NOTES ON ERGODIC AVERAGES WITH POLYNOMIAL ITERATES

By Andreas Koutsogiannis

1. Introduction

In these notes we will present some recent results in ergodic theory that have to do with ergodic averages. More specifically, we will mainly deal with averages for a single transformation T with polynomial iterates, i.e., averages of the form

$$\frac{1}{N}\sum_{n=1}^{N} T^{[p_1(n)]}f_1 \cdot \ldots \cdot T^{[p_k(n)]}f_k, \tag{1}$$

where p_i's are real polynomials (see below for quantifiers).

Our goal is to study the L^2-convergence (norm-convergence) of (1), with iterates coming from appropriate polynomial families, in order to obtain applications to other areas of mathematics as combinatorics, number theory, topological dynamics etc.

1.1. **History of the problem.** We always work on a *measure preserving system*, i.e., a quadruple $(X, \mathcal{B}, \mu, T)$ where X is a set, $\mathcal{B}$ is a σ-algebra on X, μ is a probability measure on $\mathcal{B}$, and T is an invertible (this assumption sometimes can be skipped but for reasons of simplicity we will always assume it in these notes) measure preserving transformation, i.e., $\mu(T^{-1}A) = \mu(A)$ for all $A \in \mathcal{B}$.

The simplest average that one can study is

$$\frac{1}{N}\sum_{n=1}^{N} T^n f, \tag{2}$$

where $f \in L^2$ (note that in the following, whenever we deal with multiple terms, the assumption for our functions would be that these are elements of L^∞), $Tf(x) := f(Tx)$, while T^n is the composition of T with itself n times.

The first result in understanding (2) is due to von Neumann:

Theorem 1.1 (von Neumann, 1932). *Under the previous assumptions, for every $f \in L^2$ we have that*

$$\frac{1}{N}\sum_{n=1}^{N} T^n f \to Pf,$$

as $N \to \infty$, where P is the orthogonal projection to the space of the left T-invariant functions $\{f : Tf = f\}$ and the convergence takes place in L^2.

The proof of this result is a simple, and nowadays classical, splitting of the Hilbert space L^2 and it is elementary. We also have that

$$Pf = \int f \, d\mu \quad \text{if and only if } \ T \ \text{ is } \textit{ergodic},$$

meaning that the only T-invariant sets of $\mathcal{B}$ are the ones with trivial measure, i.e., in $\{0,1\}$.

Of course, someone can study more complicated expressions, like the following multiple average:

$$\frac{1}{N}\sum_{n=1}^{N} T^n f_1 \cdot T^{2n} f_2 \cdot \ldots \cdot T^{kn} f_k, \tag{3}$$

where $k \in \mathbb{N}$ and $f_1, \ldots, f_k \in L^\infty$.

The study of the precise form of the L^2-limit of (3) proved to be a hard problem. Relatively recently, Host and Kra managed to show, in [13], the existence of the limit providing simultaneously a closed form of it which is rather complicated to be stated here (it has to do with conditional expectations on nilfactors which we won't cover in full detail in these short notes).

One may wonder why we are interested in studying the behavior of (3). The answer mainly lies in the following result that Furstenberg got, using ergodic theoretical methods, by studying (3):

Theorem 1.2 (Furstenberg, 1977, [11]). *Under the standard assumptions, for all $A \in \mathcal{B}$ with $\mu(A) > 0$, we have that*

$$\liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu(A \cap T^{-n}A \cap T^{-2n}A \cap \ldots \cap T^{-kn}A) > 0. \tag{4}$$

This breakthrough made the ergodic theory to blossom and gave numerous deep and interesting applications to many areas of Mathematics (see below its connection with Szemerédi's celebrated theorem).

We remark that the lim inf that appears in (4) is actually a limit by [13]. Note that the connection of (4) with (3) can be reflected by the relations $T\mathbf{1}_A = \mathbf{1}_{T^{-1}A}$, and $\int T^n \mathbf{1}_A \, d\mu = \mu(T^{-n}A)$.

Via Furstenberg's *correspondence principle* (see below) we will get one of the most deep and interesting results about the set of natural numbers, namely Szemerédi's theorem, which states that every "large" subset of natural numbers is *AP rich*, i.e., contains arbitrarily long arithmetic progressions.

Before we state Furstenberg's correspondence principle (actually, we will state a reformulation of it which is due to Bergelson [1]) we define the motion of *upper density* of a subset of natural numbers:

Definition. Let $E \subseteq \mathbb{N}$. We define the *upper density* of E to be

$$\bar{d}(E) := \limsup_{N\to\infty} \frac{|E \cap \{1, \ldots, N\}|}{N}$$

which is a number between 0 and 1.

(The quantity $|E \cap \{1, \ldots, N\}|$ measures how many elements from the first N natural numbers the set E captures.)

Theorem 1.3 (Furstenberg correspondence principle, [11], [1]). *For any $E \subseteq \mathbb{N}$, there exists a system $(X, \mathcal{B}, \mu, T)$ and a set $A \in \mathcal{B}$ with $\mu(A) = \bar{d}(E)$ such that*

$$\bar{d}(E \cap (E - n_1) \cap \ldots \cap (E - n_k)) \geq \mu(A \cap T^{-n_1}A \cap \ldots \cap T^{-n_k}A) \tag{5}$$

for all $k \in \mathbb{N}$ and $n_1, \ldots, n_k \in \mathbb{N}$.

Note that by starting with a set $E \subseteq \mathbb{N}$ with $\bar{d}(E) > 0$, by Theorem 1.3 we can find a system $(X, \mathcal{B}, \mu, T)$ and a set $A \in \mathcal{B}$ with $\mu(A) = \bar{d}(E) > 0$. For an arbitrary $k \in \mathbb{N}$, using Theorem 1.2, we can find $n_0 \in \mathbb{N}$ such that

$$\mu(A \cap T^{-n_0}A \cap \ldots \cap T^{-kn_0}A) > 0,$$

so, by Theorem 1.3 we have that

$$\bar{d}(E \cap (E - n_0) \cap \ldots \cap (E - kn_0)) > 0.$$

In particular,

$$E \cap (E - n_0) \cap \ldots \cap (E - kn_0) \neq \emptyset,$$

hence, there exists $x_0 \in E$ such that

$$x_0, x_0 + n_0, \ldots, x_0 + kn_0 \in E.$$

Summing up the previous arguments, we have reproved the following:

Theorem 1.4 (Szemerédi, 1975). *Every subset of natural numbers with positive upper density contains arbitrarily long arithmetic progressions.*

In order to state the next result, we have to recall the notion of a weakly mixing system:

Definition. Let $(X, \mathcal{B}, \mu, T)$ be a system. T is called *weakly mixing*, which we will denote with w.m., (and the whole system $(X, \mathcal{B}, \mu, T)$ is called *weakly mixing*) if for all $f, g \in L^2$ we have that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left| \int T^n f \cdot g \, d\mu - \int f \, d\mu \cdot \int g \, d\mu \right| = 0.$$

We remark at this point that a w.m. transformation is ergodic while the opposite is not in general true.

In the same paper, [11], under the weakly mixing assumption of T Furstenberg showed the following convergence result:

Theorem 1.5 (Furstenberg, [11]). *If $(X, \mathcal{B}, \mu, T)$ is a w.m. system, then for any $k \in \mathbb{N}$ and $f_1, \ldots, f_k \in L^\infty$ we have*

$$\frac{1}{N} \sum_{n=1}^{N} T^n f_1 \cdot T^{2n} f_2 \cdot \ldots \cdot T^{kn} f_k \to \prod_{i=1}^{k} \int f_i \, d\mu, \tag{6}$$

as $N \to \infty$, where the convergence takes place in L^2.

1.2. **From linear iterates to polynomial ones.** Bergelson was the first to view the iterates $n, 2n, \ldots, kn$ as linear polynomials $p_1, \ldots, p_k$ with the property $p_i - p_j \neq$ constant for all $i \neq j$.

This naturally led him to the following definition for more general, than linear, polynomials.

Definition. The non-constant polynomials $p_1(t), \ldots, p_k(t)$ in $\mathbb{Z}[t]$ are called *essentially distinct* if $p_i - p_j \neq$ constant for all $i \neq j$.

Exploted the van der Corput trick, which we will see below, Bergelson showed the following:

Theorem 1.6 (Bergelson, [2]). *If $(X, \mathcal{B}, \mu, T)$ is a w.m. system, then for any $k \in \mathbb{N}$, $p_1(t), \ldots, p_k(t)$ essentially distinct polynomials in $\mathbb{Z}[t]$ and $f_1, \ldots, f_k \in L^\infty$ we have*

$$\frac{1}{N}\sum_{n=1}^{N} T^{p_1(n)}f_1 \cdot \ldots \cdot T^{p_k(n)}f_k \to \prod_{i=1}^{k} \int f_i \, d\mu, \tag{7}$$

as $N \to \infty$, where the convergence takes place in L^2.

To show this, Bergelson used the following reformulation, due to himself, of van der Corput trick:

Lemma 1.7 (van der Corput, Bergelson, [2]). *Let $(x_n)_n$ be a bounded sequence in a Hilbert space and suppose that for any $h \geq h_0$ (for some fixed, large, $h_0 \in \mathbb{N}$) we have*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \langle x_n, x_{n+h}\rangle = 0, \tag{8}$$

then

$$\lim_{N\to\infty} \left\| \frac{1}{N}\sum_{n=1}^{N} x_n \right\| = 0.$$

Remark. To study (7) via van der Corput trick, we assume without loss of generality that some $\int f_i \, d\mu = 0$, we set

$$x_n := T^{p_1(n)}f_1 \cdot \ldots \cdot T^{p_k(n)}f_k$$

and we try to show that for large enough h we have (8).

This is achieved by induction (which is formally called *PET induction*) on the "complexity" of the polynomial family since the quantity $\langle x_n, x_{n+h}\rangle$ leads to differences (i.e., derivatives) hence to reduction of the complexity.

Ten years after Theorem 1.6, studying ergodic averages with polynomial iterates, Bergelson and Leibman, in [4], obtained far-reaching polynomial multidimensional (in $\mathbb{Z}^d$) extensions of Szemerédi's theorem (showing existence of "polynomial progressions" in any "large" subset of $\mathbb{Z}^d$). They managed to show the corresponding polynomial relation to (4) and they also stated the following conjecture:

Conjecture ([4]). Let $k \in \mathbb{N}$. For any measure preserving system $(X, \mathcal{B}, \mu, T_1, \ldots, T_k)$, where each T_i is measure preserving and $T_iT_j = T_jT_i$, and all polynomials $p_1(t), \ldots, p_k(t)$ in $\mathbb{Z}[t]$ the following expression

$$\frac{1}{N}\sum_{n=1}^{N} T_1^{p_1(n)} f_1 \cdot \ldots \cdot T_k^{p_k(n)} f_k$$

has a limit, in L^2, as $N \to \infty$ for all $f_1, \ldots, f_k \in L^\infty$.

This conjecture was answered positively in stages. Walsh was the one that eventually showed:

Theorem 1.8 (Walsh, [18]). *The conjecture of Bergelson-Leibman holds true.*

He actually showed that this is true for products of transformations, with any polynomial power, which more generally we don't have to assume that they commute but that they produce a nilpotent group.

Of course we have no information of the limit function in question.

In these notes we will solely deal with a single transformation. We will find a class of polynomials for which we have convergence in $\prod_{i=1}^{k} \int f_i \, d\mu$ with no assumption on the system. This lack of assumptions makes our result, even though it is stated with ergodic theory language, a combinatorial object. Also, the strong nature of the result will provide us with many interesting applications in other areas of mathematics.

2. Passing to integer parts - Main result

Someone of course can think of extending the polynomial iterates to iterates of the form $[p(n)]$, where $[x]$ denotes the integer part function, or floor function at $x \in \mathbb{R}$ which gives the closest integer which is less of equal to x and $p(t)$ is a real polynomial.

The transition from $p(t) \in \mathbb{Z}[t]$ to $[q(t)]$ for $q(t) \in \mathbb{R}[t]$ is not immediate, since PET induction is not immediately applicable. This is mainly due to error terms that appear in the corresponding differences $\langle x_n, x_{n+h} \rangle$ (recall that $[x]-[y] = [x-y]+e$, where $e \in \{0, 1\}$).

Following the work of Lesigne (for one term) and the work of Wierdl (for two terms) one can show that the expression

$$T^{[a_d n^d + \ldots + a_1 n + a_0]}$$

"looks" like

$$S_{a_d n^d + \ldots + a_1 n + a_0} = S_{a_d}^{n^d} \cdots S_{a_1}^{n} S_{a_0},$$

where S is the suspension flow (with respect to T). This aforementioned "looks like" statement that we mentioned above hides a periodic property that polynomials with no non-constant irrational coefficients and an equidistribution property that polynomials with some non-constant irrational coefficient have (for more details see [15]).

So, extending Lesigne's and Wierdl's argument for an arbitrary number of terms, using Walsh's result (the one about products of transformations) we get:

Theorem 2.1 (K, [15]). *Let $k \in \mathbb{N}$. For any measure preserving system $(X, \mathcal{B}, \mu, T_1, \dots, T_k)$, where each T_i is measure preserving and $T_iT_j = T_jT_i$, and all polynomials $p_1(t), \dots, p_k(t)$ in $\mathbb{R}[t]$ the expression*

$$\frac{1}{N}\sum_{n=1}^{N} T_1^{[p_1(n)]}f_1 \cdot \ldots \cdot T_k^{[p_k(n)]}f_k \tag{9}$$

has a limit as $N \to \infty$, in L^2, as $N \to \infty$ for all $f_1, \dots, f_k \in L^\infty$.

We remark at this point that we actually have the same result (i.e., L^2-convergence of the corresponding Equation (9)) for products of transformations for all real polynomials.

Hence, from now on we won't have to worry about existence of the limit of averages with integer parts of polynomial iterates.

2.1. **Results in general systems.** To study the existence of a limit and its precise value-expression are two different problems.

In case we knew for some specific polynomial families the precise expression of the limit of (9), we would be able to get deeper applications for the corresponding systems. In the special case where we could obtain a result like this for general systems under no assumptions, ergodicity etc, we would have a combinatorial result.

The first result ever in this direction for multiple averages with polynomial iterates of the form $[p(n)]$, $2[p(n)], \dots, k[p(n)]$ for some special polynomial $p(t) \in \mathbb{R}[t]$ (see below) is due to Frantzikinakis:

Theorem 2.2 (Frantzikinakis, [7]). *Let $p \in \mathbb{R}[t]$ with $p(t) \neq cq(t) + d$, for all $c, d \in \mathbb{R}$ and $q \in \mathbb{Q}[t]$. Then, for every $k \in \mathbb{N}$, system $(X, \mathcal{B}, \mu, T)$ and $f_1, \dots, f_k \in L^\infty(\mu)$, we have*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} T^{[p(n)]}f_1 \cdot T^{2[p(n)]}f_2 \cdot \ldots \cdot T^{k[p(n)]}f_k = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} T^{n}f_1 \cdot T^{2n}f_2 \cdot \ldots \cdot T^{kn}f_k, \tag{10}$$

where the convergence takes place in L^2.

This result of Frantzikinakis is an intermediate result in order for him to study multiple averages with iterates coming from Hardy field functions (see [7]). This is the first result on polynomial non-linear iterates where we know the precise expression of the limit (via the work of Host-Kra, [13]).

Now, one can see that one out of the plethora of applications that Theorem 2.2 has, is a Szemerédi-type result implication. Indeed, using Theorems 1.2 and 1.3, Theorem 2.2 implies that any set $E \subseteq \mathbb{N}$ with $\bar{d}(E) > 0$ contains arithmetic progressions of the form $\{[p(n)], 2[p(n)], \dots, k[p(n)]\}$ for any $p \in \mathbb{R}[t]$ with $p(t) \neq cq(t) + d$, for all $c, d \in \mathbb{R}$ and $q \in \mathbb{Q}[t]$, any $k \in \mathbb{N}$ and some n depending on k (for more applications check [7]).

Switching gears to multiple polynomials, one can generalize the condition that Frantzikinakis has for a single polynomial to the following:

Definition. For $k \in \mathbb{N}$, let $\{p_1, \dots, p_k\}$ be a family of real polynomials. We say that this family is *strongly independent* (or that the polynomials $p_1, \dots, p_k$ are *strongly independent*) if any non-trivial real linear combination of the polynomials p_i has a non-constant irrational coefficient.

Note that a family with one element, $\{p\}$, where $p \in \mathbb{R}[t]$, is strongly independent iff $p(t) \neq cq(t) + d$ for all $c, d \in \mathbb{R}$ and $q \in \mathbb{Q}[t]$ (or $\mathbb{Z}[t]$ equivalently).

Examples. The family of polynomials $\{\sqrt{2}t^3+t^2, \sqrt{3}t^3-t\}$ is strongly independent while the families $\{\sqrt{5}t^3+t^2+\sqrt{6}t, t^2, \sqrt{7}t\}$ and $\{\sqrt{2}t^2+t, \sqrt{5}t^2-t\}$ are not.

At this point we also remark that a very nice family of polynomials, in the sense that we have many results for it, with integer coefficients of different degrees is trivially not strongly independent. This fact is natural though as it is known that polynomials of different degrees, while having nice properties and we know many converging results for averages with iterates such polynomials, don't behave in the expected way for general systems (see below for a more detailed clarification of this statement). Hence, someone, in order to get a general, for all systems result, has to restrict to a more special families of polynomials, as the strongly independent ones.

The main result of these notes is the following:

Theorem 2.3 (Karageorgos - K, [14]). *Let $k\in\mathbb{N}$, $p_1,\dots,p_k\in\mathbb{R}[t]$ be strongly independent real polynomials, $(X,\mathcal{B},\mu,T)$ be an ergodic system and $f_1,\dots,f_k\in L^\infty(\mu)$. Then*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}T^{[p_1(n)]}f_1\cdot\ldots\cdot T^{[p_k(n)]}f_k=\prod_{i=1}^{k}\int f_i\,d\mu, \tag{11}$$

where the convergence takes place in $L^2(\mu)$.

So, not only we know the precise expression of the limit but it is also the expected one.

We note at this point that only for aesthetic reasons we stated Theorem 2.3 under the ergodicity assumption for if we deal with a general T, using the ergodic decomposition of μ (i.e., splitting the space into fibers where in each one the system is ergodic), we get that the limit of (11) is equal to the product of the conditional expectations of f_i's with respect to the σ-algebra of the T-invariant sets (which we denote with $\mathbb{E}(f_i|\mathcal{I}(T))$).

Remark. The assumption of Theorem 2.3 that the polynomials are strongly independent is necessary, since even for $k=1$, $p(t)=\sqrt{2}t$ and ergodic rotations on the torus, (11) typically fails.

Hence, even for families of polynomials with integer coefficients it is not true in general that one has convergence as in (11), i.e., to the expected limit for a general ergodic system (see remark after Theorem 2.4). Such a result requires more assumptions on the system, as the total ergodicity one (see [8]). Hence, one is "forced" to work with real polynomials in order to have this nice convergence behavior.

As a consequence of Theorem 2.3, via Hölder's inequality, we get the following recurrence result:

Theorem 2.4 (Karageorgos - K, [14]). *Let $k\in\mathbb{N}$ and $p_1,\dots,p_k\in\mathbb{R}[t]$ be strongly independent real polynomials. Then for every system $(X,\mathcal{B},\mu,T)$ and $A\in\mathcal{B}$ we have*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu\Big(A\cap T^{-[p_1(n)]}A\cap\ldots\cap T^{-[p_k(n)]}A\Big)\geq(\mu(A))^{k+1}. \tag{12}$$

Remark. The assumption that the polynomials are strongly independent is necessary since even for $k=1$ and $p(t)=t^2$, (12) typically fails.

Hence, Theorem 2.4 is another indication that one has to work with real polynomials in order to have nice lower bounds as in (12) for general systems.

Note at this point that following the arguments of the proof of Theorem 2.3 we can show its uniform version, meaning that one can replace the standard Cesàro averages, $\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}$, with the respective uniform ones, $\lim_{N-M\to\infty}\frac{1}{N-M}\sum_{n=M+1}^{N}$, and the natural upper density, $\bar{d}$, with the respective upper Banach density, d^* [1].

Then, one has that the uniform version of Theorem 2.4 implies that for any $A\in\mathcal{B}$ with $\mu(A)>0$, and every $\varepsilon>0$ the set

$$R_\varepsilon(A)=\Big\{n\in\mathbb{Z}:\ \mu\Big(A\cap T^{-[p_1(n)]}A\cap\ldots\cap T^{-[p_k(n)]}A\Big)>(\mu(A))^{k+1}-\varepsilon\Big\}$$

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

We note that this general result, which holds under no assumption on the system, implies that a family of strongly independent real polynomials has a much different behavior than a family of linear integer polynomials, since its stands in contrast with the Bergelson-Host-Kra-Ruzsa counterexample to the "higher-order Khintchine recurrence theorem". Indeed, in [3], the aforementioned authors found an ergodic system $(X,\mathcal{B},\mu,T)$ and a set $A\in\mathcal{B}$ with $\mu(A)>0$ such that

$$\mu\Big(A\cap T^{-n}A\cap T^{-2n}A\cap T^{-3n}A\cap T^{-4n}A\Big)\leq\frac{\mu(A)^5}{2}\quad\text{for all}\quad n\neq 0$$

(so, for $p_i(t)=it$ we have that the syndeticity conclusion of the respective $R_\varepsilon(A)$ fails for certain ergodic systems when $k\geq 4$, while for certain non-ergodic systems it fails even when $k\geq 2$. For examples covering both cases, see [3]).

Using Theorem 2.4 and Furstenberg's corresponding principle, we have the following.

Theorem 2.5 (Karageorgos - K, [14]). *Let $k\in\mathbb{N}$ and $p_1,\ldots,p_k\in\mathbb{R}[t]$ be strongly independent real polynomials. Then for every $E\subseteq\mathbb{N}$ we have*

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\bar{d}(E\cap(E-[p_1(n)])\cap\ldots\cap(E-[p_k(n)]))\geq(\bar{d}(E))^{k+1}.$$

An immediate implication of the aforementioned result is the following.

Theorem 2.6 (Karageorgos - K, [14]). *Let $k\in\mathbb{N}$ and $p_1,\ldots,p_k\in\mathbb{R}[t]$ be strongly independent real polynomials. Then every $E\subseteq\mathbb{N}$ with $\bar{d}(E)>0$ contains arithmetic configurations of the form*

$$\{m,m+[p_1(n)],m+[p_2(n)],\ldots,m+[p_k(n)]\}$$

for some $m\in\mathbb{Z}$ and $n\in\mathbb{N}$ with $[p_i(n)]\neq 0$, for all $1\leq i\leq k$.

We note that one can get the aforementioned result for integer polynomials with no constant term from the polynomial Szemerédi theorem (Theorem A_0, [4]), but in the generality that we present it here it is not clear to us at all if the theorem follows from previous results in the literature.

---

[1] For a set $A\subseteq\mathbb{Z}$, we define its *upper Banach density*, $d^*(A)$, as $d^*(A)=\limsup_{N-M\to\infty}\frac{|A\cap\{M+1,\ldots,N\}|}{N-M}$.

3. Background material

3.1. **Nilmanifolds.** In this subsection we recall some basic facts concerning nilmanifolds and equidistribution results on them.

3.1.1. *Definitions and basic properties.* Let G be a k-step nilpotent Lie group, meaning $G_{k+1} = \{e\}$ for some $k \in \mathbb{N}$, where $G_k = [G, G_{k-1}]$ denotes the k-th commutator subgroup, and let Γ be a discrete cocompact subgroup of G. The compact homogeneous space $X = G/\Gamma$ is called a *k-step nilmanifold* (or just *nilmanifold*).

The group G acts on G/Γ by left translations where the translation by an element $b \in G$ is given by $T_b(g\Gamma) = (bg)\Gamma$. We denote by m_X the normalized *Haar measure* on X, meaning the unique probability measure that is invariant under the action of G by left translations and $\mathcal{G}/\Gamma$ denotes the Borel σ-algebra of G/Γ. If $b \in G$, we call the system $(G/\Gamma, \mathcal{G}/\Gamma, m_X, T_b)$ a *k-step nilsystem* (or just *nilsystem*) and the elements of G *nilrotations.*

3.1.2. *Equidistribution on nilmanifolds.* Let $\exp : \mathfrak{g} \to G$ be the exponential map, where $\mathfrak{g}$ is the Lie algebra of G for a connected and simply connected Lie group G. For $b \in G$ and $s \in \mathbb{R}$ we define the element b^s of G as follows: If $Z \in \mathfrak{g}$ is such that $\exp(Z) = b$, then $b^s = \exp(sZ)$ (this is well defined since exp is a bijection).

If $(a(n))_n$ is a sequence of real numbers and $X = G/\Gamma$ is a nilmanifold with G connected and simply connected, we say that the sequence $(b^{a(n)}x)_n$ is *equidistributed* in a sub-nilmanifold Y of X, if for every $F \in C(Y)$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} F(b^{a(n)}x) = \int F \, dm_Y.$$

If the sequence $(a(n))_n$ takes only integer values, we are not obliged to assume that G is connected and simply connected.

A nilrotation $b \in G$ is *ergodic*, or *acts ergodically* on X, if the sequence $(b^n\Gamma)_n$ is dense in X. If $b \in G$ is ergodic, then for every $x \in X$ the sequence $(b^n x)_n$ is equidistributed in X (a nontrivial fact which follows by unique ergodicity).

Let $X = G/\Gamma$ be a nilmanifold and $b \in G$. Then the orbit closure $\overline{(b^n\Gamma)_n}$ of b has the structure of a nilmanifold. Furthermore, the sequence $(b^n\Gamma)_n$ is equidistributed in $\overline{(b^n\Gamma)_n}$. If G is connected and simply connected and $b \in G$, then $\overline{(b^s\Gamma)_{s\in\mathbb{R}}}$ is a nilmanifold. Furthermore, the nilflow $(b^s\Gamma)_{s\in\mathbb{R}}$ is equidistributed in $\overline{(b^s\Gamma)_{s\in\mathbb{R}}}$.

If G is a nilpotent group, then a sequence $g : \mathbb{N} \to G$ of the form $g(n) = b_1^{p_1(n)} \cdots b_k^{p_k(n)}$, where $b_i \in G$ and p_i are polynomials taking integer values at the integers for every $1 \leq i \leq k$ is called a *polynomial sequence* in G. A *polynomial sequence on the nilmanifold* $X = G/\Gamma$ is a sequence of the form $(g(n)\Gamma)_n$ where $g : \mathbb{N} \to G$ is a polynomial sequence in G.

The following qualitative equidistribution result was established by Leibman in [17]:

Theorem 3.1 (Theorems B, C, [17])**.** *Suppose that $X = G/\Gamma$ is a nilmanifold with G connected and simply connected and $(g(n))_n$ is a polynomial sequence in G. Let $Z = G/([G,G]\Gamma)$ and $\pi : X \to Z$ be the natural projection. Then the following statements hold:*

(i) *For every $x \in X$ the sequence $(g(n)x)_n$ is equidistributed in a finite union of subnilmanifolds of X.*

(ii) *For every $x \in X$ the sequence $(g(n)x)_n$ is equidistributed in X if and only if the sequence $(g(n)\pi(x))_n$ is equidistributed in Z.*

If $X = G/\Gamma$ is a nilmanifold with G connected and simply connected, then Z is a connected compact abelian Lie group, hence a torus, meaning $\mathbb{T}^s$ for some $s \in \mathbb{N}$, and as a consequence every nilrotation in Z is isomorphic to a rotation on $\mathbb{T}^s$.

3.2. **Ergodic Theory.** We gather below some basic notions and facts from ergodic theory that we use throughout the paper.

3.2.1. *Factors.* A *homomorphism* from a system $(X, \mathcal{X}, \mu, T)$ onto a system $(Y, \mathcal{Y}, \nu, S)$ is a measurable map $\pi : X \to Y$, such that $\mu \circ \pi^{-1} = \nu$ and $S \circ \pi(x) = \pi \circ T(x)$ for $x \in X$. When we have such a homomorphism we say that the system $(Y, \mathcal{Y}, \nu, S)$ is a *factor* of the system $(X, \mathcal{X}, \mu, T)$. If the factor map $\pi : X \to Y$ can be chosen to be injective, then we say that the systems $(X, \mathcal{X}, \mu, T)$ and $(Y, \mathcal{Y}, \nu, S)$ are *isomorphic.* A factor can also be characterised by $\pi^{-1}(\mathcal{Y})$ which is a T-invariant sub-σ-algebra of $\mathcal{X}$. By a classical abuse of terminology we denote by the same letter the σ-algebras $\mathcal{Y}$ and $\pi^{-1}(\mathcal{Y})$.

3.2.2. *Characteristic Factors.* Let $(X, \mathcal{X}, \mu, T)$ be a system. We say that the σ-algebra $\mathcal{Y}$ of $\mathcal{X}$ is a *characteristic factor* for the family of integer sequences $\{(a_1(n))_n, \dots, (a_k(n))_n\}$ if $\mathcal{Y}$ is T-invariant and

$$\lim_{N\to\infty} \left\| \frac{1}{N}\sum_{n=1}^{N} T^{a_1(n)} f_1 \cdot \ldots \cdot T^{a_k(n)} f_k - \frac{1}{N}\sum_{n=1}^{N} T^{a_1(n)} \tilde{f}_1 \cdot \ldots \cdot T^{a_k(n)} \tilde{f}_k \right\|_{L^2(\mu)} = 0,$$

where $\tilde{f}_i = \mathbb{E}(f_i|\mathcal{Y})$, for $f_i \in L^\infty(\mu)$ for all $1 \le i \le k$ [2].

3.2.3. *Seminorms and Nilfactors.* We follow [13] and [5] for the inductive definition of the seminorms $\|\!|\cdot|\!\|_k$. More specifically, the definition that we use here follows from [13] (in the ergodic case), [5] (in the general case) and the use of von Neumann's ergodic theorem.

Let $(X, \mathcal{B}, \mu, T)$ be a system and $f \in L^\infty(\mu)$. We define inductively the seminorms $\|\!|f|\!\|_k$ as follows: For $k = 1$ we set

$$\|\!|f|\!\|_1 := \|\mathbb{E}(f|\mathcal{I}(T))\|_{L^2(\mu)}\,.$$

For $k \ge 1$, we let

$$\|\!|f|\!\|_{k+1}^{2^{k+1}} := \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \|\!|\bar{f} \cdot T^n f|\!\|_k^{2^k}\,.$$

It was shown in [13] that for every integer $k \ge 1$ all these limits exist and $\|\!|\cdot|\!\|_k$ defines a seminorm on $L^\infty(\mu)$.

Using these seminorms we can construct factors $\mathcal{Z}_k = \mathcal{Z}_k(T)$ of X characterized by the property:

$$\text{for } f \in L^\infty(\mu), \quad \mathbb{E}(f|\mathcal{Z}_{k-1}) = 0 \text{ if and only if } \|\!|f|\!\|_k = 0.$$

It was also shown in [13] that for every $k \in \mathbb{N}$ the factor $\mathcal{Z}_k$ has an algebraic structure, in fact we can assume that it is a k-step nilsystem. This is the content of the following Structure theorem, which we recall in the ergodic case and follows by Theorem 10.1 in [13]:

[2]Equivalently, if $\mathbb{E}(f_i|\mathcal{Y}) = 0$ for some $1 \le i \le k$, then $\lim_{N\to\infty} \left\| \frac{1}{N}\sum_{n=1}^{N} T^{a_1(n)} f_1 \cdot \ldots \cdot T^{a_k(n)} f_k \right\|_{L^2(\mu)} = 0.$

Theorem 3.2 (**Host & Kra**, [13]). *Let $(X,\mathcal{B},\mu,T)$ be an ergodic system and $k\in\mathbb{N}$. Then the factor $\mathcal{Z}_k(T)$ is an inverse limit of k-step nilsystems.* [3]

Because of this result we call $\mathcal{Z}_k$ the *k-step nilfactor* of the system. The smallest factor that is an extension of all finite step nilfactors is denoted by $\mathcal{Z}=\mathcal{Z}(T)$, meaning, $\mathcal{Z}=\bigvee_{k\in\mathbb{N}}\mathcal{Z}_k$, and is called the *nilfactor* of the system.

4. Proof of Theorem 2.3

In this section we will state the intermediate results that we used in order to prove Theorem 2.3.

The main argument is an equidistribution result involving nil-orbits of several sequences of strongly independent polynomials (first proved for Hardy field functions by Frantzikinakis, [6, Theorem 1.3]).

Theorem 4.1 (Karageorgos - K, [14]). *Let $k\in\mathbb{N}$ and $p_1,\ldots,p_k\in\mathbb{R}[t]$ be strongly independent real polynomials.*

(i) *If $X_i=G_i/\Gamma_i$, $1\leq i\leq k$, are nilmanifolds with G_i connected and simply connected, then for every $b_i\in G_i$ and $x_i\in X_i$ the sequence*

$$\left(b_1^{p_1(n)}x_1,\ldots,b_k^{p_k(n)}x_k\right)_n$$

is equidistributed in the nilmanifold

$$\overline{(b_1^s x_1)}_{s\in\mathbb{R}}\times\cdots\times\overline{(b_k^s x_k)}_{s\in\mathbb{R}}.$$

(ii) *If $X_i=G_i/\Gamma_i$, $1\leq i\leq k$, are nilmanifolds, then for every $b_i\in G_i$ and $x_i\in X_i$ the sequence*

$$\left(b_1^{[p_1(n)]}x_1,\ldots,b_k^{[p_k(n)]}x_k\right)_n$$

is equidistributed in the nilmanifold

$$\overline{(b_1^n x_1)}_n\times\cdots\times\overline{(b_k^n x_k)}_n.$$

Actually, by [6, Lemma 5.1], Part (ii) of Theorem 4.1 follows from Part (i).
Part (i) of Theorem 4.1 on the other hand, follows by the following two statements:

Proposition 4.2 (Karageorgos - K, [14]). *Let $k\in\mathbb{N}$ and $p_1,\ldots,p_k\in\mathbb{R}[t]$ be strongly independent real polynomials. Let $X=G/\Gamma$ be a nilmanifold with G connected and simply connected and elements $b_i\in G$ acting ergodically on X. Then the sequence*

$$\left(b_1^{p_1(n)}\Gamma,\ldots,b_k^{p_k(n)}\Gamma\right)_n$$

is equidistributed in the nilmanifold X^k.

[3]By this we mean that there exist T-invariant sub-σ-algebras $\mathcal{Z}_{k,i}$, $i\in\mathbb{N}$, of $\mathcal{B}$ such that $\mathcal{Z}_k=\bigcup_{i\in\mathbb{N}}\mathcal{Z}_{k,i}$ and for every $i\in\mathbb{N}$, the factors induced by the σ-algebras $\mathcal{Z}_{k,i}$ are isomorphic to k-step nilsystems.

This proposition uses the deep result of Leibman, Theorem 3.1, on the equidistribution of polynomial sequences in a nilmanifold. From the proof of this proposition, by Weyl's criterion, we get a first condition that our polynomials have to satisfy in order for Theorem 4.1 to hold.

The last ingredient in proving Part (i) of Theorem 4.1 is the following lemma:

Lemma 4.3 (Lemma 5.2, [6]). *Let $k \in \mathbb{N}$ and $X = G/\Gamma$ be a nilmanifold with G connected and simply connected. Then for every $b_1, \ldots, b_k \in G$ there exists an $s_0 \in \mathbb{R}$ such that for all $1 \leq i \leq k$ the element $b_i^{s_0}$ acts ergodically on the nilmanifold $\overline{(b_i^s\Gamma)}_{s\in\mathbb{R}}$.*

The proof of this lemma (together with the proof of the precious proposition) gives us the precise condition, that of "strongly independence", that our polynomials have to satisfy for Theorem 4.1 to hold.

The last step before the proof of Theorem 2.3 is to show that the nilfactor is the characteristic factor for our "nice" polynomial iterates.

Definition ([7]). Let $k \in \mathbb{N}$ and for $N \in \mathbb{N}$, let $\mathcal{P}_N = \{p_{1,N}, \ldots, p_{k,N}\}$ be a family of polynomials with real coefficients. We say that the collection $(P_N)_N$ is *nice* if for every $N \in \mathbb{N}$ the polynomials $p_{i,N}$ and $p_{i,N} - p_{j,N}$, $i \neq j$, are non-constant and their leading coefficients are independent of N.

Note that a strongly independent family of polynomials is nice.

Lemma 4.4 (Lemma 4.7, [7]). *Let $(\{p_{1,N}, \ldots, p_{k,N}\})_N$ be a nice collection of polynomial families, $(X, \mathcal{B}, \mu, T)$ be a system and suppose that one of the functions $f_1, \ldots, f_k \in L^\infty(\mu)$ is orthogonal to the nilfactor $\mathcal{Z}$. Then for any Følner sequence $(\Phi_N)_N$ in $\mathbb{Z}$ [4] and any bounded two parameter sequence $(c_{N,n})_{N,n}$ of real numbers we have*

$$\lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} c_{N,n} T^{[p_{1,N}(n)]} f_1 \cdot \ldots \cdot T^{[p_{k,N}(n)]} f_k = 0, \tag{13}$$

where the convergence takes place in $L^2(\mu)$.

We close this subsection with the proof of Theorem 2.3:

Proof of Theorem 2.3, [14]. We start by using Lemma 4.4 in order to get that the nilfactor $\mathcal{Z}$ is characteristic for the corresponding multiple ergodic average. Via Theorem 3.2 we can assume without loss of generality that our system is an inverse limit of nilsystems. By a standard approximation argument, we can further assume that our system is a nilsystem.

Let $(X = G/\Gamma, \mathcal{G}/\Gamma, m_X, T_b)$ be a nilsystem, where $b \in G$ is ergodic, and $F_1, \ldots, F_k \in L^\infty(m_X)$. Our objective now is to show that if $\{p_1, \ldots, p_k\}$ is a strongly independent family of polynomials then

$$\lim_{N\to\infty} \sum_{n=1}^{N} F_1(b^{[p_1(n)]}x) \cdot \ldots \cdot F_\ell(b^{[p_k(n)]}x) = \int F_1 \, dm_X \cdot \ldots \cdot \int F_k \, dm_X \tag{14}$$

[4] A *Følner sequence in* $\mathbb{Z}$ is a sequence $(\Phi_n)_n$ of finite sunsets of $\mathbb{Z}$ such that for any $m \in \mathbb{Z}$ we have $\lim_{n\to\infty} \frac{|(\Phi_n + m) \cap \Phi_n|}{|\Phi_n|} = 1$.

where the convergence takes place in $L^2(m_X)$. By density, we can assume that the functions $F_1, \ldots, F_k$ are continuous. Then we can apply Theorem 4.1 to the nilmanifold X^k, the nilrotation $\tilde{b} = (b, \ldots, b) \in G^k$, the point $\tilde{x} = (x, \ldots, x) \in X^k$, and the continuous function $\tilde{F}(x_1, \ldots, x_k) = F_1(x_1) \cdot \ldots \cdot F_k(x_k)$, to get that

$$\lim_{N\to\infty} \sum_{n=1}^{N} \tilde{F}(b^{[p_1(n)]}x, \ldots, b^{[p_k(n)]}x) = \int \tilde{F}\; dm_{X^k}$$

and this gives the desired limit in (14), completing the proof. □

5. From averages along natural to prime numbers

In this last section we will prove the corresponding expressions of Theorems 2.2 and 2.3, and so, their applications as well, along prime numbers. More specifically we will show:

Theorem 5.1. *Let $q \in \mathbb{R}[t]$ with $q(t) \neq c\tilde{q}(t) + d$ for all $c, d \in \mathbb{R}$ and $\tilde{q} \in \mathbb{Q}[t]$. Then for every $k \in \mathbb{N}$, system $(X, \mathcal{B}, \mu, T)$ and $f_1, \ldots, f_k \in L^\infty(\mu)$, we have that*

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\in\mathbb{P}\cap[1,N]} T^{[q(p)]}f_1 \cdot T^{2[q(p)]}f_2 \cdot \ldots \cdot T^{k[q(p)]}f_k = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^n f_1 \cdot T^{2n} f_2 \cdot \ldots \cdot T^{kn} f_k,$$

where the convergence takes place in $L^2(\mu)$ and $\pi(N) = |\mathbb{P} \cap [1, N]|$ denotes the number of primes up to N.

Theorem 5.2. *Let $k \in \mathbb{N}$, $p_1, \ldots, p_k \in \mathbb{R}[t]$ be strongly independent real polynomials, $(X, \mathcal{B}, \mu, T)$ be an ergodic system and $f_1, \ldots, f_k \in L^\infty(\mu)$. Then*

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\in\mathbb{P}\cap[1,N]} T^{[p_1(p)]}f_1 \cdot \ldots \cdot T^{[p_k(p)]}f_k = \prod_{i=1}^{k} \int f_i\; d\mu,$$

where the convergence takes place in $L^2(\mu)$.

We start by recalling the definition of the *von Mangoldt function*, $\Lambda : \mathbb{N} \to \mathbb{R}$, where $\Lambda(n) = \begin{cases} \log(p) & , \text{ if } n = p^k \text{ for some } p \in \mathbb{P} \text{ and some } k \in \mathbb{N} \\ 0 & , \text{ elsewhere} \end{cases}$.

It is more natural though for us to work instead of Λ with the function $\Lambda' : \mathbb{N} \to \mathbb{R}$, where $\Lambda'(n) = \mathbf{1}_{\mathbb{P}}(n) \cdot \Lambda(n) = \mathbf{1}_{\mathbb{P}}(n) \cdot \log(n)$.

The function Λ', according to the following lemma, will allow us to relate averages along primes with weighted averages over the integers.

Lemma 5.3 ([9])**.** *If $a : \mathbb{N} \to \mathbb{C}$ is bounded, then*

$$\lim_{N\to\infty} \left| \frac{1}{\pi(N)} \sum_{p\in\mathbb{P}\cap[1,N]} a(p) - \frac{1}{N} \sum_{n=1}^{N} \Lambda'(n) \cdot a(n) \right| = 0.$$

The proof of this lemma, which can be found in [9], uses the prime number theorem and it is relatively immediate.

Note that in order for someone to show convergence along primes of averages of the sequence $(a(n))_n$, according to this lemma, has to show convergence along natural numbers of averages of the sequence $(\Lambda'(n) \cdot a(n))_n$.

For $w > 2$, let

$$W = \prod_{p \in \mathbb{P} \cap [1, w-1]} p$$

be the product of primes bounded above by w. For $r \in \mathbb{N}$, let

$$\Lambda'_{w,r}(n) = \frac{\phi(W)}{W} \cdot \Lambda'(Wn + r),$$

where ϕ is the Euler function, be the *modified von Mangoldt function.*

The proposition below, the proof of which relies on a deep result due to Green and Tao ([12]) on the inverse conjecture for the Gowers norms, will provide us with a crucial intermediate step in order to prove Theorems 5.1 and 5.2, as well as as well as to get their implications (we will actually use a very weak version of it for all these results).

Proposition 5.4 (Proposition 3.2, [16]). *Let $k, m \in \mathbb{N}$, $(X, \mathcal{B}, \mu, T_1, \dots, T_m)$ be a system, where T_i's commute, $p_{i,j} \in \mathbb{R}[t]$ be real polynomials, $1 \leq i \leq m$, $1 \leq j \leq k$ and $f_1, \dots, f_k \in L^\infty(\mu)$.*

Then,

$$\max_{1 \leq r \leq W,\, (r,W)=1} \left\| \frac{1}{N} \sum_{n=1}^N (\Lambda'_{w,r}(n) - 1) \cdot (\prod_{i=1}^m T_i^{[p_{i,1}(Wn+r)]}) f_1 \cdot \ldots \cdot (\prod_{i=1}^m T_i^{[p_{i,k}(Wn+r)]}) f_k \right\|_{L^2(\mu)}$$

converges to 0 as $N \to \infty$ and then $w \to \infty$.

Proof of Theorem 5.1. We borrow the arguments from the proof of Theorem 1.3 from [10] (see also Theorem 1.3 in [16]). By Lemma 5.3 it suffices to show that the sequence

$$A(N) := \frac{1}{N} \sum_{n=1}^N \Lambda'(n) \cdot T^{[q(n)]} f_1 \cdot T^{2[q(n)]} f_2 \cdot \ldots \cdot T^{k[q(n)]} f_k$$

converges in $L^2(\mu)$ to the same limit as the sequence $\frac{1}{N} \sum_{n=1}^N T^n f_1 \cdot T^{2n} f_2 \cdot \ldots \cdot T^{kn} f_k$ as $N \to \infty$. For w (which gives a corresponding W), $r \in \mathbb{N}$, we define

$$B_{w,r}(N) := \frac{1}{N} \sum_{n=1}^N T^{[q(Wn+r)]} f_1 \cdot T^{2[q(Wn+r)]} f_2 \cdot \ldots \cdot T^{k[q(Wn+r)]} f_k.$$

For any $\varepsilon > 0$, using Proposition 5.4 with $m = k$, $T_i = T$, $1 \leq i \leq k$ and $p_{i,j} = \begin{cases} 0 & \text{, if } i \leq k - j \\ q & \text{, elsewhere} \end{cases}$, for sufficiently large N and some w_0 we have

$$\left\| A(W_0 N) - \frac{1}{\phi(W_0)} \sum_{1 \leq r \leq W_0,\, (r, W_0)=1} B_{w_0,r}(N) \right\|_{L^2(\mu)} < \varepsilon.$$

Note at this point that for all $W, r \in \mathbb{N}$ we have that $q(Wt + r) \notin c\mathbb{Q}[t] + d$ for $c, d \in \mathbb{R}$, for otherwise q would have the same property contradicting our assumption.

By Theorem 2.2, we have that for any $1 \leq r \leq W_0$ the sequence $(B_{w_0,r}(N))_N$ converges to the same limit as the sequence $\frac{1}{N}\sum_{n=1}^{N} T^n f_1 \cdot T^{2n} f_2 \cdot \ldots \cdot T^{kn} f_k$, and since

$$\lim_{N\to\infty} \|A(W_0 N + r) - A(W_0 N)\|_{L^2(\mu)} = 0$$

for every $1 \leq r \leq W_0$, we get the result. □

Proof of Theorem 5.2. The proof is analogous to the previous one. In this case we define $A(N) := \frac{1}{N}\sum_{n=1}^{N} \Lambda'(n) \cdot T^{[p_1(n)]} f_1 \cdot \ldots \cdot T^{[p_k(n)]} f_k$ and for $w, r \in \mathbb{N}$, $B_{w,r}(N) := \frac{1}{N}\sum_{n=1}^{N} T^{[p_1(Wn+r)]} f_1 \cdot \ldots \cdot T^{[p_k(Wn+r)]} f_k$. We use Proposition 5.4 with $m = k$, $T_i = T$, $1 \leq i \leq k$, $p_{i,j} = \begin{cases} 0 & , \text{ if } i \neq j \\ p_i & , \text{ if } i = j \end{cases}$ and we note that the family $\{\tilde{p}_1, \ldots, \tilde{p}_k\}$, where $\tilde{p}_i(t) = p_i(Wt + r)$, is strongly independent for all $W, r \in \mathbb{N}$. (Indeed, if for some $(\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k \setminus \{\vec{0}\}$, $d \in \mathbb{R}$, $q \in \mathbb{Q}[t]$ and $W, r \in \mathbb{N}$ we had $\sum_{i=1}^{k} \lambda_i p_i(Wt + r) = q(t) + d$, then $\sum_{i=1}^{k} \lambda_i p_i(t) = \tilde{q}(t) + d$, where $\tilde{q}(t) = q((t - r)/W) \in \mathbb{Q}[t]$, a contradiction to the strong independence assumption.) The result now follows similarly to the previous proof since by Theorem 2.3, we have that for any $1 \leq r \leq W_0$ the sequence $(B_{w_0,r}(N))_N$ converges, in $L^2(\mu)$, to $\prod_{i=1}^{k} \int f_i \, d\mu$. □

References

[1] V. Bergelson. Ergodic Ramsey theory. *Logic and combinatorics (Arcata, Calif., 1985)*, 63–87, *Contemp. Math.*, **65**, *Amer. Math. Soc., Providence, RI*, 1987.

[2] V. Bergelson. Weakly mixing PET. *Ergodic Theory Dynam. Systems* **7** (1987), no. 3, 337–349.

[3] V. Bergelson, B. Host, B. Kra, with an appendix by I. Ruzsa. Multiple recurrence and nilsequences. *Inventiones Math.* **160** (2005), no. 2, 261–303.

[4] V. Bergelson, A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *Journal of AMS* **9** (1996), no. 3, 725–753.

[5] Q. Chu, N. Frantzikinakis, B. Host. Ergodic averages of commuting transformations with distinct degree polynomial iterates. *Proc. of the London Math. Society.* (3), **102** (2011), 801–842.

[6] N. Frantzikinakis. Equidistribution of sparse sequences on nilmanifolds. *Journal d'Analyse Mathematique*, **109** (2009), 353–395.

[7] N. Frantzikinakis. Multiple recurrence and convergence for Hardy field sequences of polynomial growth. *Journal d'Analyse Mathematique* **112** (2010), 79–135.

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

[9] N. Frantzikinakis, B. Host, B. Kra. Multiple recurrence and convergence for sequences related to the prime numbers. *J. Reine Angew. Math.* **611** (2007), 131–144.

[10] N. Frantzikinakis, B. Host, B. Kra. The polynomial multidimensional Szemerédi theorem along shifted primes. *Israel J. Math.* **194** (2013), no. 1, 331–348.

[11] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Analyse Math.* **31** (1977), 204-–256.
[12] B. Green and T. Tao. Linear equations in primes. *Ann. of Math.* (2) **171** (2010), 1753–1850.
[13] B. Host, B. Kra. Nonconventional ergodic averages and nilmanifolds. *Annals of Math.* **161** (2005), no. 1, 397–488.
[14] D. Karageorgos, A. Koutsogiannis. Integer part independent polynomial averages and applications along primes. To appear in Studia Mathematica.
[15] A. Koutsogiannis. Integer part polynomial correlation sequences. *Ergodic Theory and Dynamical Systems* **38** (2018), no. 4, 1525–1542.
[16] A. Koutsogiannis. Closest integer polynomial multiple recurrence along shifted primes. *Ergodic Theory Dynam. Systems* **38** (2018), no. 2, 666–685.
[17] A. Leibman. Pointwise Convergence of ergodic averages for polynomial sequences of translations on a nilmanifold. *Ergodic Theory Dynam. Systems* **25** (2005), no. 1, 201–213.
[18] M. Walsh. Norm convergence of nilpotent ergodic averages. *Annals of Mathematics* **175** (2012), no. 3, 1667–1688.

(Andreas Koutsogiannis) The Ohio State University, Department of mathematics, Columbus, Ohio, USA
Email address: koutsogiannis.1@osu.edu