POLYNOMIAL AVERAGES CONVERGE TO THE PRODUCT OF INTEGRALS

NIKOS FRANTZIKINAKIS AND BRYNA KRA

Abstract. We answer a question posed by Vitaly Bergelson, showing that in a totally ergodic system, the average of a product of functions evaluated along polynomial times, with polynomials of pairwise differing degrees, converges in L^2 to the product of the integrals. Such averages are characterized by nilsystems and so we reduce the problem to one of uniform distribution of polynomial sequences on nilmanifolds.

1. Introduction

1.1. **Bergelson's Question.** In [B96], Bergelson asked if the average of a product of functions in a totally ergodic system (meaning that each power of the transformation is ergodic) evaluated along polynomial times converges in L^2 to the product of the integrals. More precisely, if $(X, \mathcal{X}, \mu, T)$ is a totally ergodic probability measure preserving system, $p_1, p_2, \ldots, p_k$ are polynomials taking integer values on the integers with pairwise distinct non-zero degrees, and $f_1, f_2, \ldots, f_k \in L^\infty(\mu)$, does

$$\lim_{N\to\infty} \Big\| \frac{1}{N}\sum_{n=0}^{N-1} f_1(T^{p_1(n)}x) f_2(T^{p_2(n)}x)\ldots f_k(T^{p_k(n)}x) - \prod_{i=1}^{k}\int f_i\, d\mu \Big\|_{L^2(\mu)}$$

equal 0?

We show that the answer to this question is positive, under slightly more general assumptions. We start with some definitions in order to precisely state the theorem.

An *integer polynomial* is a polynomial taking integer values on the integers. A family of integer polynomials $\{p_1(n), p_2(n), \ldots, p_k(n)\}$ is said to be *independent* if for all integers $m_1, m_2, \ldots, m_k$ with at least some $m_j \neq 0$, $j \in \{1, 2, \ldots, k\}$, the polynomial $\sum_{j=1}^{k} m_j p_j(n)$ is not constant.

We prove:

The second author was partially supported by NSF grant DMS-0244994.

Theorem 1.1. *Let $(X, \mathcal{X}, \mu, T)$ be a totally ergodic measure preserving probability system and assume that $\{p_1(n), p_2(n), \dots, p_k(n)\}$ is an independent family of polynomials. Then for $f_1, f_2, \dots, f_k \in L^\infty(\mu)$,*

$$\lim_{N\to\infty} \Big\| \frac{1}{N}\sum_{n=0}^{N-1} f_1(T^{p_1(n)}x) f_2(T^{p_2(n)}x) \dots f_k(T^{p_k(n)}x) - \prod_{i=1}^{k} \int f_i \, d\mu \Big\|_{L^2(\mu)} \tag{1}$$

equals 0.

The assumption that the polynomial family is independent is necessary, as can be seen by considering an irrational rotation on the circle. An ergodic rotation on a finite group with at least two elements demonstrates that the hypothesis of total ergodicity is necessary; in this example, the average for any independent family with $k > 1$ polynomials does not converge to the product of the integrals for appropriate choice of the functions f_i.

If one assumes that T is weakly mixing, Bergelson [B87] showed that for all polynomial families, the limit in (1) exists and is constant. However, without the assumption of weak mixing one can easily show that the limit need not be constant, even when restricting to polynomials of degree one. For the polynomial families (n, n^2) and (n^2, n^2+n), the convergence to the product of the integrals was proved by Furstenberg and Weiss [FW96]. The existence of the limit in a totally ergodic system for an arbitrary family of integer polynomials was shown in Host and Kra [HK02], but further analysis is needed to describe the form of the limit.

1.2. Reduction to a problem of uniform distribution.

In [HK02], Host and Kra showed that for any family of polynomials, the characteristic factor of the average in (1) in a totally ergodic system is an inverse limit of nilsystems. We need a few definitions to make this statement precise.

Given a group G, we denote the commutator of $g, h \in G$ by $[g, h] = g^{-1}h^{-1}gh$. If $A, B \subset G$, then $[A, B]$ is defined to be $\{[a, b] : a \in A, b \in B\}$. A group G is said to be *k-step nilpotent* if its $(k + 1)$ commutator $[G, G^{(k)}]$ is trivial. If G is a k-step nilpotent Lie group and Γ is a discrete cocompact subgroup, then the compact space $X = G/\Gamma$ is said to be a *k-step nilmanifold*. The group G acts on G/Γ by left translation and the translation by a fixed element $a \in G$ is given by $T_a(g\Gamma) = (ag)\Gamma$. Let μ denote the unique probability measure on X that is invariant under the action of G by left translations (called the *Haar measure*) and let $\mathcal{G}/\Gamma$ denote the Borel σ-algebra of G/Γ. Fixing an element

$a \in G$, we call the system $(G/\Gamma, \mathcal{G}/\Gamma, \mu, T_a)$ a *k-step nilsystem* and call the map T_a a *nilrotation.*

A *factor* of the measure preserving system $(X, \mathcal{X}, \mu, T)$ is a measure preserving system $(Y, \mathcal{Y}, \nu, S)$ so that there exists a measure preserving map $\pi : X \to Y$ taking μ to ν and such that $S \circ \pi = \pi \circ T$. In a slight abuse of terminology, when the underlying measure space is implicit we call S a factor of T.

In this terminology, Host and Kra's result means that there exists a factor $(Z, \mathcal{Z}, m)$ of X, where $\mathcal{Z}$ denotes the Borel σ-algebra of Z and m its Haar measure, so that the action of T on Z is an inverse limit of nilsystems and furthermore, whenever $\mathbb{E}(f_j|\mathcal{Z}) = 0$ for some $j \in \{1, 2, \ldots, k\}$, the average in (1) is itself 0. Since an inverse limits of nilsystems can be approximated arbitrarily well by a nilsystem, it suffices to verify Theorem 1.1 for nilsystems. Moreover, since measurable functions can be approximated arbitrarily well in L^2 by continuous functions, Theorem 1.1 is equivalent to the following generalization of Weyl's polynomial uniform distribution theorem (see Section 4 for the statement of Weyl's Theorem):

Theorem 1.2. *Let $X = G/\Gamma$ be a nilmanifold, $(G/\Gamma, \mathcal{G}/\Gamma, \mu, T_a)$ a nilsystem and suppose that the nilrotation T_a is totally ergodic. If $\{p_1(n), p_2(n), \ldots, p_k(n)\}$ is an independent polynomial family, then for almost every $x \in X$ the sequence $(a^{p_1(n)}x, a^{p_2(n)}x, \ldots, a^{p_k(n)}x)$ is uniformly distributed in X^k.*

If G is connected, we can reduce Theorem 1.2 to a uniform distribution problem that is easily verified using the standard uniform distribution theorem of Weyl. The general (not necessarily connected) case is more subtle. Using a result of Leibman [L02], in Section 2, we reduce the problem to studying the action of a polynomial sequence on a factor space with abelian identity component. The key step (Section 3) is then to prove that nilrotations acting on such spaces are isomorphic to affine transformations on some finite dimensional torus. In Section 4, we complete the proof by checking the result for affine transformations.

2. Reduction to an abelian connected component

Suppose that G is a nilpotent Lie group and Γ is a discrete, cocompact subgroup. Throughout, we let G_0 denote the connected component of the identity element and denote the identity element by e.

A sequence $g(n) = a_1^{p_1(n)} a_2^{p_2(n)} \ldots a_k^{p_k(n)}$ with $a_1, a_2, \ldots, a_k \in G$ and $p_1, p_2, \ldots, p_k$ integer polynomials is called a *polynomial sequence* in G. We are interested in studying uniform distribution properties of polynomial sequences on the nilmanifold $X = G/\Gamma$.

Leibman [L02] showed that the uniform distribution of a polynomial sequence in a connected nilmanifold reduces to uniform distribution in a certain factor:

Theorem. [Leibman] *Let $X = G/\Gamma$ be a connected nilmanifold and let $g(n) = a_1^{p_1(n)} a_2^{p_2(n)} \ldots a_k^{p_k(n)}$ be a polynomial sequence in G. Let $Z = X/[G_0, G_0]$ and let $\pi\colon X \to Z$ be the natural projection. If $x \in X$ then $\{g(n)x\}_{n\in\mathbb{Z}}$ is uniformly distributed in X if and only if $\{g(n)\pi(x)\}_{n\in\mathbb{Z}}$ is uniformly distributed in Z.*

We remark that if G is connected, then the factor $X/[G_0, G_0]$ is an abelian group. However, this does not hold in general as the following examples illustrate:

Example 1. On the space $G = \mathbb{Z}\times\mathbb{R}^2$, define multiplication as follows: if $g_1 = (m_1, x_1, x_2)$ and $g_2 = (n_1, y_1, y_2)$, let

$$g_1 \cdot g_2 = (m_1 + n_1, x_1 + y_1, x_2 + y_2 + m_1 y_1).$$

Then G is a 2-step nilpotent group and $G_0 = \{0\} \times \mathbb{R}^2$ is abelian. The discrete subgroup $\Gamma = \mathbb{Z}^3$ is cocompact and $X = G/\Gamma$ is connected. Moreover, $[G_0, G_0] = \{\mathbf{e}\}$ and so $X/[G_0, G_0] = X$.

Example 2. On the space $G = \mathbb{Z}\times\mathbb{R}^3$, define multiplication as follows: if $g_1 = (m_1, x_1, x_2, x_3)$ and $g_2 = (n_1, y_1, y_2, y_3)$, let

$$g_1 \cdot g_2 = (m_1 + n_1, x_1 + y_1, x_2 + y_2 + m_1 y_1, x_3 + y_3 + m_1 y_2 + \frac{1}{2} m_1^2 y_1).$$

Then G is a 3-step nilpotent group and $G_0 = \{0\} \times \mathbb{R}^3$ is abelian. The discrete subgroup $\Gamma = \mathbb{Z}^3 \times (\mathbb{Z}/2)$ is cocompact and $X = G/\Gamma$ is connected. Again, $X/[G_0, G_0] = X$.

We use Leibman's theorem to reduce the problem on uniform distribution to the case that G_0 is abelian:

Proposition 2.1. *Theorem 1.2 follows if it holds for all nilsystems $(G/\Gamma, \mathcal{G}/\Gamma, \mu, T_a)$ with G_0 abelian and T_a totally ergodic.*

Proof. Given $a \in G$ and $x \in X = G/\Gamma$, let $a_1 = (a, e, \ldots, e), a_2 = (e, a, e, \ldots, e), \ldots, a_k = (e, e, \ldots, a) \in G^k$, $\tilde{x} = (x, \ldots, x) \in X^k$, and $g(n) = T_{a_1}^{p_1(n)} T_{a_2}^{p_2(n)} \cdots T_{a_k}^{p_k(n)}$. We need to check that for μ-a.e. $x \in X$ the polynomial sequence $g(n)\tilde{x}$ is uniformly distributed in X^k. By Leibman's Theorem, it suffices to check that $g(n)\pi(\tilde{x})$ is uniformly distributed in the nilmanifold Z^k, where $Z = X/[G_0, G_0]$ and $\pi\colon G \to G/[G_0, G_0]$ is the natural projection. Since $(G/[G_0, G_0])_0$ is abelian and a factor of a totally ergodic system is totally ergodic, the statement follows. $\square$

3. Reduction to an affine transformation on a torus

We reduce the problem on uniform distribution (Theorem 1.2) to studying an affine transformation on a torus. If G is a group then a map $T\colon G \to G$ is said to be *affine* if $T(g) = bA(g)$ for an endomorphism A of G and some $b \in G$. The endomorphism A is said to be *unipotent* if there exists $n \in \mathbb{N}$ so that so that $(A - \mathrm{I}d)^n = 0$. In this case we say that the affine transformation T is a unipotent affine transformation.

Proposition 3.1. *Let $X = G/\Gamma$ be a connected nilmanifold such that G_0 is abelian. Then any nilrotation $T_a(x) = ax$ defined on X with the Haar measure μ is isomorphic to a unipotent affine transformation on some finite dimensional torus.*

Proof. First observe that for every $g \in G$, the subgroup $g^{-1}G_0g$ is both open and closed in G so $g^{-1}G_0g = G_0$. Hence, G_0 is a normal subgroup of G. Similarly, since $G_0\Gamma$ is both open and closed in G, we have that $(G_0\Gamma)/\Gamma$ is open and closed in X. Since X is connected, $X = (G_0\Gamma)/\Gamma$ and so $G = G_0\Gamma$.

We claim that $\Gamma_0 = \Gamma \cap G_0$ is a normal subgroup of G. Let $\gamma_0 \in \Gamma_0$ and $g = g_0\gamma$, where $g_0 \in G_0$ and $\gamma \in \Gamma$. Since G_0 is normal in G, we have that $g^{-1}\gamma_0 g \in G_0$. Moreover,

$$g^{-1}\gamma_0 g = \gamma^{-1}g_0^{-1}\gamma_0 g_0\gamma = \gamma^{-1}\gamma_0\gamma \in \Gamma,$$

the last equality being valid since G_0 is abelian. Hence, $g^{-1}\gamma_0 g \in \Gamma_0$ and Γ_0 is normal in G.

Therefore we can substitute G/Γ_0 for G and Γ/Γ_0 for Γ; then $X = (G/\Gamma_0)/(\Gamma/\Gamma_0)$. So we can assume that $G_0 \cap \Gamma = \{e\}$. Note that we now have that G_0 is a connected compact abelian Lie group and so is isomorphic to some finite dimensional torus $\mathbb{T}^d$.

Every $g \in G$ is uniquely representable in the form $g = g_0\gamma$, with $g_0 \in G_0$, $\gamma \in \Gamma$. The map $\phi\colon X \to G_0$, given by $\phi(g\Gamma) = g_0$ is a well defined homeomorphism. Since $\phi(hg\Gamma) = h\phi(g\Gamma)$ for any $h \in G_0$, the measure $\phi(\mu)$ on G_0 is invariant under left translations. Thus $\phi(\mu)$ is the Haar measure on G_0. If $a = a_0\gamma$, $g = g_0\gamma'$ with $a_0, g_0 \in G_0$ and $\gamma, \gamma' \in \Gamma$, then $ag\Gamma = a_0\gamma g_0\gamma^{-1}\Gamma$. Since $\gamma g_0\gamma^{-1} \in G_0$, we have that $\phi(ag\Gamma) = a_0\gamma g_0\gamma^{-1}$. Hence ϕ conjugates T_a to $T'_a\colon G_0 \to G_0$ defined by

$$T'_a(g_0) = \phi T_a\phi^{-1} = a_0\gamma g_0\gamma^{-1}.$$

Since G_0 is abelian this is an affine map; its linear part $g_0 \mapsto \gamma g_0\gamma^{-1}$ is unipotent since G is nilpotent. Letting $\psi\colon G_0 \to \mathbb{T}^d$ denote the isomorphism between G_0 and $\mathbb{T}^d$, we have that T_a is isomorphic to the unipotent affine transformation $S = \psi T'_a\psi^{-1}$ acting on $\mathbb{T}^d$. $\square$

We illustrate this with the examples of the previous section:

Example 3. Let X be as in Example 1 and let $a = (m_1, a_1, a_2)$. Since $G_0/\Gamma_0 = \mathbb{T}^2$ we see that T_a is isomorphic to the unipotent affine transformation $S\colon \mathbb{T}^2 \to \mathbb{T}^2$ given by

$$S(x_1, x_2) = (x_1 + a_1, x_2 + m_1x_1 + a_2).$$

Example 4. Let X be as in Example 2 and $a = (m_1, a_1, a_2, a_3)$. Since $G_0/\Gamma_0 = \mathbb{R}^3/(\mathbb{Z}^2\times\mathbb{Z}/2)$, and $\psi\colon G_0/\Gamma_0 \to \mathbb{T}^3$ defined by $\psi(x_1, x_2, x_3) = (x_1, x_2, 2x_3)$ is an isomorphism, we see that T_a is isomorphic to the unipotent affine transformation $S\colon \mathbb{T}^3 \to \mathbb{T}^3$ given by

$$S(x_1, x_2, x_3) = (x_1 + a_1, x_2 + m_1x_1 + a_2, x_3 + 2m_1x_2 + m_1^2x_1 + 2a_3).$$

Proposition 3.2. *Theorem 1.2 follows if it holds for all nilsystems $(G/\Gamma, \mathcal{G}/\Gamma, \mu, T_a)$ such that T_a is isomorphic to an ergodic, unipotent, affine transformation on some finite dimensional torus.*

Proof. We first note that since $X = G/\Gamma$ admits a totally ergodic nilrotation T_a, it must be connected. Indeed, let X_0 be the identity component of X. Since X is compact, it is a disjoint union of d copies of translations of X_0 for some $d \in \mathbb{N}$. Since a permutes these copies, a^d preserves X_0. By assumption the translation by $T_{a^d} = T_a^d$ is ergodic and so $X_0 = X$.

By Proposition 2.1 we can assume that G_0 is abelian. Since X is connected, the result follows from Proposition 3.1. □

4. Uniform distribution for an affine transformation

We are left with showing that Theorem 1.2 holds when the nilsystem is isomorphic to an ergodic, unipotent, affine system on a finite dimensional torus. Before turning into the proof, note that if G is connected then the uniform distribution property of Theorem 1.2 holds for every $x \in X$. However, this does not hold in general. We illustrate this with the following example:

Example 5. We have seen that the nilrotation of Example 1 is isomorphic to the affine transformation $S\colon \mathbb{T}^2 \to \mathbb{T}^2$ given by

$$S(x_1, x_2) = (x_1 + a_1, x_2 + m_1x_1 + a_2).$$

If $m_1 = 2$ and $a_1 = a_2 = a$ is irrational then S is totally ergodic and $S^n(x_1, x_2) = (x_1 + na, x_2 + 2nx_1 + n^2a)$. Then

$$\big(S^n(0,0), S^{n^2}(0,0)\big) = (na, n^2a, n^2a, n^4a)$$

is not uniformly distributed on $\mathbb{T}^4$. On the other hand

$$\big(S^n(x_1,x_2), S^{n^2}(x_1,x_2)\big) = \\ (x_1+na, x_2+2nx_1+n^2a, x_1+n^2a, x_2+2n^2x_1+n^4a,)$$

is uniformly distributed on $\mathbb{T}^4$ as long as a and x_1 are rationally independent.

The main tool used in the proof of Theorem 1.2 is the following classic theorem of Weyl [W16] on uniform distribution:

Theorem. [Weyl] *(i) Let $a_n \in \mathbb{R}^d$. Then a_n is uniformly distributed in $\mathbb{T}^d$ if and only if*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi i m\cdot a_n} = 0$$

for every nonzero $m \in \mathbb{Z}^d$, where $m \cdot a_n$ denotes the inner product of m and a_n.

(ii) If $a_n = p(n)$ where p is a real valued polynomial with at least one nonconstant coefficient irrational then

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi i a_n} = 0.$$

Before turning to the proof of Theorem 1.2, we prove a lemma that simplifies the computations:

Lemma 4.1. *Let $T\colon \mathbb{T}^d \to \mathbb{T}^d$ be defined by $T(x) = Ax + b$, where A is a $d \times d$ unipotent integer matrix and $b \in \mathbb{T}^d$. Assume furthermore that T is ergodic. Then T is a factor of an ergodic affine transformation $S\colon \mathbb{T}^d \to \mathbb{T}^d$, where $S = S_1 \times S_2 \times \cdots \times S_s$ and for $r = 1, 2, \dots, s$, $S_r\colon \mathbb{T}^{d_r} \to \mathbb{T}^{d_r}$ ($\sum_{r=1}^{s} d_r = d$) has the form*

$$S_r(x_{r1}, x_{r2}, \dots, x_{rd_r}) = (x_{r1} + b_r, x_{r2} + x_{r1}, \dots, x_{rd_r} + x_{rd_r-1})$$

for some $b_r \in \mathbb{T}$.

Proof. Let J be the Jordan canonical form of A with Jordan blocks J_r of dimension d_r for $r = 1, 2, \dots, s$. Since A is unipotent, all diagonal entries of J are equal to 1. There exists a matrix P with rational entries such that $PA = JP$. After multiplying P by an appropriate integer, we can assume that it too has integer entries. So P defines an endomorphism $P\colon \mathbb{T}^d \to \mathbb{T}^d$ such that $PT = SP$, where $S\colon \mathbb{T}^d \to \mathbb{T}^d$ is given by $S(x) = J(x) + c$ for $c = P(b)$. Hence, T is a factor of S. By making a change of variables $x \to x + a$ for some suitable $a \in \mathbb{T}^d$, we can assume that S has the advertised form.

It remains to show that S is ergodic. Since J is unipotent, using a theorem of Hahn ([H63], Theorem 4) we get that ergodicity of S is equivalent to showing that for every nontrivial character χ in the dual of $\mathbb{T}^d$ we have the implication

$$\chi(Jx) = \chi(x) \text{ for every } x \in \mathbb{T}^d \Rightarrow \chi(c) \neq 1.$$

Suppose that $\chi(Jx) = \chi(x)$. Using the relation $PA = JP$ we get that $\chi'(Ax) = \chi'(x)$ where $\chi'(x) = \chi(Px)$. Since $T(x) = Ax + b$ is assumed to be ergodic, again using Hahn's theorem we get that $\chi'(b) \neq 1$. The relation $PA = JP$ implies that $\chi(c) \neq 1$ and the proof is complete. □

Proof of Theorem 1.2. By Proposition 3.2 it suffices to verify the uniform distribution property for all ergodic, unipotent, affine transformations on $\mathbb{T}^d$. First observe that relation (1) of Theorem 1.1 is preserved when passing to factors. Hence, using Lemma 4.1 we can assume that $T = T_1 \times T_2 \times \cdots \times T_s$, where $T_r\colon \mathbb{T}^{d_r} \to \mathbb{T}^{d_r}$ ($\sum_{r=1}^{s} d_r = d$) is given by

$$T_r(x_{r1}, x_{r2}, \ldots, x_{rd_r}) = (x_{r1} + b_r, x_{r2} + x_{r1}, \ldots, x_{rd_r} + x_{rd_r-1}),$$

for $r = 1, 2, \ldots, s$. Since T is ergodic the set $\{b_1, b_2, \ldots, b_s\}$ is rationally independent. For convenience, set $x_{r0} = b_r$ for $r = 1, 2, \ldots s$.

We claim that if x is chosen so that the set $C = \{x_{rj} : 1 \leq r \leq s, 0 \leq j \leq d_r\}$ is rationally independent, then the polynomial sequence $g(n)\tilde{x} = (T^{p_1(n)}x, T^{p_2(n)}x, \ldots, T^{p_k(n)}x)$ is uniformly distributed on $\mathbb{T}^{dk}$ (we include x_{rd_r} in C only for simplicity). To see this we use the first part of Weyl's theorem; letting $Q_{rjl}(n)$ denote the j-th coordinate of $T_r^{p_l(n)}x$ and

$$R(n) = \sum_{r,j,l} m_{rjl} Q_{rjl}(n) \tag{2}$$

where $\{m_{rjl} : 1 \leq r \leq s, 1 \leq j \leq d_r, 1 \leq l \leq k\}$ are integers, not all of them zero, it suffices to check that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i R(n)} = 0. \tag{3}$$

To prove (3) we use the second part of Weyl's theorem; it suffices to show that the polynomial $R(n)$ has at least one nonconstant coefficient irrational. We compute

$$Q_{rjl}(n) = x_{rj} + \binom{p_l(n)}{1} x_{rj-1} + \cdots + \binom{p_l(n)}{j-1} x_{r1} + \binom{p_l(n)}{j} x_{r0}. \tag{4}$$

We can put $R(n)$ in the form

$$R(n) = \sum_{r,j} R_{rj}(n) x_{rj}, \tag{5}$$

where R_{rj} are integer polynomials and $1 \le r \le s$, $0 \le j \le d_r$. This representation is unique since the x_{rj} are rationally independent. So it remains to show that some R_{rj} is nonconstant. To see this, choose any r_0 such that $m_{r_0 jl} \neq 0$ for some j, l, and define j_0 to be the maximum $1 \le j \le d_{r_0}$ such that $m_{r_0 jl} \neq 0$ for some $1 \le l \le k$. We show that $R_{r_0, j_0 - 1}$ is nonconstant. By the definition of j_0 we have $m_{r_0 jl} = 0$ for $j > j_0$. For $j \le j_0$ we see from (4) that the variable $x_{r_0 j_0 - 1}$ appears only in the polynomials $Q_{r_0 j_0 l}$ with coefficient $p_l(n)$, and if $j_0 > 1$ also in the polynomials $Q_{r_0 (j_0 - 1) l}$ with coefficient 1. It follows from (2) and (5) that

$$R_{r_0 j_0 - 1}(n) = \sum_{l=1}^{k} m_{r_0 j_0 l} p_l(n) + c,$$

where $c = \sum_{l=1}^{k} m_{r_0 j_0 l}$ if $j_0 > 1$, and $c = 0$ if $j_0 = 1$. Since the polynomial family $\{p_i(n)\}_{i=1}^{k}$ is independent and $m_{r_0 j_0 l} \neq 0$ for some l, the polynomial $R_{r_0 j_0 - 1}$ is nonconstant. We have thus established uniform distribution for a set of x of full measure, completing the proof. $\square$

Acknowledgment: The authors thank the referee for his help in organizing and simplifying the presentation, and in particular for the simple proof of Proposition 3.1.

References

[B87] V. Bergelson. Weakly mixing PET. *Erg. Th. & Dyn. Sys.*, **7** (1987), 337-349.

[B96] V. Bergelson. Ergodic Ramsey theory an update. *Ergodic Theory of $\mathbb{Z}^d$-actions*, Eds.: M. Pollicott, K. Schmidt. Cambridge University Press, Cambridge (1996), 1-61.

[FW96] H. Furstenberg and B. Weiss. A mean ergodic theorem for $\frac{1}{N}\sum_{n=1}^{n} f(T^n x) g(T^{n^2} x)$. *Convergence in Ergodic Theory and Probability*, Eds.: V. Bergelson, P. March, J. Rosenblatt. Walter de Gruyter & Co, Berlin (1996), 193-227.

[H63] F. J. Hahn. On affine transformations of compact abelian groups. *Amer. J. Math.*, **85**, No. 3, (1963), 428-446.

[HK02] B. Host and B. Kra. Convergence of polynomial ergodic averages. To appear, *Isr. J. Math.*

[L02] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of rotations of a nilmanifold. Preprint, 2002.

[W16] H. Weyl. Über die Gleichverteilung von Zahlen mod Eins. *Math. Ann.*, **77** (1916), 313-352.

Department of Mathematics, McAllister Building, The Pennsylvania State University, University Park, PA 16802

E-mail address: nikos@math.psu.edu

E-mail address: kra@math.psu.edu