdots + Q_t(x+P_t(y)) = 0,$$

we have

$$R(x, y) = -\sum_{j \neq k} \sum_{i=0}^{t} a_{ij}(x + P_i(y))^j \in \sum_{j \neq k} W_j,$$

and so $W_k^c = W_k \cap \sum_{j \neq k} W_j$ is nonempty. Thus (ii) implies (i) by contrapositive. The argument can be reversed, and so (i) and (ii) are in fact equivalent. $\qquad\square$

For homogeneous progression, it is quite straightforward to obtain an upper bound on algebraic complexity.

Proposition 4.2. *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ be a homogeneous polynomial progression. Then $\mathcal{A}_i(\vec{P}) \leqslant t - 1$ for each $0 \leqslant i \leqslant t$.*

This bound is sharp, as evidenced by the example of arithmetic progressions.

Proof. By homogeneity of $\vec{P}$, the only algebraic relations of degree t that $\vec{P}$ could satisfy are of the form

$$(25) \qquad a_0 \binom{x}{t} + a_1 \binom{x + P_i(y)}{t} + \dots + a_t \binom{x + P_t(y)}{t} = 0.$$

A relation (25), together with the formula

$$\binom{x + P_i(y)}{t} = \binom{x}{t} + \binom{x}{t-1} P_i(y) + \binom{x}{t-2} \binom{P_i(y)}{2} + \dots + \binom{P_i(y)}{t},$$

implies

$$a_1 \binom{P_i(y)}{k} + \dots + a_t \binom{P_t(y)}{k} = 0$$

for $1 \leq k \leq t$. From the invertibility of Vandermonde matrix it follows that this is only possible when $a_1 = \dots = a_t = 0$, in which case $a_0 = 0$ as well. Hence $\vec{P}$ satisfies no nontrivial relation of degree t.

$\qquad\square$

Proposition 4.1 implies that homogeneous progressions satisfy

$$(26) \qquad V_k = \bigoplus_{i=1}^{k} W_i = \bigoplus_{i=1}^{k} W_i'.$$

In the inhomogeneous case, we instead have

$$(27) \qquad V_k = \sum_{i=1}^{k} W_i = \left(\bigoplus_{i=1}^{k} W_i' \right) \oplus (W^c \cap V_k)$$

for some nontrivial subspace $W^c \cap V_k$. The nontriviality of this subspace is the main source of difficulty preventing us from generalising Theorem 1.11 to inhomogeneous progressions.

Given the rather abstract nature of the spaces W_k, W_k' and W_k^c, we illustrate their definitions with concrete examples. For the homogeneous progression $(x,\ x + y,\ x + 2y,\ x + y^3)$, we have

$$W_1' = W_1 = \mathrm{Span}_{\mathbb{R}}\{x, y, y^3\} \quad \text{and} \quad W_2' = W_2 = \mathrm{Span}_{\mathbb{R}}\left\{ \binom{x}{2}, xy + \binom{y}{2}, y^2, xy^3 + \binom{y^3}{2} \right\},$$

while for the inhomogeneous progression $(x,\ x + y,\ x + 2y,\ x + y^2)$, we have

$$W_1 = \mathrm{Span}_{\mathbb{R}}\{x, y, y^2\} \quad \text{and} \quad W_2 = \mathrm{Span}_{\mathbb{R}}\left\{ \binom{x}{2}, xy + \binom{y}{2}, y^2, xy^2 + \binom{y^2}{2} \right\}$$

but

$$W_1' = \mathrm{Span}_{\mathbb{R}}\{x, y\}, \quad W_2' = \mathrm{Span}_{\mathbb{R}}\left\{ \binom{x}{2}, xy + \binom{y}{2}, xy^2 + \binom{y^2}{2} \right\} \quad \text{and} \quad W^c = \mathrm{Span}_{\mathbb{R}}\{y^2\}.$$

The nontriviality of W^c for the latter progression is intrinsically related to the algebraic relation (24).

The spaces V_k and W_k are subspaces of $\mathbb{R}[x, y]$. We also need an analogous family of subspaces of $\mathbb{R}^{t+1}$. Let

$$\mathcal{P}_k = \mathrm{Span}_{\mathbb{R}}\left\{ \left(\binom{x}{j}, \binom{x + P_1(y)}{j}, ..., \binom{x + P_t(y)}{j} \right) : x, y \in \mathbb{R},\ 1 \leqslant j \leqslant k \right\}$$

$$= \mathrm{Span}_{\mathbb{R}}\left\{ \left(\binom{x}{k}, \binom{x + P_1(y)}{k}, ..., \binom{x + P_t(y)}{k} \right) : x, y \in \mathbb{R} \right\}$$

$$= \mathrm{Span}_{\mathbb{R}}\left\{ (x^k, (x + P_1(y))^k, ..., (x + P_t(y))^k) : x, y \in \mathbb{R} \right\}.$$

We shall also use the notation

$$\vec{P}^k(x, y) = (x^k, (x + P_1(y))^k, ..., (x + P_t(y))^k) \text{ and } \binom{\vec{P}(x, y)}{k} = \left(\binom{x}{k}, \binom{x + P_1(y)}{k}, ..., \binom{x + P_t(y)}{k} \right).$$

The equivalence of three formulations of $\mathcal{P}_k$ may not be obvious at first glance. The first two formulations are equivalent because if $(a_0, ..., a_t)$ is the coefficient of $\binom{x}{i}\binom{y}{l}$ in $\binom{\vec{P}(x,y)}{j}$, then it will be the coefficient of $\binom{x}{i+k-j}\binom{y}{l}$ in $\binom{\vec{P}(x,y)}{k}$ whenever $j \leqslant k$. The equivalence of the last two formulations follows by induction on k.

Let $t_k = \dim W_k$ and $t_k' = \dim W_k'$ for each $k \in \mathbb{N}$. The spaces W_k and $\mathcal{P}_k$ can be related as follows. Let $\{Q_{k,1}, ..., Q_{k,t_k}\}$ be a basis for W_k. Then

$$\left(\binom{x}{k}, \binom{x + P_1(y)}{k}, ..., \binom{x + P_t(y)}{k} \right) = \sum_{j=1}^{t_k} \vec{v}_{k,j} Q_{k,j}(x, y)$$

for some linearly independent vectors $\vec{v}_{k,1}, ..., \vec{v}_{k,t_k} \in \mathbb{R}^{t+1}$. We let $\tau_k(Q_{k,j}) = \vec{v}_{k,j}$, and extend this map to all of W_k by linearity. This map depends on the choice of the basis for W_k. It is surjective by the definition of $\mathcal{P}_k$ and injective by the linear independence of $\vec{v}_{k,1}, ..., \vec{v}_{k,t_k}$. Hence it is an isomorphism. In particular, Proposition 4.1 implies that $W_k' \cong \mathcal{P}_k$ whenever $\vec{P}$ is homogeneous.

For instance, for $(x, \ x + y, \ x + 2y, \ x + y^3)$, the isomorphisms τ_1 and τ_2 are given by

$$\tau_1(x) = (1,1,1,1), \quad \tau_1(y) = (0,1,2,0), \quad \tau_1(y^3) = (0,0,0,1)$$

and

$$\tau_2\left(\binom{x}{2}\right) = (1,1,1,1), \quad \tau_2\left(xy + \binom{y}{2}\right) = (0,1,2,0),$$

$$\tau_2(y^2) = (0,0,1,0), \quad \tau_2\left(xy^3 + \binom{y^3}{2}\right) = (0,0,0,1).$$

We treat $\mathbb{R}^{t+1}$ as an $\mathbb{R}$-algebra with coordinatewise multiplication $\vec{v}\cdot\vec{w} = (v(0)w(0), ..., v(t)w(t))$ for $\vec{v} = (v(0), ..., v(t))$ and $\vec{w} = (w(0), ..., w(t))$. We similarly let $A \cdot B = \{\vec{a} \cdot \vec{b} : \vec{a} \in A, \vec{b} \in B\}$ be the product set of A and B for any $A, B \subseteq \mathbb{R}^{t+1}$. With these definitions, we observe that $\mathcal{P}_{i+j} \leqslant \mathcal{P}_i \cdot \mathcal{P}_j$, but the converse is in general not true. We also set $\vec{e}_i$ to be the coordinate vector with 1 in the i-th place and 0 elsewhere.

5. Relating Host-Kra complexity to algebraic complexity

Having introduced the notation for the spaces $\mathcal{P}_i$, we are ready to show precisely how determining Host-Kra complexity for homogeneous progressions can be reduced to a certain equidistribution problem on nilmanifolds. We start by defining a group which contains the orbit (18). Groups of this form have previously been defined in [Lei09, GT10, CS12, Kuc20], among others.

Definition 5.1 (Leibman group). *Let $t \in \mathbb{N}_+$ and G be a connected group with a filtration $G_\bullet$ of degree s. For an integral polynomial progression $\vec{P} \in \mathbb{R}[x,y]^{t+1}$, we define the associated Leibman group to be*

$$G^P = \langle g_i^{\vec{v}_i} : g_i \in G_i, \vec{v}_i \in \mathcal{P}_i, 1 \leqslant i \leqslant s \rangle,$$

where $h^{\vec{v}} = (h^{v(0)}, ..., h^{v(t)})$ for any $h \in G$ and $\vec{v} = (v(0), ..., v(t)) \in \mathbb{R}^{t+1}$. We also set $\Gamma^P = G^P \cap G^{t+1}$. If $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$, then we denote

$$g^P(x,y) = (g(x), g(x + P_1(y)), ..., g(x + P_t(y)))$$

and observe that g^P takes values in G^P.

Lemma 5.2. *Let $t \in \mathbb{N}_+$ and G be a connected group with a filtration $G_\bullet$ of degree s. Suppose that $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ is an integral polynomial progression with $\mathcal{A}_i(\vec{P}) = s'$ for some $s' \in \mathbb{N}$. Then G^P contains $1^i \times G_{s'+1} \times 1^{t-i}$.*

Proof. The assumption $\mathcal{A}_i(\vec{P}) = s'$ implies that $(x + P_i(y))^{s'+1}$ is linearly independent from $(x + P_k(y))^{s'+1}$ for $k \neq i$, hence $\mathcal{P}_{s'+1}$ contains $\vec{e}_i$. The Lemma then follows by the definition of G^P. $\qquad\square$

We are now ready to state an infinitary version of the main technical result in the paper. This result constitutes the first part of Theorem 1.15.

Theorem 5.3. *Let $t \in \mathbb{N}_+$ and G be a connected group with filtration $G_\bullet$. Suppose that $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ is irrational and that $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ is a homogeneous polynomial progression. Then g^P is equidistributed on the nilmanifold G^P/Γ^P.*

Importantly, Theorem 5.3 fails for inhomogeneous progressions in that for each inhomogeneous progression $\vec{P}$, we can find a nilmanifold G/Γ, a filtration $G_\bullet$, and an irrational sequence $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ for which the orbit of g^P is contained in a proper subnilmanifold of G^P/Γ^P. An example of this is given in Section 9.

We have all the tools to prove Theorem 5.3 by now. However, we will later need a finitary version of Theorem 5.3, and so instead of proving twice what is essentially the same result, we shall only give the finitary proof later on and deduce Theorem 5.3 from it. For now, however, we can show how the $\mathcal{HK}_i(\vec{P}) \leqslant \mathcal{A}_i(\vec{P})$ part of Theorem 1.11 follows from Theorem 5.3.

Corollary 5.4. *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ be a homogeneous polynomial progression. For any $0 \leqslant i \leqslant t$, we have*

$$\mathcal{HK}_i(\vec{P}) \leqslant \mathcal{A}_i(\vec{P}).$$

The converse inequality will follow from showing that algebraic complexity equals Weyl complexity, and that Weyl complexity is less than or equal to Host-Kra complexity, both of which are done in Section 11.

Proof of Corollary 5.4 using Theorem 5.3. Let $\mathcal{A}_i(\vec{P}) = s$. Let $(X, \mathcal{X}, \mu, T)$ be a totally ergodic system, $f_0, ..., f_t \in L^\infty(\mu)$, and suppose that $\mathbb{E}(f_i|\mathcal{Z}_s) = 0$. By Theorem 1.4, the expression

$$(28) \qquad \lim_{N \to \infty} \mathbb{E}_{n \in [N]} \int_X f_0 \cdot T^{P_1(n)} f_1 \cdot ... \cdot T^{P_t(n)} f_t d\mu$$

remains unchanged if we project the functions $f_0, ..., f_t$ onto the factor $\mathcal{Z}_{s_0}$ for some $s_0 \in \mathbb{N}$. If $s_0 < s$, then $\mathbb{E}(f_i|\mathcal{Z}_{s_0}) = 0$ and the limit (28) is 0, so we can assume that $s_0 \geqslant s$. Since the factor $\mathcal{Z}_{s_0}$ is an inverse limit of s_0-step nilsystems, we can approximate X by totally ergodic nilsystems.

Let $(G/\Gamma, \mathcal{G}/\Gamma, \nu, T_a)$ be a totally ergodic nilsystem, and $G_\bullet$ be the lower central series filtration on G. Using (12), it suffices to show that if $f_0, ..., f_t \in L^\infty(\nu)$ and f_i vanishes on each coset of $G_{s+1}\Gamma$, then

$$\lim_{N \to \infty} \mathbb{E}_{n \in [N]} \int_{G/\Gamma} f_0(b\Gamma) \cdot f_1(a^{P_1(n)} b\Gamma) \cdot ... \cdot f_t(a^{P_t(n)} b\Gamma) d\nu(b\Gamma) = 0.$$

Let $G^0_\bullet$ be the filtration on G^0 given by $G^0_\bullet = G_\bullet \cap G^0$, and let $g_b \in \mathrm{poly}(\mathbb{Z}, G^0_\bullet)$ be the irrational sequence defined in Proposition 3.2 for which $a^n b\Gamma = g_b(n)\Gamma$. The irrationality of g_b, Lemma 3.1 and Theorem 5.3 imply that

$$\lim_{N \to \infty} \mathbb{E}_{n \in [N]} \int_{G/\Gamma} f_0(b\Gamma) \cdot f_1(a^{P_1(n)} b\Gamma) \cdot ... \cdot f_t(a^{P_t(n)} b\Gamma) d\nu(b\Gamma)$$

$$= \int_{G^0/\Gamma^0} \lim_{N \to \infty} \mathbb{E}_{m,n \in [N]} f_0(g_b(m)\Gamma^0) \cdot f_1(g_b(m + P_1(n))\Gamma^0) \cdot ... \cdot f_t(g_b(m + P_t(n))\Gamma^0) d\nu(b\Gamma^0)$$

$$= \int_{(G^0)^P/(\Gamma^0)^P} f_0 \otimes ... \otimes f_t d\nu^P,$$

where $(G^0)^P$ is the Leibman group for $\vec{P}$ and ν^P is the Haar measure on $(G^0)^P/(\Gamma^0)^P$.

The assumption that f_i vanishes on each coset of $G_{s+1}\Gamma$ in G/Γ together with Lemma 3.4 imply that f_i vanishes on each coset of $G_{s+1}^0\Gamma^0$ inside G^0/Γ^0. By Lemma 5.2, the group $(G^0)^P$ contains $H = 1^i \times G_{s+1}^0 \times 1^{t-i}$; therefore

$$\left| \int_{(G^0)^P/(\Gamma^0)^P} f_0 \otimes \ldots \otimes f_t \right| \leqslant \int_{(G^0)^P/H(\Gamma^0)^P} \left| \int_{xH(\Gamma^0)^P} f_0 \otimes \ldots \otimes f_t \right|$$

$$\leqslant \left(\prod_{j \neq i} \|f_j\|_\infty \right) \int_{(G^0)^P/H(\Gamma^0)^P} \left| \int_{x_i G_{s+1}^0 \Gamma^0} f_i \right| = 0,$$

implying that $\mathcal{Z}_s$ is characteristic for the weak convergence of $\vec{P}$ at i. $\qquad\square$

Corollary 5.4 implies that if a progression $\vec{P}$ satisfies $\mathcal{A}_i(\vec{P}) = s$, then $\mathcal{Z}_s$ is characteristic for the weak or L^2 convergence of $\vec{P}$ at i for any totally ergodic system. We now prove Corollary 1.12, which extends this result to ergodic systems with a slight modification in the $s = 0$ case. The proof is almost identical to the proof of Proposition 4.1 in [Fra08].

Proof of Corollary 1.12. Let $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be a homogeneous progression with $\mathcal{A}_i(\vec{P}) = s$ and $(X, \mathcal{X}, \mu, T)$ be ergodic. By Theorem 1.4, there exists a Host-Kra factor that is characteristic for the weak and L^2 convergence of $\vec{P}$. Since each Host-Kra factor is an inverse limit of nilsequences, we can approximate X by an ergodic nilsystem $(G/\Gamma, \mathcal{G}/\Gamma, \nu, T_a)$. The compactness of G/Γ and the assumption that G is generated by the connected component G_0 and a imply that $a^r \in G^0$ for some $r \in \mathbb{N}_+$; and hence

$$(29) \qquad \mathbb{E}_{n \in [rN]} \prod_{i=1}^{t} T_a^{P_i(n)} f_i = \mathbb{E}_{j \in [r]} \mathbb{E}_{n \in [N]} \prod_{i=1}^{t} T_a^{P_i(r(n-1)+j)} f_i$$

$$= \mathbb{E}_{j \in [r]} \mathbb{E}_{n \in [N]} \prod_{i=1}^{t} (T_a^r)^{\tilde{P}_{i,j}(n)} (T_a^{P_i(j)} f_i),$$

where $\tilde{P}_{i,j}(n) = \frac{P_i(r(n-1)+j)-P_i(j)}{r}$. A tedious linear algebraic argument shows that for any $0 \leqslant j < r$, the progression $\vec{P}$ is homogeneous if and only if

$$\vec{\tilde{P}}_j(x,y) = (x,\ x + \tilde{P}_{1,j}(y),\ \ldots,\ x + \tilde{P}_{t,j}(y))$$

is, and that there is a one-to-one correspondence between algebraic relations that $\vec{P}$ and $\vec{\tilde{P}}_j$ satisfy. In particular, $\vec{\tilde{P}}_j$ is also homogeneous and $\mathcal{A}_i(\vec{\tilde{P}}_j) = s$.

If $s > 0$, suppose that $\mathbb{E}(f_i|\mathcal{Z}_s(T_a)) = 0$. Then the equality $\mathcal{Z}_s(T_a) = \mathcal{Z}_s(T_a^r)$ and the T_a-invariance of $\mathcal{Z}_s$ imply that $\mathbb{E}(T_a^{P_i(j)} f_i|\mathcal{Z}_s(T_a^r)) = 0$. We deduce from Corollary 5.4 and the total ergodicity of T_a^r on each connected components of G/Γ that the expression in (29) converges to 0 as $N \to \infty$.

If $s = 0$, suppose that $\mathbb{E}(f_i|\mathcal{K}_{rat}(T_a)) = 0$. The total ergodicity of T_a^r implies that $\mathcal{K}_{rat}(T_a) = \mathcal{Z}_0(T_a^r)$, and so $\mathbb{E}(T_a^{P_i(j)} f_i|\mathcal{Z}_0(T_a^r)) = 0$. Again, it follows from Corollary 5.4

and the total ergodicity of T_a^r on each connected components of G/Γ that the expression in (29) converges to 0 as $N \to \infty$.

$\square$

Theorem 5.3 also allows us to prove the second part of Corollary 1.13.

Proof of Corollary 1.13(ii). Let $(X, \mathcal{X}, \mu, T)$ be a totally ergodic system, and suppose that $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ is an integral progression with algebraic complexity at most 1. This implies that $\vec{P}$ is homogeneous since each inhomogeneous algebraic relation must have degree at least 2. For each $0 \leqslant i \leqslant t$, let $P_i(y) = \sum_{j=1}^d a_{i,j} Q_j(y)$ and $L_i(y_1, ...y_d) = \sum_{j=1}^d a_{i,j} y_j$ for some $a_{i,j} \in \mathbb{Z}$ and integral polynomials $Q_1, ..., Q_d$. Letting

$$\vec{L}(x, y_1, ..., y_d) = (x, \ x + L_1(y_1, ..., y_d), \ ..., \ x + L_t(y_1, ..., y_d)),$$

we observe that $\vec{P}(x, y) = \vec{L}(x, Q_1(y), ..., Q_d(y))$. It follows that $\vec{L}$ also has an algebraic complexity at most 1, since each algebraic relation of degree $(j_0, ..., j_t)$ between terms of $\vec{L}$ would immediately imply an algebraic relation of the same degree between terms of $\vec{P}$ after substituting $y_i = Q_i(y)$.

Using the same argument as in the proof of Corollary 5.4, we reduce the question of understanding

$$(30) \qquad \lim_{N \to \infty} \mathbb{E}_{n \in [N]} \int_X \prod_{i=0}^t T^{P_i(n)} f_i \, d\mu$$

to understanding

$$(31) \qquad \lim_{N \to \infty} \mathbb{E}_{x,y \in [N]} F(g^P(x, y))$$

for each essentially bounded function $F : (G/\Gamma)^{t+1} \to \mathbb{C}$ and an irrational sequence $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ for some filtration $G_\bullet$ on G. Following the same method to analyse

$$(32) \qquad \lim_{N \to \infty} \mathbb{E}_{y_1, ..., y_d \in [N]} \int_X \prod_{i=0}^t T^{L_i(y_1, ..., y_d)} f_i \, d\mu,$$

we deduce that understanding (32) comes down to estimating

$$(33) \qquad \lim_{N \to \infty} \mathbb{E}_{x, y_1, ..., y_d \in [N]} F(g^L(x, y_1, ..., y_d)),$$

where

$$g^L(x, y_1, ..., y_d) = (g(x), g(x + L_1(y_1, ..., y_d)), ..., g(x + L_t(y_1, ..., y_d))).$$

By Theorem 5.3, the limit in (31) equals $\int_{G^P/\Gamma^P} F$; by Theorem 11 of [GT10], the limit in (33) is $\int_{G^L/\Gamma^L} F$ for some subgroup $G^L \leqslant G^{t+1}$. From the fact that $\max_i \mathcal{A}_i(\vec{P}) \leqslant 1$ we deduce that $G^P = \langle h_1^{\vec{v}_1}, G_2^{t+1} : h_1 \in G_1, \vec{v}_1 \in \mathcal{P}_1 \rangle$; similarly, the construction of the group G^L in [GT10] and the fact that $\vec{L}$ has algebraic complexity at most 1 reveals that $G^L = \langle h_1^{\vec{v}_1}, G_2^{t+1} : h_1 \in G_1, \vec{v}_1 \in \mathcal{L}_1 \rangle$, where

$$\mathcal{L}_1 = \mathrm{Span}_{\mathbb{R}} \{(x, x + L_1(y_1, ..., y_d), ..., x + L_t(y_1, ..., y_d)) : x, y_1, ..., y_d \in \mathbb{R}\}.$$

We observe that $\mathcal{P}_1 = \mathcal{L}_1$; from this it follows that $G^P = G^L$, and so the limits in (31) and (33) are equal. This implies that (30) and (32) equal as well.

$\square$

6. Finitary nilmanifold theory

Before we can prove a finitary version of Theorem 5.3, we need to introduce necessary finitary concepts required for this task. Most concepts and definitions in this and next section are taken from [GT10, GT12, CS12]. Throughout this section, we assume that G is connected, and that each nilmanifold G/Γ comes with a filtration $G_\bullet$ and a Mal'cev basis χ adapted to $G_\bullet$. We call a nilmanifold endowed with filtration and a Mal'cev basis *filtered*. A *Mal'cev basis* is a basis for the Lie algebra of G with some special properties; since we do not explicitly work with the notion of Mal'cev basis or its rationality in the paper, we refer the reader to [GT12] for definitions of these concepts. What matters for us is that each Mal'cev basis induces a diffemomorphism $\psi : G \to \mathbb{R}^m$, called *Mal'cev coordinate map*, which satisfies the following properties:

(i) $\psi(\Gamma) = \mathbb{Z}^m$;

(ii) $\psi(G_i) = \{0\}^{m-m_i} \times \mathbb{R}^{m_i}$, where $m_i = \dim G_i$.

Thus, ψ provides a natural coordinate system on G that respects the filtration $G_\bullet$ and the lattice Γ. Similarly to ψ, we define maps $\psi_i : G_i \to \mathbb{R}^{m_i - m_{i+1}}$ by assigning to each element of G_i its Mal'cev coordinates indexed by $m - m_i + 1, ..., m - m_{i+1}$. With this definition, we have $\psi_i(x) = 0$ if and only if $x \in G_{i+1}$, and $\psi_i(x) \in \mathbb{Z}^{m_i - m_{i+1}}$ if and only if $x \in \Gamma_i$.

Definition 6.1 (Complexity of nilmanifolds). *A filtered nilmanifold G/Γ has complexity M if the degree s of the filtration $G_\bullet$, the dimension m of the group G, and the rationality of the Mal'cev basis χ are all bounded by M.*

We remark that complexity of nilmanifolds has nothing to do with the four notions of complexity of polynomial progressions that we examine. Neither does complexity of nilsequences defined below.

Definition 6.2 (Nilsequences). *A function $f : \mathbb{Z} \to \mathbb{C}$ is a* nilsequence *of degree s and complexity M if $f(n) = F(g(n)\Gamma)$, where $F : G/\Gamma \to \mathbb{R}$ is an M-Lipschitz function on a filtered nilmanifold G/Γ of degree s and complexity M, and $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$.*

Definition 6.3 (Quantitative equidistribution). *Let $D \in \mathbb{N}_+$ and $\delta > 0$. A sequence $g \in \mathrm{poly}(\mathbb{Z}^D, G)$ is (δ, N)-equidistributed on G/Γ if*

$$\left| \mathbb{E}_{n \in [N]^D} F(g(n)\Gamma) - \int_{G/\Gamma} F \right| \leqslant \delta ||F||_{\mathrm{Lip}}$$

for all Lipschitz functions $F : G/\Gamma \to \mathbb{C}$, where $||F||_{\mathrm{Lip}}$ is the Lipschitz norm on F with respect to a metric defined in [GT12].

It has been shown in Theorem 2.5 that equidistribution is related to horizontal characters. Given the Mal'cev coordinate map $\psi : G \to \mathbb{R}^m$, each horizontal character can be written in the form $\eta(x) = k \cdot \psi(x)$ for some $k \in \mathbb{Z}^m$. We call $|\eta| = |k| = |k_1| + ... + |k_m|$ the *modulus* of η. Similarly, each i-th level character $\eta_i : G_i \to \mathbb{R}$ is of the form $\eta_i(x) = k \cdot \psi_i(x)$ for some $k \in \mathbb{Z}^{m_i - m_{i+1}}$, and we define its modulus to be $|\eta_i| = |k| = |k_1| + ... + |k_{m_i - m_{i+1}}|$.

We shall also need to quantify the notion of polynomials that are "almost constant" mod $\mathbb{Z}$, using a definition from [GT12]. In what follows, $||x||_{\mathbb{R}/\mathbb{Z}} = \min\{|x - n| : n \in \mathbb{Z}\}$ is the *circle norm* of $x \in \mathbb{R}$.

Definition 6.4 (Smoothness norm). *Let*

$$Q(n_1, ..., n_D) = \sum_{i=0}^{d} \sum_{i_1+...+i_D=i} a_{i_1,...,i_D} \binom{n}{i_1} \cdots \binom{n}{i_d}$$

be a polynomial in $\mathbb{R}[n_1, ..., n_D]$. *For* $N \in \mathbb{N}_+$, *we define the* smoothness norm *of* Q *to be*

$$||Q||_{C^\infty[N]} = \max\{N^{i_1+...+i_D}||a_{i_1,...,i_D}||_{\mathbb{R}/\mathbb{Z}} : i_1, ..., i_D \in \mathbb{N}, \ 1 \leq i_1 + \cdots + i_D \leq d\}.$$

In particular, $||Q||_{C^\infty[N]}$ is bounded from above as $N \to \infty$ if and only if Q is constant mod $\mathbb{Z}$.

With these definitions, we are ready to state a quantitative version of Theorem 2.5

Theorem 6.5 (Quantitative Leibman's equidistribution theorem, Theorem 2.9 of [GT12]). *Let* $\delta > 0$, $M \geqslant 2$ *and* $D, N \in \mathbb{N}_+$ *with* $D \leqslant M$. *Let* G/Γ *be a filtered nilmanifold of complexity* M *and* $g \in \mathrm{poly}(\mathbb{Z}^D, G_\bullet)$. *Then there exists* $C_M > 0$ *such that at least one of the following is true:*

(i) g *is* (δ, N)-*equidistributed in* G/Γ;

(ii) *there exists a nontrivial horizontal character* η *of modulus* $|\eta| \ll \delta^{-C_M}$ *for which* $||\eta \circ g||_{C^\infty[N]} \ll \delta^{-C_M}$.

We now need to quantify the notion of irrationality.

Definition 6.6 (Quantitative irrationality). *Let* G/Γ *be a filtered nilmanifold of degree* s, *and suppose* $A, N > 0$. *An element* $g_i \in G_i$ *is* (A, N)-*irrational if for every nontrivial* i-*th level character* $\eta : G_i \to \mathbb{R}$ *of modulus* $|\eta| \leqslant A$, *we have* $||\eta(g_i)||_{\mathbb{R}/\mathbb{Z}} \geqslant A/N^i$. *It is* A-*irrational if for every nontrivial* i-*th level character* $\eta : G_i \to \mathbb{R}$ *of modulus* $|\eta| \leqslant A$, *we have* $\eta \circ g_i \notin \mathbb{Z}$. *We say that a sequence* $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ *is* (A, N)-*irrational (respectively* A-*irrational) if* g_i *is* (A, N)-*irrational (respectively* A-*irrational) for each* $1 \leqslant i \leqslant s$. *Similarly, we say that the nilsequence* $n \mapsto F(g(n)\Gamma)$ *is* (A, N)- *or* A-*irrational if the polynomial sequence* g *is.*

Clearly, (A, N)-irrationality is stronger than A-rationality, but for some of our applications the latter notion will be sufficient.

We are now ready to state the finitary version of Theorem 5.3, which is the main technical result of this paper, and derive Theorem 5.3 from it.

Theorem 6.7. *Let* $t \in \mathbb{N}_+$ *and* $A, M, N \geqslant 2$. *Let* G/Γ *be a filtered nilmanifold of complexity* M. *Suppose that* $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ *is* (A, N)-*irrational,* $F : (G/\Gamma)^{t+1} \to \mathbb{C}$ *is* M-*Lipschitz, and* $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ *is a homogeneous polynomial progression. Then*

$$\mathbb{E}_{x,y \in [N]} F(g^P(x, y)\Gamma^{t+1}) = \int_{G^P/\Gamma^P} F + O_M(A^{-c_M})$$

for some $c_M > 0$.

Proof of Theorem 5.3 using Theorem 6.7. Let $F : (G/\Gamma)^{t+1} \to \mathbb{R}$ be a continuous function. By Stone-Weierstrass theorem, Lipschitz functions on a compact set form a dense subset of the algebra of continuous functions. Approximating F by a sequence of Lipschitz functions

if necessary, we can assume without loss of generality that F is Lipschitz. We let M be the maximum of the complexity of G/Γ and the Lipschitz norm of F.

Let $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ be an irrational sequence. For each $N \in \mathbb{N}_+$, we let A_N be the maximal real number A for which g is (A_N, N)-irrational. We claim that $A_N \to \infty$ as $N \to \infty$. If not, then there exists some number $A > 0$ and an index $i \in \mathbb{N}_+$ with the property that g_i is not (A, N)-irrational for all $N \in \mathbb{N}_+$. We fix this i. It follows that there exists a sequence of nontrivial i-th level characters $\eta_N : G_i \to \mathbb{R}$ of modulus at most A such that $||\eta_N(g_i)||_{\mathbb{R}/\mathbb{Z}} < A/N^i$. Since there are only finitely many i-th level characters of modulus bounded by A, we conclude that there exists a nontrivial i-th level character η of modulus at most A such that $||\eta(g_i)||_{\mathbb{R}/\mathbb{Z}} < A/N^i$ for all $N \in \mathbb{N}_+$. Taking $N \to \infty$, we see that $\eta(g_i) \in \mathbb{Z}$, contradicting the irrationality of g_i.

It therefore follows from Theorem 6.7 that

$$\mathbb{E}_{x,y\in[N]}F(g^P(x,y)\Gamma^{t+1}) = \int_{G^P/\Gamma^P} F + O_M(A_N^{-c_M})$$

Since M is constant, letting $N \to \infty$ sends the error term to 0, implying that g^P is equidistributed on G^P/Γ^P as claimed. $\qquad\square$

7. Reducing true complexity to an equidistribution question

In Sections 3-6, we have shown how the question of determining Host-Kra complexity for homogeneous progressions can be reduced to showing that g^P is equidistributed on G^P/Γ^P. Determining true complexity for homogeneous progression comes down to the exact same equidistribution question. All the arguments in this section can be viewed as finitary analogues of arguments in previous sections.

Since we are now primarily concerned with functions from $\mathbb{Z}/N\mathbb{Z}$ to $\mathbb{C}$, we shall need an N-periodic version of certain previously defined concepts. In this section, N is always a prime, and the group G is connected. A function $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ is called *1-bounded* whenever $||f||_\infty \leqslant 1$.

Definition 7.1 (Periodic sequences). *Let $G_\bullet$ be a filtration on G. A sequence $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ is N-periodic if $g(n+N)g(n)^{-1} \in \Gamma$ for each $n \in \mathbb{Z}$, and it is periodic if it is N-periodic for some $N > 0$. A nilsequence $n \mapsto F(g(n)\Gamma)$ is N-periodic (resp. periodic) if g is.*

Given a homogeneous polynomial progression $\vec{P} \in \mathbb{R}[x,y]^{t+1}$, we want to show that $\mathcal{A}_i(\vec{P}) = \mathcal{T}_i(\vec{P})$ for each $0 \leqslant i \leqslant t$. The forward inequality is straightforward to derive (see Theorem 1.13 in [Kuc20]); it is the reverse inequality that poses a challenge. We thus want to prove the following.

Theorem 7.2. *Let $t \in \mathbb{N}_+$, $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be a homogeneous polynomial progression, $0 \leqslant i \leqslant t$, and suppose that $\mathcal{A}_i(\vec{P}) = s$. For every $\epsilon > 0$, there exist $\delta > 0$ and $N_0 \in \mathbb{N}$ such that for all primes $N > N_0$ and all 1-bounded functions $f_0, ..., f_t : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, we have*

$$|\mathbb{E}_{x,y\in\mathbb{Z}/N\mathbb{Z}}f_0(x)f_1(x+P_1(y))...f_t(x+P_t(y))| < \epsilon$$

whenever $||f_i||_{U^{s+1}} < \delta$.

We know that each progression is controlled by *some* Gowers norm. The result below plays the same role in deriving Theorem 7.2 as Theorem 1.4 plays in the proof of Corollary 5.4.

Proposition 7.3 (Proposition 2.2 of [Pel19])**.** *Let $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression. There exists $s \in \mathbb{N}_+$ with the following property: for every $\epsilon > 0$, there exist $\delta > 0$ and $N_0 \in \mathbb{N}$ such that for all primes $N > N_0$ and all 1-bounded functions $f_0, ..., f_t : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, we have*

$$|\mathbb{E}_{x,y \in \mathbb{Z}/N\mathbb{Z}} f_0(x) f_1(x + P_1(y)) ... f_t(x + P_t(y))| < \epsilon$$

whenever $||f_i||_{U^{s+1}} < \delta$ for some $0 \leqslant i \leqslant t$.

Next, we want to perform a finitary analogue of the approximation-by-nilsystems argument. This can be achieved with the help of a periodic version of celebrated arithmetic regularity lemma from [GT10] in which the same polynomial sequence g is used in the decomposition of several functions.

Lemma 7.4 (Lemma 2.13 of [Kuc20])**.** *Let $s, t \in \mathbb{N}+$, $\epsilon > 0$, and $\mathcal{F} : \mathbb{R}_+ \to \mathbb{R}_+$ be a growth function. There exists $M = O_{\epsilon, \mathcal{F}}(1)$, a filtered nilmanifold G/Γ of degree s and complexity at most M, and an N-periodic, $\mathcal{F}(M)$-irrational sequence $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ satisfying $g(0) = 1$ such that for all 1-bounded functions $f_0, ..., f_t : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, there exist decompositions*

$$f_i = f_{i,nil} + f_{i,sml} + f_{i,unf}$$

where

(i) $f_{i,nil}(n) = F_i(g(n)\Gamma)$ for M-Lipschitz function $F_i : G/\Gamma \to \mathbb{C}$,

(ii) $||f_{i,sml}||_2 \leqslant \epsilon$,

(iii) $||f_{i,unf}||_{U^{s+1}} \leqslant \frac{1}{\mathcal{F}(M)}$,

(iv) the functions $f_{i,nil}$, $f_{i,sml}$ and $f_{i,unf}$ are 4-bounded,

The last piece that we need is a finitary, periodic version of Theorem 6.7.

Proposition 7.5. *Let $t \in \mathbb{N}_+$ and $A, M, N \geqslant 2$. Let G/Γ be a filtered nilmanifold and complexity M. Suppose that $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ is an A-irrational, N-periodic polynomial sequence, $F : (G/\Gamma)^{t+1} \to \mathbb{C}$ is M-Lipschitz and 1-bounded, and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ is a homogeneous polynomial progression. Then*

$$\mathbb{E}_{x,y \in \mathbb{Z}/N\mathbb{Z}} F(g^P(x,y)\Gamma^{t+1}) = \int_{G^P/\Gamma^P} F + O_M(A^{-c_M})$$

for some $c_M > 0$.

Proof of Proposition 7.5 using Theorem 6.7. Let $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ be A-irrational and N-periodic. We claim that g is (A, Nk)-irrational for all sufficiently large $k \in \mathbb{N}_+$. If not, then there exists $1 \leqslant i \leqslant s$ such that for each $k \in \mathbb{N}_+$ there exists an i-th level character $\eta_{i,k} : G_i \to \mathbb{R}$ of complexity at most A satisfying $||\eta_{i,k}(g_i)||_{\mathbb{R}/\mathbb{Z}} < A/(Nk)^i$. The N-periodicity of g_i implies that $g_i^{N^i} \in \Gamma_i \mod G_{i+1}^\nabla$ (Lemma 5.3 of [CS12]); hence $\eta_{i,k}(g_i) \in \frac{1}{N^i}\mathbb{Z}$. Thus, $\eta_{i,k}(g_i) \in \mathbb{Z}$ whenever $k^i > A$. In particular, since we can take k arbitrarily large, there exists a nontrivial i-th level character $\eta_{i,k}$ of complexity at most A for which $\eta_{i,k}(g_i) \in \mathbb{Z}$, contradicting the A-irrationality of g. Hence g is (A, Nk)-irrational for all sufficiently large $k \in \mathbb{N}_+$.

Applying Theorem 6.7, we deduce that

$$\mathbb{E}_{x,y\in\mathbb{Z}/N\mathbb{Z}}F(g^P(x,y)\Gamma^{t+1}) = \mathbb{E}_{x,y\in[Nk]}F(g^P(x,y)\Gamma^{t+1}) + O(1/k)$$
$$= \int_{G^P/\Gamma^P} F + O_M(A^{-c_M}) + O(1/k)$$

for all sufficiently large $k \in \mathbb{N}_+$. Taking $k \to \infty$ finishes the proof. $\qquad\square$

Theorem 7.2 is a special case of Theorem 8.1 of [Kuc20], the proof of which is analogous to the derivation of Corollary 5.4 from Theorem 5.3. Here, we only sketch the steps taken in the derivation of Theorem 8.1 of [Kuc20], and we refer the reader to [Kuc20] for all the details. First, we use Proposition 7.3 and Lemma 7.4 to replace the functions $f_0, ..., f_t$ by irrational, periodic nilsequences. Second, we use Proposition 7.5 to approximate the sum by an integral of some Lipschitz function F over G^P/Γ^P. Third, we use the fact that $\mathcal{A}_i(\vec{P}) = s$ to conclude that $1^i \times G_{s+1} \times 1^{t_i}$ is a subgroup of G^P. Fourth, we use disintegration theorem to bound $\int_{G^P/\Gamma^P}$ by averages of some Lipschitz function F_i over cosets of $G_{s+1}\Gamma$. Fifth, we use the assumption that f_i has a small U^{s+1} norm to conclude that averages of F_i over cosets of $G_{s+1}\Gamma$ are small. From this follows the smallness of

$$\mathbb{E}_{x,y\in\mathbb{Z}/N\mathbb{Z}}f_0(x)f_1(x + P_1(y))\cdots f_t(x + P_t(y)).$$

The proof of Theorem 8.1 of [Kuc20] makes this argument precise and illustrates how all the error quantities are taken care of.

Finally, Proposition 7.5 together with Theorem 9.1 of [Kuc20] imply part (i) of Corollary 1.13.

8. The proof of Theorem 6.7

To complete the proofs of Corollary 5.4 and Theorem 7.2, it remains to derive Theorem 6.7. Before we prove Theorem 6.7 for an arbitrary homogeneous progression, we want to deduce the theorem in the special case of $\vec{P} = (x, \ x + y, \ x + 2y, \ x + y^3)$. This will help illustrate the method, and we will later compare this progression with $(x, \ x+y, \ x+2y, \ x+y^2)$ to see what is failing in the inhomogeneous case. The method is an adaptation of the proof of Theorem 1.11 from [GT10], however the linear algebraic component coming from the fact that we are dealing with polynomial progressions is much more involved. The method used here is somewhat similar to the methods used in [Kuc20]; here, however, we perform downward induction on the degree of subgroups G_i whereas in [Kuc20], we induct downwardly on the degree of monomials in $\eta \circ g^P$.

Proposition 8.1. *Let $A, M, N \geqslant 2$. Let G/Γ be a filtered nilmanifold of degree 2 and complexity M. Suppose that $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ is an (A, N)-irrational sequence satisfying $g(0) = 1$, $F : (G/\Gamma)^{t+1} \to \mathbb{C}$ is M-Lipschitz, and $\vec{P} = (x, \ x + y, \ x + 2y, \ x + y^3)$. Then*

$$\mathbb{E}_{x,y\in[N]}F(g^P(x,y)\Gamma^4) = \int_{G^P/\Gamma^P} F + O_M(A^{-c_M})$$

for some $c_M > 0$.

The assumption that G has a filtration of degree 2 is made to simplify the exposition, and because all the difficulties that emerge in higher-step cases are already present here.

We shall need the following lemma.

Lemma 8.2. *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be a homogeneous polynomial progression, $\epsilon > 0$, and $s, N \in \mathbb{N}_+$. Let $W_i \leqslant \mathbb{R}[x,y]$ be as defined in Section 4, and for each $1 \leqslant i \leqslant s$, let $Q_{i,1}, ..., Q_{i,t_i}$ be a basis for W_i composed of integral polynomials. Suppose that a_{ij} are real numbers such that the polynomial*

$$Q(x,y) = \sum_{i=1}^{s} \sum_{j=1}^{t_i} a_{ij} Q_{i,j}(x,y) = \sum_{k,l} c_{kl} \binom{x}{k} \binom{y}{l}$$

satisfies $||Q||_{C^\infty[N]} \leqslant \epsilon$. Then there exists a positive integer $q = O(1)$ with the property that $||q a_{sj}||_{\mathbb{R}/\mathbb{Z}} \ll \epsilon N^{-s}$ for all $1 \leqslant j \leqslant t_s$.

Proof. For $k \in \mathbb{N}_+$, we let W_k, V_k be as in Section 4. We also define

$$\tilde{W}_k = \mathrm{Span}_{\mathbb{R}}\{(x + P_i(y))^k : 0 \leqslant i \leqslant t\} \quad \text{and} \quad U_k = \mathrm{Span}_{\mathbb{R}}\left\{ \binom{x}{i}\binom{y}{j} : i + j < k \right\}.$$

The homogeneity of $\vec{P}$ implies that $W_k \cong V_k/V_{k-1} \cong \tilde{W}_k$, and we want to show that $\dim W_k/U_k = \dim W_k = t_k$. We observe that $W_k/U_k \cong V_k/U_k V_{k-1} \cong \tilde{W}_k/U_k \cong \tilde{W}_k$, where the last isomorphism follows from the fact no polynomial in $\tilde{W}_k$ has a nonzero monomial of degree less than k. The claim $\dim W_k/U_k = t_k$ follows.

Let $\tilde{Q}(x,y) = \sum_{k+l \geqslant s} c_{kl} \binom{x}{k}\binom{y}{l}$. Thus, $\tilde{Q} = Q \bmod U_s$, and it satisfies $||\tilde{Q}||_{C^\infty[N]} \leqslant \epsilon$. Setting $Q_{i,j}(x,y) = \sum_{k,l} b_{klij} \binom{x}{k}\binom{y}{l}$, we similarly let $\tilde{Q}_{i,j}(x,y) = \sum_{k+l \geqslant s} b_{klij} \binom{x}{k}\binom{y}{l}$. We deduce from $\dim W_k/U_k = t_k = \dim W_k$ that $\tilde{Q}_{s,1}, ..., \tilde{Q}_{s,t_s}$ are linearly independent.

From the definitions of Q and b_{klij} it follows that $c_{kl} = \sum_{i,j} b_{klij} a_{ij}$, and that $||c_{kl}||_{\mathbb{R}/\mathbb{Z}} \leqslant \epsilon N^{-(k+l)} \leqslant \epsilon N^{-s}$ whenever $k + l \geqslant s$.

Let u be the number of pairs (k,l) with $k + l \geqslant s$ for which $c_{kl} \neq 0$. The fact that $\dim W_s/U_s = t_s$ implies that $u \geqslant t_s$. Indexing these pairs as $(k_1, l_1), ..., (k_u, l_u)$ in some arbitrary fashion, we obtain an $u \times s$ matrix $B = (b_{k_r l_r ij})_r$ as well as a t_s-dimensional column vector $a = (a_{sj})_j$ and a u-dimensional column vector $c = (c_{k_r l_r})_r$ such that $Ba = c$. The linear independence of $\tilde{Q}_{s,1}, ..., \tilde{Q}_{s,t_s}$ implies that there exists an invertible $t_s \times t_s$ submatrix $\tilde{B}$ of B and a t_s-dimensional column vector $\tilde{c}$ such that $\tilde{B}a = \tilde{c}$. Since the entries of $\tilde{B}$ are integers of size $O(1)$, the entries of $\tilde{B}^{-1}$ are rational numbers of height $O(1)$. Therefore, there exists a positive integer $q = O(1)$ for which the entries of the matrix $q\tilde{B}^{-1}$ are integers of size $O(1)$. The equality $a = \tilde{B}^{-1}\tilde{c}$ and the condition $||c_{kl}||_{\mathbb{R}/\mathbb{Z}} \leqslant \epsilon N^{-s}$ whenever $k + l \geqslant s$ imply that $||q a_{sj}||_{\mathbb{R}/\mathbb{Z}} \ll \epsilon N^{-s}$ for $1 \leqslant j \leqslant t_s$, as claimed. $\qquad \square$

Proof of Proposition 8.1. Let $\vec{P} = (x, \ x+y, \ x+2y, \ x+y^3)$. We set

$$\vec{v}_1 = (1,1,1,1), \quad \vec{v}_2 = (0,1,2,0), \quad \vec{v}_3 = (0,0,0,1) \quad \text{and} \quad \vec{v}_4 = (0,0,1,0)$$

and observe that

$$\vec{P}(x,y) = \vec{v}_1 x + \vec{v}_2 y + \vec{v}_3 y^3$$

$$\binom{\vec{P}(x,y)}{2} = \vec{v}_1 \binom{x}{2} + \vec{v}_2 \left(xy + \binom{y}{2} \right) + \vec{v}_3 \left(xy^3 + \binom{y^3}{2} \right) + \vec{v}_4 y^2.$$

Thus, we have

$$\mathcal{P}_1 = \mathrm{Span}_{\mathbb{R}}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} \quad \text{and} \quad \mathcal{P}_2 = \mathcal{P}_3 = ... = \mathrm{Span}_{\mathbb{R}}\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\} = \mathbb{R}^4$$

as well as

$$G^P = G^{\vec{v}_1} G^{\vec{v}_2} G^{\vec{v}_3} G_2^4,$$

where $H^{\vec{w}} = \langle h^{\vec{w}} : h \in H \rangle$.

We shall prove Proposition 8.1 by applying Theorem 6.5. Suppose that g^P is not $(c_M A^{-C_M}, N)$-equidistributed on G^P/Γ^P for some constants $0 < c_M < 1 < C_M$. By Theorem 6.5, there exists a nontrivial horizontal character $\eta : G^P \to \mathbb{R}$ of modulus at most cA, for which $||\eta \circ g^P||_{C^\infty[N]} \leqslant cA$ for some constant $c > 0$ that depends on c_M and C_M. The constant C_M is chosen in such a way as to match the exponents in the case (ii) of Theorem 6.5. We however have control over how we choose the constant c_M, and we shall pick it small enough to show that g^P not being $(c_M A^{-C_M}, N)$-equidistributed contradicts the (A, N)-irrationality of g.

Rewriting the expression for $\eta \circ g^P$, we see that

$$\eta \circ g^P(x,y) = \eta(g_1^{\vec{v}_1})x + \eta(g_1^{\vec{v}_2})y + \eta(g_1^{\vec{v}_3})y^3$$
$$+ \eta(g_2^{\vec{v}_1})\binom{x}{2} + \eta(g_2^{\vec{v}_2})\left(xy + \binom{y}{2} \right) + \eta(g_2^{\vec{v}_3})\left(xy^3 + \binom{y^3}{2} \right) + \vec{v}_4 y^2.$$

Applying Lemma 8.2 and the assumption $||\eta \circ g^P||_{C^\infty[N]} \leqslant cA$, and choosing c_M in such a way that $c > 0$ is sufficiently small, we deduce that there exists a positive integer $q = O(1)$ such that $||q\eta(g_i^{\vec{v}_j})||_{\mathbb{R}/\mathbb{Z}} < AN^{-i}$ for all pairs

$$(i,j) \in \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (2,4)\}.$$

We aim to show that η is trivial by showing that it vanishes on all of G^P. First, we want to show that η vanishes on G_2^4. Suppose that $\eta|_{G_2^4} \neq 0$, and define $\xi_{2,1} : G_2 \to \mathbb{R}$ by $\xi_{2,1}(h_2) = q\eta(h_2^{(1,1,1,1)})$. We claim that $\xi_{2,1}$ is a 2-nd level character. To prove this, we need to show that $\xi_{2,1}$ is a continuous group homomorphism, it vanishes on G_3, it sends (Γ_2) to $\mathbb{Z}$, and it vanishes on $[G_1, G_1]$. The first statement follows from the fact that η is a continuous group homomorphism, the second is true since G_3 is trivial, and the third follows from the fact that $q \in \mathbb{Z}$, $\eta(\Gamma^P) \leqslant \mathbb{Z}$ and $(1,1,1,1) \in \mathbb{Z}^4$. To see the last statement, we note that for any $h_1, h_1' \in G_1$, we have

$$[h_1^{\vec{v}_1}, h_1'^{\vec{v}_1}] = [h_1, h_1']^{\vec{v}_1}.$$

Since $h_1^{\vec{v}_1}, h_1'^{\vec{v}_1}$ are both elements of G^P, we have

$$\xi_{2,1}([h_1, h_1']) = \eta([h_1, h_1']^{\vec{v}_1}) = \eta([h_1^{\vec{v}_1}, h_1'^{\vec{v}_1}]) = 0,$$

implying that $\xi_{2,1}$ vanishes on $[G_1, G_1]$. Thus, $\xi_{2,1}$ is a 2-rd level character.

Performing a similar analysis while looking at the coefficients of $\binom{x}{2}, xy + \binom{y}{2}, xy^3 + \binom{y^3}{2}$ and y^2 respectively, we conclude that for all $1 \leqslant j \leqslant 4$, the maps $\xi_{2,j}(h_2) = q\eta(h_2^{\vec{v}_j})$ from G_2 to $\mathbb{R}$ are 2-nd level characters. The nontriviality of η on G_2^4 and the fact that $\vec{v}_1, \vec{v}_2, \vec{v}_3$ and $\vec{v}_4$ span $\mathcal{P}_2 = \mathbb{R}^4$ imply that for at least one value $1 \leqslant i \leqslant 4$, the character η does not vanish on $G_2^{\vec{v}_i}$. We fix this i. From $\|\xi_{2,i}(g_i)\|_{\mathbb{R}/\mathbb{Z}} = \|q\eta(g_i^{\vec{v}_j})\|_{\mathbb{R}/\mathbb{Z}} < AN^{-i}$ and the (A, N)-irrationality of g_2 we deduce that $|\xi_{2,i}| > A$. Together with the bounds $q = O(1)$ and $|\vec{v}_1| = O(1)$, this implies that $|\eta| > c'A$ for some constant $c' > 0$. Choosing c_M in such a way that $c < c'$ gives the desired contradiction. Hence η vanishes on G_2^4.

This leaves us with

$$\eta \circ g^P(x, y) = \eta(g_1^{\vec{v}_1})x + \eta(g_1^{\vec{v}_2})y + \eta(g_1^{\vec{v}_3})y^3.$$

By analysing the coefficients of x, y and y^3 as above, we see that η vanishes on elements of the form $h_1^{\vec{v}_i}$ with $h_1 \in G_1$ and $1 \leqslant i \leqslant 3$. Thus, η vanishes on all of G^P. This contradicts the nontriviality of η, and so g^P is $(c_M A^{-C_M}, N)$-equidistributed on G^P/Γ^P. $\qquad\square$

We now prove Theorem 6.7 in full generality.

Proof of Theorem 6.7. Let $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ be an integral polynomial progression, $G_\bullet$ be a filtration of degree s and $g \in \text{poly}(\mathbb{Z}, G_\bullet)$. By (26), we can find a family $\{Q_{i,j} : 1 \leqslant i \leqslant s, 1 \leqslant j \leqslant t_i\}$ of linearly independent integral polynomials such that $Q_{i,1}, ..., Q_{i,t_i}$ is a basis for $W_i = W_i'$ for $1 \leqslant i \leqslant s$. It is crucial that these polynomials are linearly independent, which follows from homogeneity of $\vec{P}$. For each i, let $\tau_i : W_i \to \mathcal{P}_i$ be the map associated with $Q_{i,1}, ..., Q_{i,t_i}$ as defined in Section 4. We also let $\vec{v}_{i,j} \in \mathbb{Z}^{t+1}$ be the vectors such that $\tau_i(Q_{i,j}) = \vec{v}_{i,j}$.

As in the proof of Proposition 8.1, suppose that g^P is not $(c_M A^{-C_M}, N)$-equidistributed on G^P/Γ^P for some constants $0 < c_M < 1 < C_M$. We apply Theorem 6.5 again to conclude that there exists a nontrivial horizontal character $\eta : G^P \to \mathbb{R}$ of modulus at most cA, satisfying $\|\eta \circ g^P\|_{C^\infty[N]} \leqslant cA$ for some constant $c > 0$ that depends on c_M and C_M. The constant C_M is chosen in such a way as to match the exponents in the case (ii) of Theorem 6.5, but the choice of c_M is up to us again. We shall pick it small enough to show that the failure of g^P to be $(c_M A^{-C_M}, N)$-equidistributed contradicts the (A, N)-irrationality of g.

Thus,

$$\eta \circ g^P(x, y) = \sum_{i=1}^{s} \sum_{j=1}^{t_i} \eta(g_i^{\vec{v}_{i,j}}) Q_{i,j}(x, y).$$

Using Lemma 8.2 and the assumption $\|\eta \circ g^P\|_{C^\infty[N]} \leqslant cA$, and choosing c_M in such a way that $c > 0$ is sufficiently small, we deduce that there exists a positive integer $q = O(1)$ such that $\|q\eta(g_i^{\vec{v}_{i,j}})\|_{\mathbb{R}/\mathbb{Z}} < AN^{-i}$ for all $1 \leqslant i \leqslant s$ and $1 \leqslant j \leqslant t_i$.

Our goal now is to show by downward induction on i that η vanishes on the group

$$H_i = \langle h_i^{\vec{v}_{i,j}} : h_i \in G_i, 1 \leqslant j \leqslant t_i \rangle$$

for all $i \in \mathbb{N}_+$. This is trivially true for $i \geqslant s + 1$. Suppose that η vanishes on H_{i+1} for some $1 \leqslant i \leqslant s$ but that it does not vanish on H_i. We define the maps $\xi_{i,j} : G_i \to \mathbb{R}$ by

$\xi_{i,j}(h_i) = \eta(qh_i^{\vec{v}_{i,j}})$ and claim that they are i-th level characters. They are continuous group homomorphisms because η is, and they vanish on G_{i+1} by induction hypothesis. Since $q \in \mathbb{Z}$ and $\vec{v}_{i,j}$ have integer entries, we also have $\xi_{i,j}(\Gamma_i) \subseteq \mathbb{Z}$. It remains to show that $\xi_{i,j}$ vanishes on $[G_l, G_{i-l}]$ for all $1 \leqslant l < i$. The fact that $\mathcal{P}_i \subseteq \mathcal{P}_l \cdot \mathcal{P}_{i-l}$ implies the existence of $\vec{u}_l \in \mathcal{P}_l$ and $\vec{u}_{i-l} \in \mathcal{P}_{i-l}$ for which $\vec{v}_{i,j} = \vec{u}_l \cdot \vec{u}_{i-l}$, and so we have

$$[G_l^{\vec{u}_l}, G_{i-l}^{\vec{u}_{i-l}}] = [G_l, G_{i-l}]^{\vec{u}_l \cdot \vec{u}_{i-l}} \bmod G_{i+1}^{t+1},$$

from which it follows that $\xi_{i,j}|_{[G_l, G_{i-l}]} = 0$. Therefore each $\xi_{i,j}$ is an i-th level character.

The nontriviality of η on H_i and the fact that $\mathcal{P}_i$ is spanned by the vectors $\vec{v}_{i,1}, ..., \vec{v}_{i,t_i}$ imply that for at least one value $1 \leqslant j \leqslant t_i$, the character η does not vanish on $G_i^{\vec{v}_{i,j}}$, and so $\xi_{i,j}$ is nontrivial. From $||\xi_{i,j}(g_i)||_{\mathbb{R}/\mathbb{Z}} = ||q\eta(g_i^{\vec{v}_{i,j}})||_{\mathbb{R}/\mathbb{Z}} < AN^{-i}$ and the (A, N)-irrationality of g_i we deduce that $|\xi_{i,j}| > A$. Together with the bounds $q = O(1)$ and $|\vec{v}_{i,j}| = O(1)$, this implies that $|\eta| > c'A$ for some constant $c' > 0$. We choose c_M in such a way that $c < c'$; this contradicts the nontriviality of η on H_i. This proves the inductive step; hence η vanishes on all of G^P, contradicting the nontriviality of η. It follows that g^P is $(c_M A^{-C_M}, N)$-equidistributed on G^P/Γ^P. $\qquad\square$

9. THE FAILURE OF THEOREM 6.7 IN THE INHOMOGENEOUS CASE

Having derived Theorem 6.7, we want to show why an analogous statement fails in the inhomogeneous case. We let

$$(34) \qquad \vec{P}(x, y) = (x, \ x + y, \ x + 2y, \ x + y^2),$$

with a square instead of a cube in the last position. It is an inhomogeneous progression because of the inhomogeneous relation (10). Suppose that $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ is an irrational polynomial sequence with $g(0) = 1$ on a connected group G with a filtration $G_\bullet$ of degree 2. We shall try to show that g^P is equidistributed on G^P/Γ^P the same way as we argued in Proposition 8.1, and we indicate where and why the argument fails.

Once again, we let

$$\vec{v}_1 = (1, 1, 1, 1), \quad \vec{v}_2 = (0, 1, 2, 0), \quad \vec{v}_3 = (0, 0, 0, 1) \quad \text{and} \quad \vec{v}_4 = (0, 0, 1, 0),$$

and we observe that $\mathcal{P}_1 = \mathrm{Span}_{\mathbb{R}}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ and $\mathcal{P}_2 = \mathrm{Span}_{\mathbb{R}}\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$. Hence $G^P = G^{\vec{v}_1} G^{\vec{v}_2} G^{\vec{v}_3} G_2^4$. Suppose that g^P is not $(c_M A^{-C_M}, N)$-equidistributed on G^P/Γ^P for some constants $0 < c_M < 1 < C_M$. Theorem 6.5 once again implies the existence of a nontrivial horizontal character $\eta : G^P \to \mathbb{R}$ of modulus at most cA, for which $||\eta \circ g^P||_{C^\infty[N]} \leqslant cA$ for some constant $c > 0$ that depends on c_M and C_M.

Rewriting the expression for $\eta \circ g^P$, we see that

$$\eta \circ g^P(x, y) = \eta(g_1^{\vec{v}_1})x + \eta(g_1^{\vec{v}_2})y + \eta(g_1^{\vec{v}_3})y^2$$
$$+ \eta(g_2^{\vec{v}_1})\binom{x}{2} + \eta(g_2^{\vec{v}_2})\left(xy + \binom{y}{2}\right) + \eta(g_2^{\vec{v}_3})\left(xy^2 + \binom{y^2}{2}\right) + \vec{v}_4 y^2$$
$$= \eta(g_1^{\vec{v}_1})x + \eta(g_1^{\vec{v}_2})y + (\eta(g_1^{\vec{v}_3}) + \eta(g_2^{\vec{v}_4}))y^2$$
$$+ \eta(g_2^{\vec{v}_1})\binom{x}{2} + \eta(g_2^{\vec{v}_2})\left(xy + \binom{y}{2}\right) + \eta(g_2^{\vec{v}_3})\left(xy^2 + \binom{y^2}{2}\right).$$

Applying Lemma 8.2 and the assumption $||\eta \circ g^P||_{C^\infty[N]} \leqslant cA$, and choosing c_M in such a way that $c > 0$ is sufficiently small, we deduce that there exists a positive integer $q = O(1)$ such that

$$\tag{35} ||q\eta(g_i^{\vec{v}_j})||_{\mathbb{R}/\mathbb{Z}} < AN^{-i}$$

for all pairs

$$(i,j) \in \{(1,1),(1,2),(2,1),(2,2),(2,3)\}.$$

By looking at the coefficient of $\binom{x}{2}$, $xy + \binom{y}{2}$ and $xy^2 + \binom{y^2}{2}$, we deduce that the maps

$$h_2 \mapsto q\eta(h_2^{\vec{v}_1}),\ q\eta(h_2^{\vec{v}_2}),\ q\eta(h_2^{\vec{v}_3})$$

are trivial 2-nd level characters; the argument goes the exact same way as in the proof of Proposition 8.1. Thus, η vanishes on all elements of the form $h_2^{\vec{w}_2}$ with $h_2 \in G_2$ and

$$\vec{w}_2 \in \mathcal{P}_2' = \mathrm{Span}_{\mathbb{R}}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}.$$

By looking at the coefficients of x and y, we similarly show that η vanishes on all elements of the form $h_1^{\vec{w}_1}$ with $h_1 \in G_1$ and

$$\vec{w}_1 \in \mathcal{P}_1' = \mathrm{Span}_{\mathbb{R}}\{\vec{v}_1, \vec{v}_2\}.$$

We are left with

$$\eta \circ g^P(x,y) = \left(\eta(g_1^{\vec{v}_3}) + \eta(g_2^{\vec{v}_4})\right) y^2.$$

We would like to be able to say that η vanishes on all elements of the form $h_1^{\vec{w}_1}$ and $h_2^{\vec{w}_2}$ with $h_i \in G_i$ and $\vec{w}_i \in \mathcal{P}_i$; this would imply that η is trivial. For this to be case, it would suffice to show that both $\eta(g_1^{\vec{v}_3})$ and $\eta(g_2^{\vec{v}_4})$ satisfy an estimate (35), and then use (A,N)-irrationality of g_1 and g_2 to conclude that the characters $h_1 \mapsto q\eta(h_1^{\vec{v}_3})$ and $h_2 \mapsto q\eta(h_2^{\vec{v}_4})$ are trivial. Alas, this need not be true. In Proposition 8.1, the number $\eta(h_1^{\vec{v}_3})$ was the coefficient of y^3 while $\eta(h_2^{\vec{v}_4})$ was the coefficient of y^2, from which it followed that they both satisfied (35). Now, however, all we can show is that

$$\tag{36} ||q(\eta(g_1^{\vec{v}_3}) + \eta(g_2^{\vec{v}_4}))||_{\mathbb{R}/\mathbb{Z}} < AN^{-1}$$

because $\eta(g_1^{\vec{v}_3}) + \eta(g_2^{\vec{v}_4})$ is the coefficient of y^2. But it need not follow that either of $\eta(g_1^{\vec{v}_3})$ and $\eta(g_2^{\vec{v}_4})$ satisfies (35); in particular, g^P may take values in a proper rational subgroup of G^P.

We illustrate this with a specific example, akin to the example in Section 11 of [Kuc20]. Suppose that $G = G_1 = \mathbb{R}^2$, $G_2 = 0 \times \mathbb{R}$, $G_3 = 0 \times 0$. The sequence $g(n) = (an, b\binom{n}{2})$ is adapted to the filtration $G_\bullet$, and it is irrational if and only if a and b are irrational. We identify G^4 with $\mathbb{R}^8$ via the map

$$G^4 \to \mathbb{R}^8$$
$$((x_1,y_1),(x_2,y_2),(x_3,y_3),(x_4,y_4)) \mapsto (x_1,x_2,x_3,x_4,y_1,y_2,y_3,y_4).$$

Setting

$$\vec{v}_{11} = \vec{e}_1 + \vec{e}_2 + \vec{e}_3 + \vec{e}_4, \quad \vec{v}_{12} = \vec{e}_2 + 2\vec{e}_3, \quad \vec{v}_{13} = \vec{e}_4,$$
$$\vec{v}_{21} = \vec{e}_5 + \vec{e}_6 + \vec{e}_7 + \vec{e}_8, \quad \vec{v}_{22} = \vec{e}_6 + 2\vec{e}_7, \quad \vec{v}_{23} = \vec{e}_8, \quad \vec{v}_{24} = \vec{e}_7,$$

we observe that $G^P = \mathrm{Span}_{\mathbb{R}}\{\vec{v}_{11}, \vec{v}_{12}, \vec{v}_{13}, \vec{v}_{21}, \vec{v}_{22}, \vec{v}_{23}, \vec{v}_{24}\}$.

With these definitions, the coefficient of y^2 in g^P becomes $a\vec{v}_{13} + b\vec{v}_{24} = a\vec{e}_4 + b\vec{e}_7$. If $a, b, 1$ are rationally independent, then the closure of g^P is the image of the 7-dimensional subspace G^P in $(\mathbb{R}/\mathbb{Z})^8$. If a and b are rationally dependent, then the closure of g^P is the image in $(\mathbb{R}/\mathbb{Z})^8$ of the 6-dimensional subspace

$$\tilde{G} = \mathrm{Span}_{\mathbb{R}}\{\vec{v}_{11}, \vec{v}_{12}, a\vec{v}_{13} + b\vec{v}_{24}, \vec{v}_{21}, \vec{v}_{22}, \vec{v}_{23}\}.$$

Finally, if some rational linear combination of a and b is a rational number q/r in its lower terms with $r > 1$, then the closure of g^P is a union of at most r translates of a 6-dimensional subtorus of G^P/Γ^P. For instance, if $a = \sqrt{2}$ and $b = \sqrt{2} + \frac{1}{3}$, then we define

$$(37) \qquad \tilde{G} = \mathrm{Span}_{\mathbb{R}}\{\vec{v}_{11}, \vec{v}_{12}, \vec{v}_{13} + \vec{v}_{24}, \vec{v}_{21}, \vec{v}_{22}, \vec{v}_{23}\},$$

and observe that the sequences g_0^P, g_1^P, g_2^P defined by $g_i^P(x,y) = g^P(x, 3y + i)$ are equidistributed on $\tilde{G}/\tilde{\Gamma}$, $\frac{1}{3}\vec{v}_{24} + \tilde{G}/\tilde{\Gamma}$ and $\frac{1}{3}\vec{v}_{24} + \tilde{G}/\tilde{\Gamma}$ respectively. In particular, for inhomogeneous progressions it is not true that the group $\tilde{G}$ depends only on the filtration $G_\bullet$ and the progression $\vec{P}$.

While annihilating the coefficients of $\eta \circ g^P$, we were able to deal with the coefficients of x and y as well as $\binom{x}{2}$, $xy + \binom{y}{2}$ and $xy^2 + \binom{y^2}{2}$, which span the spaces W_1' and W_2' respectively. The problematic coefficient was that of y^2, belonging to the space W^c. We have remarked below (27) in Section 4 that the nontriviality of the subspace W^c prevents us from running the same argument as in Proposition 8.1 and Theorem 6.7 for inhomogeneous progressions; the problem with the coefficient of y^2 that we have encountered here illustrates this point. The reader should see from here how to generalise the aforementioned example to other inhomogeneous progression; this generalised construction proves part (ii) of Theorem 1.15.

10. Finding closure in the inhomogeneous case

Section 9 shows that we cannot always hope for the sequence g^P to equidistribute in G^P/Γ^P for an inhomogeneous progression $\vec{P}$. Here, we provide an inductive recipe for finding the closure of g^P in the case of $\vec{P}(x,y) = (x, x+y, x+2y, x+y^2)$. We believe that this argument could be generalised to an arbitrary inhomogeneous progressions; while trying to do so, however, we have encountered significant technical issues of linear algebraic nature that we have not been able to overcome.

Since the argument that we present here is already complicated enough, we prove it in an infinitary setting so as to avoid confusion coming from various quantitative parameters. In effect, we show the following.

Proposition 10.1. *Let G be a connected group with filtration $G_\bullet$ of degree s, and $\vec{P}(x,y) = (x, x+y, x+2y, x+y^2)$. Suppose that $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ is irrational. There exists a subgroup*

$\tilde{G} \leqslant G^P$ and a decomposition $g^P = \tilde{g}\gamma$, where $\tilde{g}$ takes values in $\tilde{G}$ and is equidistributed on $\tilde{G}/\tilde{\Gamma}$ whereas γ is periodic. Moreover, the group $\tilde{G}$ contains the subgroup

$$K = \langle h_i^{\vec{w}_i} : h_i \in G_i, \vec{w}_i \in \mathcal{P}_i', 1 \leqslant i \leqslant s \rangle,$$

where

$$\mathcal{P}_1' = \mathrm{Span}_{\mathbb{R}}\{(1,1,1,1),(0,1,2,0)\},$$
$$\mathcal{P}_2' = \mathrm{Span}_{\mathbb{R}}\{(1,1,1,1),(0,1,2,0),(0,0,0,1)\},$$
$$\mathcal{P}_3' = \mathcal{P}_4' = ... = \mathbb{R}^4.$$

We will need the following lemma, which is similar in spirit to Lemma 8.2.

Lemma 10.2. *Let $a_1,...,a_s$ be nonzero real numbers. Let $Q_1,...,Q_s \in \mathbb{Q}[x,y]$ be linearly independent integral polynomials, and suppose that $Q = a_1 Q_1 + ... + a_r Q_s$ takes values in $\mathbb{Q}$. Then $a_i \in \mathbb{Q}$ for all $1 \leqslant i \leqslant s$.*

Proof. Let b_{kli} be the coefficient of $\binom{x}{k}\binom{y}{l}$ in Q_i. Then

$$c_{kl} = a_1 b_{kl1} + ... + a_s b_{kls}$$

is the coefficient of $\binom{x}{k}\binom{y}{l}$ in Q, and so it is rational. Indexing the pairs $(k_1, l_1), ..., (k_u, l_u)$ in some arbitrary fashion, we obtain an $u \times s$ matrix $B = (b_{k_r l_r i})_{ri}$ as well as an s-dimensional column vector $a = (a_i)_i$ and a u-dimensional column vector $c = (c_{j_l k_l})_l$ such that $Ba = c$. The linear independence of $Q_1, ..., Q_r$ implies that B has full rank, and so there exists an invertible $s \times s$ submatrix $\tilde{B}$ of B and an s-dimensional column vector $\tilde{c}$ such that $\tilde{B}a = \tilde{c}$. Since the entries of $\tilde{B}$ are integers, the entries of $\tilde{B}^{-1}$ are rational numbers. The equality $a = \tilde{B}^{-1}\tilde{c}$ then implies that $a_i \in \mathbb{Q}$ for each $1 \leqslant i \leqslant s$. $\square$

Proof of Proposition 10.1. For each $i \geqslant 3$, we find a basis $\{Q_{i,1}, Q_{i,2}, Q_{i,3}, Q_{i,4}\}$ for W_i. The absence of an inhomogeneous algebraic relation of degree 3 or higher implies that

$$\sum_{i=3}^{s} W_i = \bigoplus_{i=3}^{s} W_i,$$

from which it follows that the set $\{Q_{i,j} : 3 \leqslant i \leqslant s,\ 1 \leqslant j \leqslant 4\}$ is linearly independent. For $3 \leqslant i \leqslant s$ and $1 \leqslant j \leqslant 4$, we let $\vec{v}_{i,j} = \tau_i(Q_{i,j})$. We also set

$$\vec{v}_1 = (1,1,1,1), \quad \vec{v}_2 = (0,1,2,0), \quad \vec{v}_3 = (0,0,0,1) \quad \text{and} \quad \vec{v}_4 = (0,0,1,0).$$

We want to find a subgroup $\tilde{G}$ of G^P on which we can guarantee equidistribution. Starting with

$$H^{(1)} = \langle h_1^{\vec{v}_3},\ h_2^{\vec{v}_4} : h_1 \in G_1, h_2 \in G_2 \rangle,$$

we inductively define a chain of subgroups

$$H^{(1)} \geqslant H^{(2)} \geqslant H^{(3)} \geqslant ...$$

as well as groups $G^{(k)} = \langle K, H^{(k)} \rangle$ and $\Gamma^{(k)} = \Gamma^P \cap G^{(k)}$. We note that $G^{(1)} = G^P$.

We also inductively define sequences $g^{(k)}$ and $h^{(k)}$, starting with $h^{(1)}(y) = g_1^{\vec{v}_3 y^2} g_2^{\vec{v}_4 y^2}$ and $g^{(1)} = g^P$. If $g^{(k)}$ is equidistributed in $G^{(k)}/\Gamma^{(k)}$, then we terminate the procedure. Otherwise Theorem 2.5 implies the existence of a nontrivial horizontal character $\eta^{(k)} : G^{(k)} \to \mathbb{R}$ that vanishes on all of $G^{(k)}$ except $H^{(k)}$, and for which $\eta^{(k)} \circ g^{(k)} = \eta^{(k)} \circ h^{(k)}$ takes values in $\mathbb{Z}$. We then take $G^{(k+1)} = \ker \eta^{(k)}$ and $H^{(k+1)} = \ker \eta^{(k)}|_{H^{(k)}}$, and we factorize $h^{(k)} = h^{(k+1)}\gamma^{(k+1)}$ using an infinitary version of Proposition 9.2 of [GT12], where $\eta^{k+1} \circ h^{(k+1)} = 0$ and $\gamma^{(k+1)}$ is periodic. We define

$$g^{(k+1)}(x,y) = g^{(k)}(x,y)(\gamma^{(k+1)}(y))^{-1}$$

and observe that

$$g^{(k+1)}(x,y) = g_1^{\vec{v}_1 x + \vec{v}_2 y} h^{(k+1)}(y) g_2^{\vec{v}_1\binom{x}{2} + \vec{v}_2(xy + \binom{y}{2}) + \vec{v}_3(xy^2 + \binom{y^2}{2})} \prod_{i=3}^{s}\prod_{j=1}^{4} g_i^{\vec{v}_{i,j} Q_{i,j}} \mod [G_1, G_2]^4.$$

The sequence $g^{(k+1)}$ takes values in $G^{(k+1)}$. We also write

$$h^{(k)}(y) = a^{(k)}(y)^{\vec{v}_4} b^{(k)}(y)^{\vec{v}_3},$$

with $a^{(k)}$ being G_2-valued and $b^{(k)}$ being G_1-valued. Letting $a^{(k)}(y) = \prod_{i=1}^{s} a_i^{(k)\binom{y}{i}}$ and similarly for $b^{(k)}$, we claim that $a_2^{(k)}$ and $b_2^{(k)}$ are irrational elements of G_2 and G_1 respectively with regard to the filtration $G_\bullet$ on G. Finally, we claim that

$$H^{(k)} = G_2^{\vec{v}_4} \mod G_1^{\vec{v}_3} \quad \text{and} \quad H^{(k)} = G_1^{\vec{v}_3} \mod G_2^{\vec{v}_4}$$

First, we observe that all these properties hold at $k = 1$. We assume that they hold for some $k \geqslant 1$, from which we aim to deduce that they also hold at $(k+1)$-th level.

If $g^{(k)}$ is equidistributed in $G^{(k)}/\Gamma^{(k)}$, then we are done. Otherwise there exists a nontrivial horizontal character $\eta^{(k)} : G^{(k)} \to \mathbb{R}$ for which $\eta^{(k)} \circ g^{(k)}$ is $\mathbb{Z}$-valued. We have

$$\begin{aligned}
\eta^{(k)} \circ g^{(k)}(x,y) &= \eta^{(k)}(g_1^{\vec{v}_1})x + \eta^{(k)}(g_1^{\vec{v}_2})y + \eta^{(k)}(h^{(k)}(y)) \\
&+ \eta^{(k)}(g_2^{\vec{v}_1})\binom{x}{2} + 2\eta^{(k)}(g_2^{\vec{v}_2})\left(xy + \binom{y}{2}\right) + \eta^{(k)}(g_2^{\vec{v}_3})\left(xy^2 + \binom{y^2}{2}\right) \\
&+ \sum_{i=3}^{k}\sum_{j=1}^{4} \eta^{(k)}(g_i^{\vec{v}_{i,j}})Q_{i,j}(x,y).
\end{aligned}$$

By looking at the coefficients of $Q_{i,j}$ for $3 \leqslant i \leqslant s$, applying Lemma 10.2, and following the same method as in the proof of Theorem 6.7, we see that $\eta^{(k)}$ vanishes on elements of the form $h_i^{\vec{v}_{i,j}}$ for $h_i \in G_i$, $3 \leqslant i \leqslant s$ and $1 \leqslant j \leqslant 4$, and so $\eta^{(k)}$ vanishes on all of $G_3 \times G_3 \times G_3 \times G_3$. This leaves us with

$$\begin{aligned}
\eta^{(k)} \circ g^{(k)}(x,y) &= \eta^{(k)}(g_1^{\vec{v}_1})x + \eta^{(k)}(g_1^{\vec{v}_2})y + \eta^{(k)}(h^{(k)}(y)) \\
&+ \eta^{(k)}(g_2^{\vec{v}_1})\binom{x}{2} + 2\eta^{(k)}(g_2^{\vec{v}_2})\left(xy + \binom{y}{2}\right) + \eta^{(k)}(g_2^{\vec{v}_3})\left(xy^2 + \binom{y^2}{2}\right).
\end{aligned}$$

We now carry on. By looking at the coefficient of $\binom{x}{2}$ and $xy + \binom{y}{2}$, we see that $\eta^{(k)}(g_2^{\vec{v}_1})$ and $\eta^{(k)}(g_2^{\vec{v}_2})$ are both integers, and so $\eta^{(k)}$ vanishes on all elements of the form $h_2^{\vec{v}_1}$ and $h_2^{\vec{v}_2}$ with $h_2 \in G_2$. By looking at the coefficients of x and y, we can similarly show that $\eta^{(k)}$ vanishes on all elements of the form $h_1^{\vec{v}_1}$ and $h_1^{\vec{v}_2}$ with $h_1 \in G_1$. We are thus left with

$$\eta^{(k)} \circ g^{(k)}(x,y) = \eta^{(k)}(h^{(k)}(y)) + \eta^{(k)}(g_2^{\vec{v}_3})\left(xy^2 + \binom{y^2}{2}\right).$$

We first deal with the last term. Since $H^{(k)} = G_1^{\vec{v}_3} \mod G_2^{\vec{v}_3}$, we have $[H^{(k)}, H^{(k)}] = [G_1^{\vec{v}_3}, G_1^{\vec{v}_3}] \mod G_3^4$. Using the fact that $\eta^{(k)}$ vanishes on both G_3^4 and $[H^{(k)}, H^{(k)}]$, we deduce that it also vanishes on $[G_1^{\vec{v}_3}, G_1^{\vec{v}_3}]$. Hence the function $\xi_{2,3} : G_2 \to \mathbb{R}$ given by $\xi_{2,3}(h) = \eta^{(k)}(h^{\vec{v}_3})$ is a 2-nd level character. By irrationality of g_2, it follows that $\xi_{2,3}$ is trivial, and so $\eta^{(k)}$ vanishes on $G_2^{\vec{v}_3}$. We have thus proved that $\eta^{(k)}$ vanishes on all of $G^{(k)}$ except $H^{(k)}$, and consequently that $\eta^{(k)} \circ g^{(k)} = \eta^{(k)} \circ h^{(k)}$.

We now show that

(38)
$$H^{(k+1)} = G_2^{\vec{v}_4} \mod G_1^{\vec{v}_3}$$

Suppose not; let U be a proper rational subgroup of $G_2^{\vec{v}_4}$ such that

$$H^{(k+1)} = U \mod G_1^{\vec{v}_3}.$$

Then

$$H^{(k+1)} \leqslant U G_1^{\vec{v}_3} \cap H^{(k)} \leqslant H^{(k)}.$$

We know from the rank-nullity theorem that $\dim H^{(k+1)} = \dim H^{(k)} - 1$, and we have $H^{(k)} = G_2^{\vec{v}_4} \mod G_1^{\vec{v}_3}$ from the inductive hypothesis. These two facts, together with the assumption that U is a proper rational subgroup of $G_2^{(0,0,1,0)}$, imply that $H^{(k+1)} = U G_1^{\vec{v}_3} \cap H^{(k)}$. It follows that

$$\eta^{(k)} \circ g^{(k)}(x,y) = \eta^{(k)}(a^{(k)}(y)^{\vec{v}_4}) + \eta^{(k)}(b^{(k)}(y)^{\vec{v}_3}) = \eta^{(k)}(a^{(k)}(y)^{\vec{v}_4})$$

We have already shown that $\eta^{(k)}$ vanishes on G_3^4. From the facts that $a^{(k)}(y) = \prod_{i=1}^{s} a_i^{(k)\binom{y}{i}}$ with $a_i^{(k)} \in G_i$, we deduce that $\eta^{(k)}(a^{(k)}(y)^{\vec{v}_4}) = \eta^{(k)}(a_1^{(k)})y + \eta^{(k)}(a_2^{(k)})\binom{y}{2}$. The map $\xi_{2,4}(h_2) = \eta^{(k)}(h_2^{\vec{v}_4})$ is a continuous group homomorphism on G_2 that vanishes on G_3 and sends Γ_2 to $\mathbb{Z}$. Since $\vec{v}_4 = (\vec{v}_2 \cdot \vec{v}_2 - \vec{v}_2)/2$, we also have

$$\xi_{2,4}([h_1, h_1']) = \frac{1}{2}\eta^{(k)}([h_1^{\vec{v}_2}, h_1'^{\vec{v}_2}]) - \frac{1}{2}\eta^{(k)}([h_1, h_1']^{\vec{v}_2}),$$

for any $h_1, h_1' \in G_1$, and so $\xi_{2,4}$ vanishes on $[G_1, G_1]$. Thus $\xi_{2,4}$ is a 2-nd level character on G_2 with respect to the filtration $G_\bullet$ on G, and since $a_2^{(k)}$ is an irrational element of G_2 with respect to this filtration, it follows that $\eta^{(k)}$ is trivial, a contradiction; hence (38) holds. The argument that

$$H^{(k+1)} = G_1^{\vec{v}_3} \mod G_2^{\vec{v}_4}$$

is similar.

Finally, we factorize $h^{(k)} = h^{(k+1)}\gamma^{(k+1)}$, where $\gamma^{(k+1)}$ is periodic and $h^{(k+1)}$ takes values in $H^{(k+1)} = \ker \eta^{(k+1)}$. It remains to show that $a_2^{(k+1)}$ and $b_2^{(k+1)}$ are irrational elements of G_2 and G_1 with respect to the filtration $G_\bullet$ on G. We observe that

$$a^{(k)} = a^{(k+1)}\gamma_a^{(k+1)} \quad \text{and} \quad b^{(k)} = b^{(k+1)}\gamma_b^{(k+1)}$$

for some periodic sequences γ_a and γ_b taking values in G_2 and G_1 respectively. Suppose that $\xi : G_2 \to \mathbb{R}$ is a 2-nd level character with respect to the filtration $G_\bullet$, for which $\xi(a_2^{(k+1)}) \in \mathbb{Z}$. The sequence $\gamma_a^{(k+1)}$ is periodic, hence $\xi \circ \gamma_a^{(k+1)}$ is $\mathbb{Q}$-valued, and so it follows that $\xi(a_2^{(k)}) \in \mathbb{Q}$ as well. Therefore there exists an integer $l > 0$ such that $l\xi(a_2^{(k)}) \in \mathbb{Z}$. Since $\xi' := l \cdot \xi$ is also a 2-nd level character, it follows from the irrationality of $a_2^{(k)}$ that ξ' is trivial. This implies that ξ is trivial as well, and hence $a_2^{(k+1)}$ is irrational. The argument showing that $b_2^{(k+1)}$ is irrational is identical.

We have thus shows inductively that $g^{(k)}$, $h^{(k)}$, $G^{(k)}$ and $H^{(k)}$ satisfy all the properties we want them to satisfy for all $k \geqslant 1$. Since $0 \leqslant \dim G^{(k+1)} < \dim G^{(k)}$, the procedure eventually terminates, at which point the sequence $g^{(k)}$ takes values in $G^{(k)}$ and is equidistributed on $G^{(k)}/\Gamma^{(k)}$. Letting $\tilde{G} = G^{(k)}$ for this value of k and $\gamma = \gamma^{(k)}...\gamma^{(1)}$, and observing that a product of periodic sequences is periodic, we finish the proof. $\qquad\square$

11. The equivalence of Weyl and algebraic complexity

While we are not able to show that Host-Kra and true complexities equal algebraic complexity for inhomogeneous progression, we can show the equivalence of Weyl and algebraic complexities for all integral progressions.

Definition 11.1 (Weyl system). *A Weyl system is an ergodic system $(X, \mathcal{X}, \mu, T)$, where X is a compact abelian Lie group and T is a unipotent affine transformation on X, i.e. $Tx = \phi(x) + a$ for $a \in X$ and an automorphism ϕ of X satisfying $(\phi - \mathrm{Id}_X)^s = 0$ for some $s \in \mathbb{N}_+$.*

We recall that an integral polynomial progression $\vec{P} \in \mathbb{R}[x, y]^{t+1}$ has Weyl complexity s at $0 \leqslant i \leqslant t$ if s the smallest natural number for which the factor $\mathcal{Z}_s$ is characteristic for the weak convergence of $\vec{P}$ at i for any Weyl system.

Every disconnected Weyl system can be written as a finite union of isomorphic tori that are cyclically permuted by the transformation T, much the same way as each disconnected nilsystem is a union of connected nilsystems (cf. Proposition 2.2 and the remark below Theorem 3.5 of [BLL07]). Therefore we can restrict our attention to connected Weyl systems. These can in turn be reduced to standard Weyl systems, which are totally ergodic by Proposition 2.2. Throughout this section, we let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$.

Definition 11.2 (Standard Weyl system of order s). *Let $s \in \mathbb{N}_+$ and $X = \mathbb{T}^s$. A standard Weyl system of order s is a system $(X, \mathcal{X}, \mu, T)$, where $\mathcal{X}$ is the Borel σ-algebra on X, μ is the Lebesgue measure, and*

$$T(a_1, ..., a_s) = (a_1 + a_0, a_2 + a_1, ..., a_s + a_{s-1})$$

for some irrational a_0.

Proposition 11.3 (Lemma 4.1 of [FK05]). *Each connected Weyl system is a factor of a product of several standard Weyl systems.*

Determining Weyl complexity therefore amounts to analysing standard Weyl systems. Since each standard Weyl system is totally ergodic, we immediately deduce the following.

Proposition 11.4. *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression. Then $\mathcal{W}_i(\vec{P}) \leqslant \mathcal{HK}_i(\vec{P})$ for all $0 \leqslant i \leqslant t$.*

We now fix a standard Weyl system $(X, \mathcal{X}, \mu, T)$ of order s with some irrational a_0. Then

$$(39) \qquad T^n(a_1, ..., a_s) = \left(a_1 + na_0, a_2 + na_1 + \binom{n}{2}a_0, ..., a_s + na_{s-1} + ... + \binom{n}{s}a_0 \right)$$

$$= g_0 + g_1 n + ... + g_s \binom{n}{s},$$

where $g_i = (a_{1-i}, ..., a_{s-i})$ and $a_{-k} = 0$ for $k > 0$. For almost all points $a = (a_1, ..., a_s) \in \mathbb{R}^s$, the numbers $1, a_0, ..., a_s$ are rationally independent, and we fix a point $a \in \mathbb{R}^s$ for which this is the case. The sequence $g(n) = T^n a$ is adapted to the filtration $G_i = \{0\}^{i-1} \times \mathbb{R}^{s-i+1}$ for $1 \leqslant i \leqslant s$ and $G_i = 0$ for $i > s$ on $G = G_0 = \mathbb{R}^s$, and it is irrational due to the irrationality of a_0. Since the $\mathcal{Z}_i$ factor of X consists of all the functions whose values depend only on the first i coordinates, we have $Z_i = G/G_{i+1}\Gamma = \mathbb{T}^i \times \{0\}^{s-i}$, where $\Gamma = \mathbb{Z}^s$.

What we aim to show is therefore the following.

Proposition 11.5. *Let $t \in \mathbb{N}_+$, $(X, \mathcal{X}, \mu, T)$ be a standard Weyl system of order s and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression. Fix $0 \leqslant i \leqslant t$ and suppose that $\mathcal{A}_i(\vec{P}) = s'$. Then the image of the group $\{0\}^i \times G_{s'+1} \times \{0\}^{t-i}$ is contained in the closure of g^P inside $(G/\Gamma)^{t+1}$.*

If $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ is a homogeneous progression, then the sequence g^P is equidistributed in G^P/Γ^P by Theorem 5.3, and Proposition 11.5 follows immediately; we want to say something about the closure of g^P in the general case. We fix an integral progression $\vec{P}$ for the rest of this section. For each $1 \leqslant i \leqslant s$, we pick linearly independent integral polynomials $Q_{i,1}, ..., Q_{i,t'_i}$ that form a basis for W'_i. We also let $\{R_1, ..., R_r\}$ be a basis for W^c consisting of integral polynomials. Thus,

$$\binom{\vec{P}}{i} = \sum_{j=1}^{t'_i} \vec{v}_{i,j} Q_{i,j} + \sum_{j=1}^{r} \vec{w}_{i,j} R_j$$

for some vectors $\vec{v}_{i,j}, \vec{w}_{i,j} \in \mathbb{Z}^{t+1}$, which follows from (27). Consequently,

$$(40) \qquad g^P = g_0 \vec{1} + \sum_{i=1}^{s} g_i \sum_{j=1}^{t'_i} \vec{v}_{i,j} Q_{i,j} + \sum_{j=1}^{r} \left(\sum_{i=1}^{s} g_i \vec{w}_{i,j} \right) R_j$$

We should explain the notation used in (40). For $h \in G$ and $\vec{v} \in \mathbb{R}^{t+1}$, we interpret $h\vec{v}$ as the element of $(\mathbb{R}^s)^{t+1}$ of the form $(hv(0), ..., hv(t))$, where $hv(i) = (h_1 v(i), ..., h_s v(i))$ is an element of $\mathbb{R}^s$ for each $h = (h_1, ..., h_s) \in \mathbb{R}^s$ and $\vec{v} = (v(0), ..., v(t))$. Thus, $h\vec{v}$ is the same as

what we previously called $h^{\vec{v}}$. We use the additive notation $h\vec{v}$ now since we are working in an abelian setting. We also denote $\vec{1} = (1, ..., 1)$.

We let $A_{i,j} = \mathrm{Span}_{\mathbb{R}}\{g_i\vec{v}_{i,j}\}$ and $B_j = \mathrm{Span}_{\mathbb{R}}\{\sum_{i=1}^{s} g_i\vec{w}_{i,j}\}$, and we denote the closure of their images in $(G/\Gamma)^{t+1}$ as $\overline{A}_{i,j}$ and $\overline{B}_j$ respectively. From the rational independence of a_i and the rationality of the entries of $\vec{v}_{i,j}$ and $\vec{w}_{i,j}$, we deduce that nonzero entries of $g_i\vec{v}_{i,j}$ and $\sum_{i=1}^{s} g_i\vec{w}_{i,j}$ are irrational; therefore the sequences $(x,y) \mapsto g_i\vec{v}_{i,j}Q_{i,j}(x,y)$ and $(x,y) \mapsto \sum_{i=1}^{s} g_i\vec{w}_{i,j}R_j(x,y)$ are equidistributed on $\overline{A}_{i,j}$ and $\overline{B}_j$ respectively. The linear independence of $Q_{i,j}, R_j$ then implies the following.

Proposition 11.6. *The closure of g^P is the image of $g_0\vec{1} + \tilde{G}$ inside $(G/\Gamma)^{t+1}$, where*

$$\tilde{G} = \sum_{i=1}^{s}\sum_{j=1}^{t'_i} A_{i,j} + \sum_{j=1}^{r} B_j.$$

In particular, the group $\tilde{G}$ contains

$$K = \sum_{i=1}^{s}\sum_{j=1}^{t'_i} A_{i,j} = \mathrm{Span}_{\mathbb{R}}\{h_i\vec{v}_{i,j} : h_i \in G_i, 1 \leqslant i \leqslant s, 1 \leqslant j \leqslant t'_i\}.$$

We observe that $K = \tilde{G} = G^P$ whenever $\vec{P}$ is homogeneous.

Corollary 11.7. *Fix $0 \leqslant i \leqslant t$ and let $\mathcal{A}_i(\vec{P}) < s$. For $k \leqslant s$, we have $\{0\}^i \times G_k \times \{0\}^{t-i} \leqslant K$ if and only if $k > \mathcal{A}_i(\vec{P})$.*

Proof. For each $1 \leqslant k \leqslant s$, we let $\mathcal{P}'_k = \mathrm{Span}_{\mathbb{R}}\{\vec{v}_{k,1}, ..., \vec{v}_{k,t'_k}\}$. Thus

$$K = \mathrm{Span}_{\mathbb{R}}\{h_k\vec{u}_k : 1 \leqslant k \leqslant s, h_k \in G_k, \vec{u}_k \in \mathcal{P}'_k\},$$

and so for $k \leqslant s$, we have the inclusion $\{0\}^i \times G_k \times \{0\}^{t-i} \leqslant K$ if and only if the vector $\vec{e}_i$ with 1 in the i-th position and 0 elsewhere is contained in $\mathcal{P}'_k$. The statement $\vec{e}_i \in \mathcal{P}'_k$ is equivalent to the inclusion $\binom{x+P_i(y)}{k} \in W'_k$. This is in turn equivalent to the statement that there are no algebraic relations of the form (8) with $\deg Q_i = k$, which is precisely the condition that $k > \mathcal{A}_i(\vec{P})$. $\qquad\square$

Corollary 11.8. *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression. Then $\mathcal{W}_i(\vec{P}) \leqslant \mathcal{A}_i(\vec{P})$ for each $0 \leqslant i \leqslant t$.*

We finish this section by showing the converse.

Proposition 11.9. *Let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral polynomial progression for which $\mathcal{A}_i(\vec{P}) = s$ for some $0 \leqslant i \leqslant t$. Then for any standard Weyl system $(X, \mathcal{X}, \mu, T)$ of order s there exist smooth functions $f_0, ..., f_t : X \to \mathbb{C}$ such that $\mathbb{E}(f_i|\mathcal{Z}_{s-1}) = 0$ but the expression (28) is 1. In particular, $\mathcal{W}_i(\vec{P}) \geqslant s$.*

Before we prove Proposition 11.9, we define $\partial Q(x) = Q(x+1) - Q(x)$ for $Q \in \mathbb{R}[x]$. From the identity $\partial\binom{x}{k} = \binom{x+1}{k} - \binom{x}{k} = \binom{x}{k-1}$ we deduce that

$$\partial\left(a_0 + a_1\binom{x}{1} + ... + a_d\binom{x}{d}\right) = a_1 + a_2\binom{x}{1} + ... + a_d\binom{x}{d-1}.$$

Proof of Proposition 11.9. Let T be as in (39) for some irrational a_0. From $\mathcal{A}_i(\vec{P}) = s$ it follows that $\vec{P}$ satisfies an algebraic relation (8) with $\deg Q_i = s$. For each $0 \leqslant k \leqslant t$, we let $Q_k(u) = b_{k,1}u + ... + b_{k,s}\binom{u}{s}$. We define $\xi(u) = e(\alpha u)$ for some irrational α, and we let

$$f_k(a_1, ..., a_s) = \xi(b_{k,1}a_1 + ... + b_{k,s}a_s).$$

Thus, we have

$$f_k(T^{x+P_k(y)}a) = \xi(a_0 Q_k(x + P_k(y)) + a_1 \partial Q_k(x + P_k(y)) + ... + a_s \partial^s Q_k(x + P_k(y))),$$

and so

$$\prod_{i=0}^{t} f_i(T^{x+P_i(y)}a) = \xi\left(\sum_{j=0}^{s} a_j \partial^j \sum_{k=0}^{t} Q_k(x + P_k(y))\right) = 1.$$

On the other hand, we have

$$|\mathbb{E}(f_i|\mathcal{Z}_{s-1})(a_1, ..., a_s)| = \left|\int_{\mathbb{T}} f_i(a_1, ..., a_s)da_s\right| = \left|\int_{\mathbb{T}} \xi(b_{i,s}a_s)da_s\right| = 0$$

for a.e. a_s. $\qquad\square$

12. The proof of Theorem 1.14

We conclude the paper with the proof of Theorem 1.14. Throughout this section, we let $t \in \mathbb{N}_+$ and $\vec{P} \in \mathbb{R}[x,y]^{t+1}$ be an integral progression of algebraic complexity at most 1. We also let $Q_1, ..., Q_k$ be integral polynomials as in the statement of Theorem 1.14. Thus, $P_i = \sum_j a_{ij}Q_j$ and $Q_i = \sum_j a'_{ij}P_j$ for $a_{ij}, a'_{ij} \in \mathbb{Z}$. The second part of the theorem follows from the first part and the Furstenberg Correspondence Principle. We therefore proceed to prove part (i), followed by part (iii). Our argument for part (i) follows closely the proof of Theorem C of [Fra08].

Proof of Theorem 1.14(i). We first prove part (i) of Theorem 1.14 in the totally ergodic case. Suppose that $(X, \mathcal{X}, \mu, T)$ is a totally ergodic system with the Kronecker factor $(Z_1, \mathcal{Z}_1, \nu, S)$. The space Z_1 can be assumed to be a connected compact abelian group with an ergodic translation $Sx = x + b$. For each $\delta > 0$, let B_δ be the δ-neighbourhood of the identity in Z_1, and let

$$\tilde{B}_\delta = \{n \in \mathbb{N} : Q_1(n)b, ..., Q_k(n)b \in B_\delta\}.$$

It follows from the ergodicity of S and linear independence of $Q_1, ..., Q_k$ that

$$\lim_{N-M\to\infty} \frac{|\tilde{B}_\delta \cap [M,N)|}{N-M} = \nu(B_\delta)^k > 0$$

for any $\delta > 0$. In particular, $\tilde{B}_\delta$ is syndetic for any $\delta > 0$, otherwise we would have $\liminf_{N-M\to\infty} \frac{|\tilde{B}_\delta \cap [M,N)|}{N-M} = 0$.

We aim to show that for any $A \in \mathcal{X}$ with $\mu(A) > 0$ and any $\epsilon > 0$, we have

(41) $$\lim_{N-M\to\infty} \mathbb{E}_{n\in\tilde{B}_\delta \cap [M,N)}\mu(A \cap T^{P_1(n)}A \cap ... \cap T^{P_t(n)}A) \geqslant \mu(A)^{t+1} - \epsilon$$

for all sufficiently small $\delta > 0$. This implies part (i) of Theorem 1.14 as follows: if there is a sequence K_N of intervals in $\mathbb{N}$ of length converging to infinity, with the property that

$$(42) \qquad \mu(A \cap T^{P_1(n)}A \cap ... \cap T^{P_t(n)}A) < \mu(A)^{t+1} - \epsilon$$

for all $n \in \bigcup_{N \in \mathbb{N}} K_N$, then the sets $\tilde{K}_N = K_N \cap \tilde{B}_\delta$ are nonempty for all sufficiently large N due to the syndecticity of B_δ (in fact, their cardinalities also converge to infinity). Since (42) holds for all $n \in \bigcup_{N \in \mathbb{N}} \tilde{K}_N$, the equality (41) fails, leading to a contradiction.

We first show that if $\mathbb{E}(f_i|\mathcal{Z}_1) = 0$, then

$$(43) \qquad \lim_{N-M \to \infty} \mathbb{E}_{n \in [M,N)} 1_{\tilde{B}_\delta}(n) \prod_{i=1}^{t} T^{P_i(n)} f_i = 0$$

in L^2 for any $f_1, ..., f_t \in L^\infty(\mu)$. From the measurability of B_δ it follows that we can approximate $1_{\tilde{B}_\delta}(n) = \prod_{i=1}^{k} 1_{B_\delta}(Q_i(n)b)$ arbitrarily well by linear combinations of $\prod_{i=1}^{k} \xi_i(Q_i(n)b)$ for some characters $\xi_1, ..., \xi_k$ on Z_1. Using the fact that each Q_i is an integral linear combination of $P_1, ..., P_t$, we can rewrite $\prod_{i=1}^{k} \xi_i(Q_i(n)b) = \prod_{i=1}^{t} \tilde{\xi}_i(P_i(n)b)$ for some characters $\tilde{\xi}_1, ..., \tilde{\xi}_t$.

In effect, it suffices to show that

$$(44) \qquad \lim_{N-M \to \infty} \mathbb{E}_{n \in [M,N)} \prod_{i=1}^{t} \tilde{\xi}_i(P_i(n)b) \prod_{i=1}^{t} T^{P_i(n)} f_i = 0.$$

We can rephrase the limit in (44) as

$$(45) \qquad \lim_{N-M \to \infty} \prod_{i=1}^{t} \tilde{\xi}_i(-y) \mathbb{E}_{n \in [M,N)} \prod_{i=1}^{t} R^{P_i(n)}(f_i(x)\tilde{\xi}_i(y)),$$

where $R = T \times S$. Let $(R_t)_t$ be the ergodic components of R and $(f_i \otimes \xi_i)(x,y) = f_i(x)\xi_i(y)$; then $\mathbb{E}(f_i \otimes \xi_i|\mathcal{Z}_1(R_t)) = 0$ whenever $\mathbb{E}(f_i|\mathcal{Z}_1(T)) = 0$ for a.e. t. It thus follows from Corollary 1.12 that if $\mathbb{E}(f_i|\mathcal{Z}_1) = 0$ for some i, then the limit in (45) is 0, which proves the claim.

We therefore deduce that

$$\lim_{N-M \to \infty} \mathbb{E}_{n \in \tilde{B}_\delta \cap [M,N)} \int_X \prod_{i=0}^{t} T^{P_i(n)} 1_A d\mu = \lim_{N-M \to \infty} \mathbb{E}_{n \in \tilde{B}_\delta \cap [M,N)} \int_{Z_1} \prod_{i=0}^{t} S^{P_i(n)} \tilde{1}_A d\nu$$

$$(46) \qquad = \lim_{N-M \to \infty} \mathbb{E}_{n \in \tilde{B}_\delta \cap [M,N)} \int_{Z_1} \prod_{i=0}^{t} S^{\sum_j a_{ij} Q_j(n)} \tilde{1}_A d\nu,$$

where $\tilde{1}_A = \mathbb{E}(1_A|\mathcal{Z}_1)$. Due to the ergodicity of S and the linear independence of $Q_1, ..., Q_k$, the limit in (46) equals

$$(47) \qquad \frac{1}{\nu(B_\delta)^k} \int_{B_\delta^k} \int_{Z_1} \prod_{i=0}^{t} \tilde{1}_A(x + \sum_j a_{ij} y_j) d\nu(x) d\nu^k(y).$$

In the limit $\delta \to 0$, the expression in (47) converges to $\int_{Z_1} (\tilde{1}_A)^{t+1}$; hence for every $\epsilon > 0$ and sufficiently small $\delta > 0$, we have

$$(48) \qquad \frac{1}{\nu(B_\delta)^k} \int_{B_\delta^k} \int_{Z_1} \prod_{i=0}^{t} \tilde{1}_A(x + \sum_j a_{ij} y_j) d\nu(x) d\nu^k(y) \geqslant \int_{Z_1} (\tilde{1}_A)^{t+1} - \epsilon.$$

Using Hölder inequality, we obtain that $\int_{Z_1}(\tilde{1}_A)^{t+1} \geqslant (\int_{Z_1}\tilde{1}_A)^{t+1} = \mu(A)^{t+1}$, which implies (41). This finished the totally ergodic case; the derivation of the ergodic case from the totally ergodic one proceeds in the same way as in the proof of Theorem C of [Fra08].

$\square$

We now proceed to the proof of part (iii) of Theorem 1.14. The argument can below can be seen as a finitary version of the argument above, with all the necessary modifications coming from working in the finitary setting. It follows the proof of the 3-term arithmetic progression case in Theorem 1.12 of [GT10].

Proof of Theorem 1.14(iii). Let $\alpha, \epsilon > 0$, and suppose that $A \subset \mathbb{Z}/N\mathbb{Z}$ has size $|A| \geqslant \alpha N$ for a prime $N > N_0(\alpha, \epsilon)$. Let $\mathcal{F} : \mathbb{R}_+ \to \mathbb{R}_+$ be a growth function to be specified later. By Theorem 5.1 of [CS12], the irrational and periodic version of the celebrated arithmetic regularity lemma of Green and Tao (Theorem 1.2 of [GT10]), there exists a positive number $M = O_{\epsilon, \mathcal{F}}(1)$ and a decomposition

$$(49) \qquad 1_A = f_{nil} + f_{sml} + f_{unf}$$

into 1-bounded functions such that

(i) $f_{nil} = F(g(n)\Gamma)$ is an $\mathcal{F}(M)$-irrational, N-periodic nilsequence of degree 1 and complexity M;

(ii) $||f_{sml}||_1 \leqslant \epsilon$;

(iii) $||f_{unf}||_{U^2} \leqslant \frac{1}{\mathcal{F}(M)}$.

Moreover, f_{nil} takes values in $[0, 1]$. Unpacking the definition of f_{nil}, we see that $F : (\mathbb{R}/\mathbb{Z})^m \to [0, 1]$ is M-Lipschitz, $1 \leqslant m \leqslant M$, and $g(n) = bn$ for some $\mathcal{F}(M)$-irrational element $b \in (\frac{1}{N}\mathbb{Z}/\mathbb{Z})^m$.

Our strategy is as follows. We shall define a weight $\tilde{\mu} : \mathbb{Z}/N\mathbb{Z} \to \mathbb{R}_{\geqslant 0}$ which satisfies

$$(50) \qquad \mathbb{E}_{y \in \mathbb{Z}/N\mathbb{Z}}\tilde{\mu}(y) = 1 + O(\epsilon)$$

and

$$(51) \qquad \mathbb{E}_{x,y \in \mathbb{Z}/N\mathbb{Z}}\tilde{\mu}(y)\prod_{i=0}^{t}1_A(x + P_i(y)) \geqslant \alpha^{t+1} - O(\epsilon).$$

Using pigeonhole principle and (50), it can be deduced from (51) that for $\Omega_{\alpha,\epsilon}(N)$ values of y, we have

$$\mathbb{E}_{x \in \mathbb{Z}/N\mathbb{Z}}\prod_{i=0}^{t}1_A(x + P_i(y)) \geqslant \alpha^{t+1} - O(\epsilon),$$

which proves part (iii) of Theorem 1.14.

We shall prove (51) by splitting each 1_A using (49) and showing that terms involving f_{sml} or f_{unf} have contributions at most $O(\epsilon)$ while the term

$$(52) \qquad \mathbb{E}_{x,y \in \mathbb{Z}/N\mathbb{Z}}\tilde{\mu}(y)\prod_{i=0}^{t}f_{nil}(x + P_i(y))$$

has size at least $\alpha^{t+1} - O(\epsilon)$. Showing that the terms involving f_{sml} or f_{unf} make negligible contributions to (51) is akin to showing (43) for all functions with $\mathbb{E}(f_i|\mathcal{Z}_1) = 0$ in the proof of part (i) of Theorem 1.14. In doing so, we shall use the idea that while we fix $\epsilon > 0$, we have control over how fast we choose $\mathcal{F}$ to grow - and we choose it to grow fast enough depending on α and ϵ to ensure that all the estimates work.

Let $\delta > 0$ be fixed later. We define $\psi : (\mathbb{R}/\mathbb{Z})^m \to \mathbb{R}_+$ to be a nonnegative, 1-bounded, $O_M(\delta^{-1})$-Lipschitz function that is 1 on $[-\frac{1}{4}\delta, \frac{1}{4}\delta]^m$ and 0 outside $[-\frac{1}{2}\delta, \frac{1}{2}\delta]^m$. We let $c = \int_{(\mathbb{R}/\mathbb{Z})^m} \psi$; thus $(\frac{1}{2}\delta)^m \leqslant c \leqslant \delta^m$. We then let $\mu(y) = \frac{\psi(by)}{c}$. Since b can be picked without the loss of generality from $[0, 1]^m$, the function μ is $O_M(\delta^{-M-1})$-Lipschitz.

We let $\tilde{\mu}(y) = \mu(Q_1(y))...\mu(Q_k(y))$. It is a weight that picks out all the values y for which $Q_1(y)b$, ..., $Q_k(y)b$ are close to being an integer, and it plays a similar role as the function $1_{\tilde{B}_\delta}$ in the proof of part (i) of Theorem 1.14, except that it is constructed using a Lipschitz function rather than an indicator function. To show (50), we observe that

$$(53) \qquad \mathbb{E}_{y \in \mathbb{Z}/N\mathbb{Z}} \tilde{\mu}(y) = \frac{1}{c^k} \mathbb{E}_{y \in [N]} \prod_{i=1}^{k} \psi(bQ_i(y)).$$

Using the $\mathcal{F}(M)$-irrationality of g, linear independence of Q_1, ..., Q_k as well as Theorem 2.5, we deduce that (53) equals

$$\frac{1}{c^k} \left(\left(\int \psi \right)^k + O_M(\delta^{-1}\mathcal{F}(M)^{-c_M}) \right) = 1 + O_M(\delta^{-M-2}\mathcal{F}(M)^{-c_M})$$

for some $c_M > 0$. The estimate (50) follows from choosing $\mathcal{F}$ growing fast enough depending on δ and picking $\delta = c'_M \epsilon$ for an appropriately chosen $c'_M > 0$.

We decompose each 1_A in (51) using (49) and split (51) into 3^t accordingly using multilinearity. We first estimate (52), and subsequently we bound contributions of f_{sml} and f_{unf}.

Taking $\mathcal{F}$ growing fast enough, we assume that $||f_{unf}||_{U^2} \leqslant \epsilon$, and thus $|\mathbb{E}_{x \in \mathbb{Z}/N\mathbb{Z}} f_{unf}(x)| = ||f_{unf}||_{U^1} \leqslant ||f_{unf}||_{U^2} \leqslant \epsilon$. From Hölder inequality and the bound on the L^1 norm of f_{sml}, we obtain a bound $|\mathbb{E}_{x \in \mathbb{Z}/N\mathbb{Z}} f_{sml}| \leqslant \epsilon$. From these bounds and (49) we deduce that $\mathbb{E}_{x \in \mathbb{Z}/N\mathbb{Z}} f_{nil}(x) \geqslant \alpha - 2\epsilon$.

We observe that by M-Lipschitzness of F and the definitions of μ, $\tilde{\mu}$ and Q_j, we have $f(x + P_i(y)) = f(x + \sum_j a_{ij} Q_j(y)) = f(x) + O_M(\delta) = f(x) + O(\epsilon)$ whenever $\tilde{\mu}(y) > 0$. It follows from this that

$$\mathbb{E}_{x, y \in \mathbb{Z}/N\mathbb{Z}} \tilde{\mu}(y) \prod_{i=0}^{t} f(x + \sum_j a_{ij} Q_j(y))$$

$$(54) \qquad = \left(\mathbb{E}_{x \in \mathbb{Z}/N\mathbb{Z}} f(x)^{t+1} + O(\epsilon) \right) \mathbb{E}_{y \in \mathbb{Z}/N\mathbb{Z}} \tilde{\mu}(y).$$

Using the estimate for (50) and Hölder inequality, we deduce that (54) is bounded from below by

$$\left(\mathbb{E}_{x \in \mathbb{Z}/N\mathbb{Z}} f_{nil}(x) \right)^{t+1} - O(\epsilon) \geqslant \alpha^{t+1} - O(\epsilon),$$

where the last inequality follows from Hölder inequality.

We now bound terms involving f_{sml}. Suppose without loss of generality that f_{sml} is in the $i = 0$ position, and let $f_1, ..., f_t \in \{f_{nil}, f_{sml}, f_{unf}\}$. Then

$$(55) \qquad |\mathbb{E}_{x,y \in \mathbb{Z}/N\mathbb{Z}} \tilde{\mu}(y) f_{sml}(x) \prod_{i=1}^{t} f_i(x + P_i(y))| \leqslant ||f_{sml}||_1 \mathbb{E}_{y \in \mathbb{Z}/N\mathbb{Z}} \tilde{\mu}(y) \leqslant \epsilon,$$

where the first inequality follows from Hölder inequality, positivity of $\tilde{\mu}$ and 1-boundedness of $f_1, ..., f_t$.

It remains to bound the contributions of f_{unf}. Using a standard argument (see e.g. the proof of Proposition 3.1 of [GT12]), we want to approximate f_{unf} by a trigonometric polynomial, which allows us to essentially replace f_{unf} by additive characters. Let $K \in \mathbb{N}_+$ be fixed later. Since μ is an $O_M(\epsilon^{-M})$-Lipschitz function, there exists a trigonometric polynomial $\mu_1 : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ such that $||\mu - \mu_1||_\infty \ll_M \epsilon^{-C_M^{(1)}} K^{-c}$ for some $0 < c, C_M^{(1)}$. Moreover, μ_1 has degree at most K^M and its coefficients satisfy $||\widehat{\mu_1}||_\infty \leqslant ||\mu||_\infty \ll_M \epsilon^{-M}$.

Let $f_0, ..., f_t \in \{f_{nil}, f_{sml}, f_{unf}\}$, with at least one of them being f_{unf}. We then bound

$$(56) \qquad |\mathbb{E}_{y \in \mathbb{Z}/N\mathbb{Z}} \tilde{\mu}(y) \prod_{i=0}^{t} f_i(x + P_i(y))| = |\mathbb{E}_{y \in \mathbb{Z}/N\mathbb{Z}} \prod_{i=1}^{k} \mu(Q_i(y)) \prod_{i=0}^{t} f_i(x + P_i(y))|$$

$$\leqslant k \max(||\mu||_\infty, ||\mu_1||_\infty)^{k-1} ||\mu - \mu||_\infty + |\mathbb{E}_{y \in \mathbb{Z}/N\mathbb{Z}} \prod_{i=1}^{k} \mu_1(Q_i(y)) \prod_{i=0}^{t} f_i(x + P_i(y))|.$$

The first term has size at most $C_M^{(2)} \epsilon^{-C_M^{(2)}} K^{-c}$ for some $C_M^{(2)} > 0$. The second term is bounded by

$$(57) \qquad K^M ||\widehat{\mu_1}||_\infty |\mathbb{E}_{y \in \mathbb{Z}/N\mathbb{Z}} \prod_{i=1}^{k} \xi_i(Q_i(y)) \prod_{i=1}^{t} f_i(x + P_i(y))|$$

for some characters ξ_i on $\mathbb{Z}/N\mathbb{Z}$. Since each Q_i is an integral linear combination of P_i's, we can rewrite $\prod_{i=1}^{k} \xi_i(Q_i(y)) = \prod_{i=1}^{t} \tilde{\xi}_i(x + P_i(y))$. We let $\tilde{f}_i = f_i \tilde{\xi}_i$. Since each $\tilde{\xi}_i$ is a linear character, we have $||f_i||_{U^2} = ||\tilde{f}_i||_{U^2}$ for each i.

We recall from Theorem 1.11 that $\vec{P}$ has true complexity 1. Combining this fact with (56), (57) and the bound $||\tilde{f}_i||_{U^2} \leqslant 1/\mathcal{F}(M)$ for some i, we deduce that there is some decreasing function $\omega : \mathbb{R}_+ \to \mathbb{R}_+$, depending only on $\vec{P}$, such that

$$(58) \qquad |\mathbb{E}_{y \in \mathbb{Z}/N\mathbb{Z}} \tilde{\mu}(y) \prod_{i=0}^{t} f_i(x + P_i(y))| \leqslant C_M^{(2)} \epsilon^{-C_M^{(2)}} K^{-c} + C_M^{(2)} \epsilon^{-M} K^M \omega(1/\mathcal{F}(M)),$$

increasing the constant $C_M^{(2)}$ if necessary. We note that the existence of ω is equivalent to the statement that $\vec{P}$ is controlled by U^2 at i. We now show that we can choose K large enough and $\mathcal{F}$ growing fast enough so that the right-hand side of (58) is bounded by $O(\epsilon)$.

For any given M, we find a constant $C_M^{(3)}$ such that $(C_M^{(3)})^c \geqslant C_M^{(2)}$ and $c C_M^{(3)} - C_M^{(2)} \geqslant 1$. We then let $K_M = C_m^{(3)} \epsilon^{-C_M^{(3)}}$, so that

$$C_M^{(2)} \epsilon^{-C_M^{(2)}} K_M^{-c} = C_M^{(2)} C_M^{(3)-c} \epsilon^{c C_M^{(3)} - C_M^{(2)}} \leqslant \epsilon.$$

Picking $\mathcal{F}$ growing sufficiently fast depending on ϵ, we can ensure that $C_M^{(2)} \epsilon^{-M} K_M^M \omega(1/\mathcal{F}(M)) \leqslant \epsilon$. We thus set $K = K_M$ for the value of M induced by ϵ and $\mathcal{F}$, and so

$$|\mathbb{E}_{y \in \mathbb{Z}/N\mathbb{Z}} \tilde{\mu}(y) \prod_{i=0}^{t} f_i(x + P_i(y))| \leqslant 2\epsilon.$$

$\square$

References

[BHK05] V. Bergelson, B. Host, and B. Kra. Multiple recurrence and nilsequences. *Invent. Math.*, **160**:261–303, 2005. With an appendix by I. Ruzsa.

[BL96] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.*, **9**:725–753, 1996.

[BLL07] V. Bergelson, A. Leibman, and E. Lesigne. Complexities of finite families of polynomials, Weyl systems, and constructions in combinatorial number theory. *J. Anal. Math.*, **103**:47–92, 2007.

[CS12] P. Candela and O. Sisask. Convergence results for systems of linear forms on cyclic groups and periodic nilsequences. *SIAM J. Discrete Math.*, **28**:786–810, 2012.

[FK05] N. Frantzikinakis and B. Kra. Polynomial averages converge to the product of integrals. *Israel J. Math.*, **148**(1):267–276, 2005.

[FK06] N. Frantzikinakis and B. Kra. Ergodic averages for independent polynomials and applications. *J. Lond. Math. Soc.*, **74**:131–142, 2006.

[Fra08] N. Frantzikinakis. Multiple ergodic averages for three polynomials and applications. *Trans. Amer. Math. Soc.*, **360**(10):5435–5475, 2008.

[Fra16] N. Frantzikinakis. Some open problems on multiple ergodic averages. *Bull. Hellenic Math. Soc.*, **60**:41–90, 2016.

[GT10] B. Green and T. Tao. An arithmetic regularity lemma, an associated counting lemma, and applications. *Bolyai Soc. Math. Stud.*, **21**:261–334, 2010.

[GT12] B. Green and T. Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math.*, **175**:465–540, 2012.

[GW10] W. T. Gowers and J. Wolf. The true complexity of a system of linear equations. *Proc. Lond. Math. Soc.*, **100**(1):155–176, 2010.

[GW11a] W. T. Gowers and J. Wolf. Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$. *Geom. Funct. Anal.*, **21**:36–69, 2011.

[GW11b] W. T. Gowers and J. Wolf. Linear forms and quadratic uniformity for functions on $\mathbb{F}_p^n$. *Mathematika*, **57**:215–237, 2011.

[GW11c] W. T. Gowers and J. Wolf. Linear forms and quadratic uniformity for functions on $\mathbb{Z}_N$. *J. Anal. Math.*, **115**(1):121–186, 2011.

[HK05a] B. Host and B. Kra. Convergence of polynomial ergodic averages. *Israel J. Math.*, **149**(1):1–19, 2005.

[HK05b] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math.*, **161**(1):397–488, 2005.

[HK18] B. Host and B. Kra. *Nilpotent structures in ergodic theory*. AMS, 2018.

[Kuc19] B. Kuca. Further quantitative bounds in the polynomial Szemerédi theorem over finite fields. 2019.

[Kuc20] B. Kuca. True complexity of polynomial progressions in finite fields. 2020.

[Lei05a] A. Leibman. Convergence of multiple ergodic averages along polynomials of several variables. *Israel J. Math.*, **146**:303–315, 2005.

[Lei05b] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold. *Ergodic Theory Dynam. Systems*, **25**(1):201–213, 2005.

[Lei09] A. Leibman. Orbit of the diagonal in the power of a nilmanifold. *Trans. Amer. Math. Soc.*, **362**(03):1619–1658, 2009.

[Pel19] S. Peluse. On the polynomial Szemerédi theorem in finite fields. *Duke Math. J.*, **168**(5):749–774, 2019.