SOME OPEN PROBLEMS ON MULTIPLE ERGODIC AVERAGES

NIKOS FRANTZIKINAKIS

1. Some useful tools and observations

1.1. **Characteristic factors.** A notion that underlies the study of the limiting behavior of several multiple ergodic averages is that of the *characteristic factor(s).* Implicit use of this notion was already made on the foundational article of H. Furstenberg [35], but the term "characteristic factor" was coined in a paper of H. Furstenberg and B. Weiss [37].

Given a probability space $(X, \mathcal{X}, \mu)$ and a collection of measure preserving transformations $T_1, \ldots, T_\ell \colon X \to X$, we say that the sub-$\sigma$-algebras $\mathcal{X}_1, \ldots, \mathcal{X}_\ell$ of $\mathcal{X}$ are *characteristic factors* for the averages

$$A_N(f_1, \ldots, f_\ell) := \frac{1}{N} \sum_{n=1}^{N} T_1^{a_1(n)} f_1 \cdot \ldots \cdot T_\ell^{a_\ell(n)} f_\ell \tag{1}$$

if the following two conditions hold:

- $\mathcal{X}_i$ is T_i-invariant for $i = 1, \ldots, \ell$,
- whenever $f_1, \ldots, f_\ell \in L^\infty(\mu)$, we have $A_N(f_1, \ldots, f_\ell) - A_N(\tilde{f}_1, \ldots, \tilde{f}_\ell) \to^{L^2(\mu)} 0$, where $\tilde{f}_i := \mathbb{E}(f_i|\mathcal{X}_i)$ for $i = 1, \ldots, \ell$.[1]

If in addition one has $\mathcal{X}_1 = \cdots = \mathcal{X}_\ell$, then we call this common sub-σ-algebra a *characteristic factor* for the averages (1).

1.2. **Gowers-Host-Kra seminorms.** When analyzing the limiting behavior of the averages (1), an intermediate goal is to choose characteristic factors that are as simple as possible, and typically simple for us means that the corresponding factor systems have very special algebraic structure. Very often this step is carried out by controlling the $L^2(\mu)$-norm of the averages (1) by the *Gowers-Host-Kra seminorms.* Similar seminorms were first introduced in combinatorics by T. Gowers [38] and their ergodic variant (that is more relevant for our study) was introduced by B. Host and B. Kra [44]. For an ergodic system $(X, \mathcal{X}, \mu, T)$ and function $f \in L^\infty(\mu)$, they are defined as follows:

$$|||f|||_1 := \Big| \int f \ d\mu \Big| \ ;$$

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

It is shown in [44] that for every $k \in \mathbb{N}$ the limit above exists, and $||| \cdot |||_k$, thus defined, is a seminorm on $L^\infty(\mu)$. For non-ergodic systems the seminorms can be similarly defined, the

Date: March 2011.

[1]Equivalently, if $\mathbb{E}(f_i|\mathcal{X}_i) = 0$ for some $i \in \{1, \ldots, \ell\}$, then $A_N(f_1, \ldots, f_\ell) \to^{L^2(\mu)} 0$.

only difference is that $\|\!| \cdot |\!\|_1$ is defined by $\|\!| f |\!\|_1 := \left\| \int f \, d\mu_x \right\|_{L^2(\mu)}$, where $\mu = \int \mu_x \, d\mu(x)$ is the ergodic decomposition of the measure μ with respect to T. If further clarification is needed, we write $\|\!| \cdot |\!\|_{k,\mu}$, or $\|\!| \cdot |\!\|_{k,T}$. We remark that if a measure preserving system is *weak mixing*[2], then $\|\!| f |\!\|_k = |\int f \, d\mu|$ for every $k \in \mathbb{N}$.

1.3. **The factors $\mathcal{Z}_k$ and their structure.** The seminorms $\|\!| \cdot |\!\|_k$ induce T-invariant sub-σ-algebras $\mathcal{Z}_{k-1}$ that satisfy

$$\text{for } \ f \in L^\infty(\mu), \quad \mathbb{E}(f|\mathcal{Z}_{k-1}) = 0 \quad \text{if and only if} \quad \|\!| f |\!\|_k = 0. \tag{2}$$

As a consequence, if for some $k_1, \dots, k_\ell \in \mathbb{N}$ one is able to produce an estimate of the form

$$\limsup_{N \to \infty} \|A_N(f_1, \dots, f_\ell)\|_{L^2(\mu)} \ll \min_{i=1,\dots,\ell} \|\!| f_i |\!\|_{k_i, T_i}, \tag{3}$$ [3]

then one knows that the factors $\mathcal{Z}_{k_1-1,T_1}, \dots, \mathcal{Z}_{k_\ell-1,T_\ell}$ are characteristic for mean convergence of the averages (1). Under such circumstances, one gets characteristic factors with the sought-after algebraic structure. This is a consequence of a result of B. Host and B. Kra [44] stating that for ergodic systems the factor system $(X, \mathcal{Z}_k, \mu, T)$ is an inverse limit of k-step nilsystems[4]. Depending on the problem, it may be more useful to think of the previous structure theorem as a decomposition result; for every ergodic system $(X, \mathcal{X}, \mu, T)$ and $f \in L^\infty(\mu)$, for every $k \in \mathbb{N}$ and $\varepsilon > 0$, there exist measurable functions f_s, f_u, f_e, with $L^\infty(\mu)$ norm at most $2\|f\|_{L^\infty(\mu)}$, such that

- $f = f_s + f_u + f_e$;
- $\|\!| f_u |\!\|_{k+1} = 0$; $\|f_e\|_{L^1(\mu)} \le \varepsilon$; and
- $(f_s(T^n x))_{n \in \mathbb{N}}$ is a k-step nilsequence[5] for μ-almost every $x \in X$.

Such a decomposition also holds for non-ergodic systems (see Proposition 3.1 in [21]).

Combining the hypothetical seminorm estimates (3) with the aforementioned structure theorem (or the decomposition result), the problem of analyzing the limiting behavior of the averages (1) is reduced to a new problem that amounts to proving certain equidistribution properties of sequences on nilmanifolds. Tools for handling such equidistribution problems have been developed in recent years, thus making such a reduction very much worthwhile. Some examples of equidistribution results of this type can be found in [2, 25, 39, 40, 49, 50, 52, 53].

1.4. **A general strategy.** Summarizing, when one is against a multiple recurrence problem in ergodic theory, or more generally any problem that can be solved by analyzing the limiting

---

[2]A measure preserving system (X, μ, T) is weak mixing if the product system $(X \times X, \mu \times \mu, T \times T)$ is ergodic, or equivalently, if for every $f \in L^\infty(\mu)$ with $\int f \, d\mu = 0$ one has $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^N |\int \bar{f} \cdot T^n f \, d\mu|^2 = 0$.

[3]As a general principle, if one can show that $A_N(f_1, \dots, f_\ell) \to^{L^2(\mu)} 0$ when all the transformations are weak mixing and one of the functions has mean 0, then one can adapt its proof to show (3).

[4]A k-step nilmanifold is a homogeneous space $X = G/\Gamma$ where G is a k-step nilpotent Lie group, and Γ is a discrete cocompact subgroup of G. A *k-step nilsystem* is a system of the form $(X, \mathcal{G}/\Gamma, m_X, T_a)$ where $X = G/\Gamma$ is a k-step nilmanifold, $a \in G$, $T_a \colon X \to X$ is defined by $T_a(g\Gamma) := (ag)\Gamma$, m_X is the normalized Haar measure on X, and $\mathcal{G}/\Gamma$ is the completion of the Borel σ-algebra of G/Γ.

[5]A *k-step nilsequence* is a uniform limit of sequences of the form $(F(b^n x))$ where $X = G/\Gamma$ is a k-step nilmanifold, $b \in G$, $x \in X$, and F is Riemann integrable on X.

behavior of the multiple ergodic averages (1), very often a useful approach is to try to work out the following three steps:[6]

- Produce seminorm estimates like the ones in (3).
- Use a structure theorem or a decomposition result to reduce matters to nilsystems.
- Use qualitative or quantitative equidistribution results on nilmanifolds to end the proof.

The reader can find several examples demonstrating this approach, or variants of it, to prove multiple recurrence and convergence results, as well as related applications in combinatorics, in the following articles: [1, 10, 13, 14, 17, 20, 21, 22, 23, 24, 26, 28, 29, 30, 31, 33, 34, 37, 41, 43, 44, 45, 46, 47, 48, 51, 54, 55, 56, 60, 61, 62]. Depending on the problem, the difficulty of each step varies; typically the first step is elementary and is carried out by successive uses of the Cauchy-Schwarz inequality and an estimate of van der Corput[7] (or Hilbert space variants of it), the second step involves the use of (an often minor) modification of the structure theorem of B. Host and B. Kra, and the third step is a combination of algebraic and analytic techniques.

1.5. **The polynomial exhaustion technique.** We explain a technique that is often used to produce seminorm estimates of the type (3). It is based on an induction scheme (often called PET induction) introduced by V. Bergelson in [**?**]. Let $F := \{a_1, \dots, a_\ell\}$ be a family of real valued sequences, and suppose that one wishes to establish seminorm estimates of the form

$$\limsup_{N\to\infty} \|A_N(f_1,\dots,f_\ell)\|_{L^2(\mu)} \ll \min_{i=1,\dots,\ell} |\!|\!| f_i |\!|\!|_{k_i} \tag{4}$$

where

$$A_N(f_1,\dots,f_\ell) := \frac{1}{N}\sum_{n=1}^{N} T^{[a_1(n)]}f_1 \cdot \ldots \cdot T^{[a_\ell(n)]}f_\ell.$$

Variations of this method could also be used to get similar estimates for some multiple ergodic averages involving commuting transformations.

1.5.1. *The method.* The main idea is to use some variation of van der Corput's fundamental estimate and bound the left hand side in (4) by an expression that involves families of sequences of smaller "complexity". Our goal is after a finite number of iterations to get families of sequences that are simple enough to handle directly. The details depend on the family of sequences at hand, but typically, after the first iteration, we get an upper bound by an average over $r \in \mathbb{N}$ of the $L^2(\mu)$-norm of multiple ergodic averages with iterates taken from the family of sequences

$$F_{a,r} := \{\text{non-constant polynomials of the form } a_i(n)-a(n),\ a_i(n+r)-a(n), i=1,\dots,\ell\} \tag{5}$$

where $a \in F$ is fixed (so in particular independent of $r \in \mathbb{N}$) and chosen so that the family $F_{a,r}$ has smaller "complexity" than F except possibly for a finite number of $r \in \mathbb{N}$.

[6] This rough plan is already implicit in the foundational paper of Furstenberg [35], the only difference is that in [35] the role of nilsystems played the much larger class of distal systems. Depending on the problem, this approach may offer some advantages, but for several recent applications it appears that the class of distal systems is just too broad to deal with directly.

[7] This states that if $a(1),\dots,a(N)$ are complex numbers bounded by 1, then for every integer R between 1 and N we have $\big|\frac{1}{N}\sum_{n=1}^{N} a(n)\big|^2 \ll \frac{1}{R}\sum_{r=1}^{R}(1-rR^{-1})\Re\big(\frac{1}{N}\sum_{n=1}^{N} a(n+r)\cdot \bar{a}(n)\big) + R^{-1} + RN^{-1}$ where $\Re(z)$ denotes the real part of z.

To be able to carry out this plan one first needs to take care of some elementary, but often not so easy, preparatory steps:

- Define a suitable collection $\mathcal{G}_0$ of families of sequences for which the desired seminorm estimates are easy to obtain directly.
- Define a suitable collection $\mathcal{G}$ of families of sequences that contains $\mathcal{G}_0$ and the family $\{a_1,\ldots,a_\ell\}$.
- Define a notion of equivalence and then a partial order $\preceq$ in $\mathcal{G}$ so that: (a) every decreasing sequence (G_n), with $G_n\in\mathcal{G}$, is eventually constant, and (b) if $G\in\mathcal{G}\setminus\mathcal{G}_0$, then there exists $b\in G$ such that $G_{b,r}\in\mathcal{G}$ and $G_{b,r}\prec G$ for all but finitely many $r\in\mathbb{N}$ ($G_{b,r}$ is defined as in (5)).

These conditions guarantee that there exists $d\in\mathbb{N}$ and an appropriate choice of sequences $b_1,b_2,\ldots,b_d$, such that the iteration

$$G\mapsto G^{r_1}:=G_{b_1,r_1}\mapsto G^{r_1,r_2}:=G^{r_1}_{b_2,r_2}\mapsto\cdots\mapsto G^{r_1,\ldots,r_d}:=G^{r_1,\ldots,r_{d-1}}_{b_d,r_d}$$

(not to be confused with an exact sequence!) produces families $G^{r_1,\ldots,r_d}$ that belong to $\mathcal{G}_0$ for a set $(r_1,\ldots,r_d)\in\mathbb{Z}^d$ that is big enough for our purposes. Practically, this means that after applying van der Corput's estimate a finite number of times, we have good chances to be able to bound the left hand side in (4) by a much simpler expression for which we can prove the desired seminorm estimates directly.

This strategy has been employed successfully in several instances and produced seminorm estimates of the form (4) for linear sequences [44], polynomial sequences [45, 51], block polynomials of fixed degree [34], and some sequences coming from smooth functions of polynomial growth [9, 26]. Notice a common desirable feature that these sequences share: after taking successive differences (meaning iterating the operation $a(n)\mapsto a(n+r)-a(n))$) a finite number of times we arrive to sequences that are either constant or piecewise asymptotically constant. This feature is not shared by several other sequences worth studying, for example, random sequences of integers, the sequence of primes, and the sequences $([n^{\log n}])$, $([n\sin n])$. In such cases one has to modify the PET induction approach or abandon it altogether and try something different.[8]

1.5.2. *An example.* Let F be a family of essentially distinct polynomials, meaning, all polynomials and their pairwise differences are not constant. Then one can define as $\mathcal{G}_0$ the collection of all families consisting of a single linear polynomial, for such families establishing an estimate of the form (4) is easy, and as $\mathcal{G}$ the collection of all families of essentially distinct polynomials.

The tricky part is to define a partial order in $\mathcal{G}$ that satisfies the third requirement mentioned in Section 1.5.1. This is done as follows: First, define the degree d of a family G of non-constant polynomials to be the maximum of the degrees of the polynomials in the family. Next, let G_i be the subfamily of polynomials of degree i in G, and let w_i denote the number of distinct leading coefficients that appear in the family G_i. The vector $(d,w_d,\ldots,w_1)$ is going to be the complexity of the family G. We identify two families that have the same complexity and we order the set of all possible complexities lexicographically, meaning, $(d,w_d,\ldots,w_1)>(d',w'_{d'},\ldots,w'_1)$ if and only if in the first instance where the two vectors disagree the coordinate of the first vector is greater than the coordinate of the second vector. We order the (equivalence

[8]For instance, for the first two sequences, it turns out to be more effective to utilize the random features of the sequences at hand, and get seminorm estimates by comparing the corresponding multiple ergodic averages with other deterministic ones that are better understood.

classes) of families of polynomials accordingly. One easily verifies that every decreasing sequence of complexities is eventually constant, and if $G \in \mathcal{G} \setminus \mathcal{G}_0$, then there exists $p \in G$ such that the family $G_{p,r}$ is in $\mathcal{G}$ and has complexity strictly smaller than that of G for all but finitely many $r \in \mathbb{N}$.

Using this strategy, seminorm estimates similar to those in (4) were established [45] for all essentially distinct polynomials except for a few cases (one has to dig into the details to see why this argument misses some cases) that were handled in [51] (for alternate proofs see Lemma 4.7 in [26] or Theorem 1.4 in [20]).

1.6. **Equidistribution of polynomial sequences on nilmanifolds.** Let $X = G/\Gamma$ be a nilmanifold, $b_1, \dots, b_\ell \in G$ and $x \in X$, and $a_1, \dots, a_\ell \colon \mathbb{N} \to \mathbb{Z}$ be sequences. In several of the applications we have in mind one is at some point called to prove that the sequence $(g(n)x)$ defined by $g(n) := b_1^{a_1(n)} \cdot \ldots \cdot b_\ell^{a_\ell(n)}$,[9] is equidistributed on some subset Y of X.[10]

Often Y is X, or a sub-nilmanifold of X, and in a few cases a union of sub-nilmanifolds of X. Such problems are typically much easier to handle when $X = \mathbb{T}^d$, since in this case one can utilize Weyl's equidistribution theorem[11] in order to reduce matters to estimating certain exponential sums. Unfortunately, such a convenient reduction is not available for all nilmanifolds, and checking equidistribution in this broader setup can be very challenging even for simple sequences.[12] Luckily, the situation is much better understood when all the sequences $a_1, \dots, a_\ell$ are given by integer polynomials, in this case we call the sequence $(g(n)x)$ *polynomial*; next, we are going to state some key results.

In the forthcoming discussion we assume that $X = G/\Gamma$ is a connected nilmanifold. By G_0 we denote the connected component of the identity element in G,[13] and we let $Z := G/([G_0, G_0]\Gamma)$ and $\pi \colon X \to Z$ be the natural projection. It is important to notice that Z has much simpler structure than X. Indeed, if G is connected, then Z is a connected compact Abelian Lie group, hence, a torus (meaning $\mathbb{T}^d$ for some $d \in \mathbb{N}$), and as a consequence every nilrotation in Z is (isomorphic to) a rotation on some torus. In general, the nilmanifold Z may be more complicated, but it is the case that every nilrotation in Z is (isomorphic to) a unipotent affine transformation on some torus[14] (see Proposition 3.1 in [29]). Iterates of such transformations can be computed explicitly[15], so one is much more comfortable to be dealing

---

[9] Such sequences cover as special cases sequences of the form $\big((c_1^{a_1(n)} x_1, \dots, c_\ell^{a_\ell(n)} x_\ell)\big)$, defined on the product of the nilmanifolds $X_1, \dots, X_\ell$. To see this let $X := X_1 \times \cdots \times X_\ell$, $x := (x_1, \dots, x_\ell)$, $b_1 := (c_1, e_2, \dots, e_l)$, …, $b_\ell := (e_1, \dots, e_{\ell-1}, c_\ell)$ where e_i denotes the identity element of the group G_i.

[10] If X is a nilmanifold we say that a sequence $(g(n))$ is equidistributed in a sub-nilmanifold Y (suppose that $g(n) \in Y$ for $n \in \mathbb{N}$) of X if for every $f \in C(Y)$ one has $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N f(g(n)) = \int f \, dm_Y$ where m_Y denotes the Haar measure on Y.

[11] This states that a sequence $(g(n))$ is equidistributed on a sub-nilmanifold Y of a torus $\mathbb{T}^d$ (suppose that $g(n) \in Y$ for $n \in \mathbb{N}$) if and only if for every non-trivial character $\chi \colon Y \to \mathbb{C}$ one has $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \chi(g(n)) = 0$.

[12] The use of representation theory has not proven to be of much help in this case.

[13] For technical reasons we assume that G_0 is simply connected and that $G = G_0\Gamma$. When X is connected, we can always arrange so that G_0 has these additional properties.

[14] A map $T \colon \mathbb{T}^d \to \mathbb{T}^d$ is said to be *affine* if $T(x) = S(x) + b$ for some homomorphism S of $\mathbb{T}^d$ and $b \in \mathbb{T}^d$. The homomorphism $S \colon \mathbb{T}^d \to \mathbb{T}^d$ is said to be *unipotent* if there exists $n \in \mathbb{N}$ so that $(S - \mathrm{Id})^n = 0$. In this case we say that the affine transformation T is a unipotent affine transformation.

[15] If $T \colon \mathbb{T}^d \to \mathbb{T}^d$ is a unipotent affine transformation, and $x \in \mathbb{T}^d$, then one can easily check that the coordinates of the sequence $(T^n x)$ are polynomial in n.

with equidistribution problems that involve unipotent affine transformations on some torus than with general nil-transformations.

The following qualitative equidistribution results were established by A. Leibman in [49]:

- A polynomial sequence $(g(n)x)_{n\in\mathbb{N}}$ is always equidistributed in a finite union of sub-nilmanifolds of X.
- A polynomial sequence $(g(n)x)_{n\in\mathbb{N}}$ is equidistributed in X if and only if the sequence $(g(n)\pi(x))_{n\in\mathbb{N}}$ is equidistributed in Z.

The second statement gives a very effective way for checking equidistribution of polynomial sequences. We illustrate this with a simple example. Suppose that $b \in G$ is an ergodic nilrotation (meaning the transformation $x \mapsto bx$ is ergodic) and we want to show that the polynomial sequence $(b^{n^2}x)$ is equidistributed in X for every $x \in X$. In the case where G is connected the nilmanifold Z is a torus, therefore, according to the previous criterion, it suffices to show that if β is an ergodic element of $\mathbb{T}^d$ (this is the case if the coordinates of β are rationally independent), then for every $x \in X$ the sequence $(x + n^2\beta)$ is equidistributed in $\mathbb{T}^d$. This is a well known fact, and can be easily verified using Weyl's equidistribution theorem and van der Corput's fundamental estimate. If G is not necessarily connected, one needs to show that if $S\colon \mathbb{T}^d \to \mathbb{T}^d$ is an ergodic unipotent affine transformation, then the sequence $S^{n^2}x$ is equidistributed for every $x \in \mathbb{T}^d$.[16] Although this is somewhat harder to establish, it follows again by Weyl's equidistribution theorem modulo some straightforward computations.

In other cases[17] one needs quantitative variants of the previous qualitative equidistribution results. Such a result was proved by B. Green and T. Tao [39]; we are not going to give the precise statement here because this would require to introduce too much additional notation.

1.7. **Pleasant and magic extensions.** Motivated by the work of T. Tao [57], H. Towsner [59], T. Austin [3], and B. Host [42], introduced new tools that help us handle some multiple ergodic averages. In particular, a key conceptual breakthrough that first appeared in [3], is that in some instances by working with suitable extensions of a family of commuting measure preserving systems (called "pleasant extensions" in [3] and "magic extensions" in [42]), characteristic factors of the corresponding multiple ergodic averages may be chosen to have particularly simple structure, a structure that is not visible when one works with the original systems (the idea of passing to an extension in order to simplify some convergence problems already appears in [37]). This is a rather counterintuitive statement since characteristic factors of extensions are extensions of characteristic factors of the original systems, so we going to explain a simple instance where such an approach works. Suppose that one wants to prove mean convergence for the averages

$$A_N(T,S,f_i) := \frac{1}{N^2}\sum_{1\leq m,n\leq N} T^m f_1 \cdot S^n f_2 \cdot T^m S^n f_3$$

[16]In some cases, for instance, when one seeks to prove equidistribution of a sequence in some unspecified nilmanifold, one can use a lifting argument in order to reduce matters to the case where G is connected. Such a simplification does not seem to be possible for the example just presented (the lifting does not preserve the ergodicity assumption).

[17]For instance, when one seeks to study equidistribution properties of the sequence $(b^{[n^{3/2}]}x)$, or tries to prove uniform convergence for the sequence $(\frac{1}{N}\sum_{n=1}^N f(b^{n^2}x))$ to the integral of f, where $b \in G$ is an ergodic element and $f \in C(X)$.

where T, S are commuting measure preserving transformations acting on the probability space $(X, \mathcal{X}, \mu)$ and $f_1, f_2, f_3 \in L^\infty(\mu)$. Although an estimate that relates the $L^2(\mu)$-norm of these averages with the Gowers-Host-Kra seminorms of the individual functions with respect to either T or S is not feasible, the following estimate is valid

$$\|A_N(T, S, f_i)\|_{L^2(\mu)} \ll \min_{i=1,2,3} \|\!|f_i|\!\|_{T,S,\mu},$$

where

$$\|\!|f|\!\|_{T,S,\mu}^4 := \lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \int f \cdot T^m \bar{f} \cdot S^n \bar{f} \cdot T^m S^n f \, d\mu$$

(it is shown in [42] that $\|\!|f|\!\|_{T,S} = \|\!|f|\!\|_{S,T}$). Now, although factors of the original systems that control the seminorms $\|\!|\cdot|\!\|_{T,S,\mu}$ may not admit particularly neat structure, it is shown in [42] that there exists a new system (X^*, μ^*, T^*, S^*) that extends the system (X, μ, T, S) and enjoys the following key extra property (the term "magic extension" from [42] alludes to this property):

$$\|\!|f^*|\!\|_{T^*,S^*,\mu^*} = 0 \Leftrightarrow f^* \perp \mathcal{I}_{T^*} \vee \mathcal{I}_{S^*}$$

where $f^* \in L^\infty(\mu^*)$ and $\mathcal{I}_T$ denotes the σ-algebra of T-invariant sets.[18] Notice also that mean convergence for the averages $A_N(T, S, f_i)$ follows if we prove mean convergence for all averages $A_N(T^*, S^*, f_i^*)$. Combining all these observations, we can easily reduce matters to proving mean convergence for the averages $A_N(T^*, S^*, f_i^*)$ when all functions f_i^* are $\mathcal{I}_{T^*} \vee \mathcal{I}_{S^*}$-measurable. This is a significant simplification of our original problem, and in fact it is now straightforward to deduce the required convergence property from the mean ergodic theorem.

This approach has proved particularly useful for handling convergence problems of multiple ergodic averages of commuting transformations with linear iterates that previously seemed intractable [3, 42, 4, 18, 19] (see also [7] for an application to continuous time averages). A drawback is that it does not give much information about the limiting function, and also, up to now, it has not proved to be as useful when some of the iterates are non-linear (for polynomial iterates though there is some progress in this direction [5, 6]).

1.8. **Furstenberg correspondence principle.** We frequently use the following correspondence principle of Furstenberg [35, 36] (the formulation given is from [8]) in order to reformulate statements in combinatorics as multiple recurrence statements in ergodic theory:

Furstenberg Correspondence Principle. *Let $\ell, d \in \mathbb{N}$, $E \subset \mathbb{Z}^d$ be a set of integers, and $\mathbf{v}_1, \ldots, \mathbf{v}_\ell \in \mathbb{Z}^d$. Then there exist a probability space $(X, \mathcal{B}, \mu)$, commuting invertible measure preserving transformations $T_1, \ldots, T_\ell \colon X \to X$, and a set $A \in \mathcal{B}$, with $\mu(A) = \bar{d}(E)$, and such that*

$$\bar{d}(E \cap (E - n_1\mathbf{v}_1) \cap \ldots \cap (E - n_\ell \mathbf{v}_\ell)) \geq \mu(A \cap T_1^{-n_1} A \cap \cdots \cap T_\ell^{-n_\ell} A), \tag{6}$$

for every $n_1, \ldots, n_\ell \in \mathbb{Z}$. Furthermore, if $\mathbf{v}_1 = \cdots = \mathbf{v}_\ell$ one can take $T_1 = \cdots = T_\ell$.

We are going to mention several applications of this principle in the next two subsections.

[18] In fact, one can take $X^* := X^4$, $T^* := (\mathrm{id}, T, \mathrm{id}, T)$, $S^* := (\mathrm{id}, \mathrm{id}, S, S)$, and define the measure μ^* by $\int f_1 \otimes f_2 \otimes f_3 \otimes f_4 \, d\mu^* := \lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \int f_1 \cdot T^m f_2 \cdot S^n f_3 \cdot T^m S^n f_4 \, d\mu$.

1.9. **Equivalent problems for sequences.** It turns out, and sometimes it is useful to be aware of this observation, that problems about mean convergence and multiple recurrence in ergodic theory are intimately related with similar problems involving bounded sequences of complex numbers. We give some explicit examples below.

Given a collection of sequences of integers $\{a_1, \ldots, a_\ell\}$, it turns out that the following two properties are equivalent:

- For every invertible measure preserving system $(X, \mathcal{X}, \mu, T)$, and $A \in \mathcal{X}$ with $\mu(A) > 0$, there exists $n \in \mathbb{N}$, such that
$$\mu(A \cap T^{-a_1(n)}A \cap \ldots \cap T^{-a_\ell(n)}A) > 0.$$
- For every non-negative bounded sequence $(z(n))$ with $\limsup_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} z(m) > 0$, there exists $n \in \mathbb{N}$, such that
$$\limsup_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} z(m) \cdot z(m + a_1(n)) \cdot \ldots \cdot z(m + a_\ell(n)) > 0.$$

Using the correspondence principle of Furstenberg it is not hard to see that the first statement implies the second. To see that the second statement implies the first it suffices to set $z(m) := \mathbf{1}_A(T^m x)$ for a suitable point $x \in X$ (μ-almost every $x \in X$ works) and use the mean ergodic theorem.

For convergence problems it is convenient to define the following notion: a sequence of complex numbers $(z(n))$ is *stationary with respect to the sequence of intervals* $([1, M_k])$, if for every $\ell \in \mathbb{N}$ and $n_1, \ldots, n_\ell \in \mathbb{Z}$, the averages
$$\frac{1}{M_k} \sum_{m=1}^{M_k} z(m + n_1) \cdot \ldots \cdot z(m + n_\ell)$$
converge as $k \to \infty$. Using a diagonal argument it is easy to show that any bounded sequence of complex numbers is stationary with respect to some subsequence of intervals. Given a collection of sequences of integers $\{a_1, \ldots, a_\ell\}$, it turns out that the following two properties are equivalent:

- For every invertible measure preserving system $(X, \mathcal{X}, \mu, T)$, and non-negative $f \in L^\infty(\mu)$, the averages
$$\frac{1}{N} \sum_{n=1}^{N} \int f \cdot T^{a_1(n)} f \cdot \ldots \cdot T^{a_\ell(n)} f \ d\mu$$
converge as $N \to \infty$.
- For every bounded sequence $(z(n))$ of complex numbers, stationary with respect to the sequence of intervals $([1, M_k])$, the averages
$$\frac{1}{N} \sum_{n=1}^{N} \Big(\lim_{k\to\infty} \frac{1}{M_k} \sum_{m=1}^{M_k} z(m) \cdot z(m + a_1(n)) \cdot \ldots \cdot z(m + a_\ell(n)) \Big)$$
converge as $N \to \infty$.

One can get similar statements for mean convergence, as well as for convergence and recurrence properties of commuting transformations (in this case one has to use sequences in several variables).

References

[1] I. Assani. Pointwise convergence of ergodic averages along cubes. *J. Analyse Math.* **110** (2010), 241–269.

[2] L. Auslander, L. Green, F. Hahn. Flows on homogeneous spaces. With the assistance of L. Markus and W. Massey, and an appendix by L. Greenberg, *Annals of Mathematics Studies*, **53**, Princeton University Press, Princeton, N.J. (1963).

[3] T. Austin. On the norm convergence of nonconventional ergodic averages. *Ergodic Theory Dynam. Systems* **30** (2010), 321–338.

[4] T. Austin. Deducing the multidimensional Szemerédi Theorem from an infinitary removal lemma. *J. Analyse Math.* **111** (2010), 131–150.

[5] T. Austin. Pleasant extensions retaining algebraic structure, I. Preprint, `arXiv:0905.0518v4`.

[6] T. Austin. Pleasant extensions retaining algebraic structure, II. Preprint, `arXiv:0910.0907v3`.

[7] T. Austin. Norm convergence of continuous-time polynomial multiple ergodic averages. Preprint, `arXiv:1103.0223`.

[8] V. Bergelson. Ergodic Ramsey Theory. *Logic and Combinatorics* (editted by S. Simpson). *Contemporary Mathematics* **65**, (1987), 63–87.

[9] V. Bergelson, I. Håland-Knutson. Weak mixing implies mixing of higher orders along tempered functions. *Ergodic Theory Dynam. Systems* **29** (2009), no. 5, 1375–1416

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

[11] V. Bergelson, B. Host, R. McCutcheon, F. Parreau. Aspects of uniformity in recurrence. *Colloq. Math.* **84/85** (2000), no. 2, 549–576.

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

[13] V. Bergelson, A. Leibman, E. Lesigne. Intersective polynomials and the polynomial Szemerédi theorem. *Adv. Math.* **219** (2008), no. 1, 369–388.

[14] V. Bergelson, A. Leibman, C. Moreira. From discrete- to continuous-time ergodic theorems. *Preprint.* Available at http://www.math.osu.edu/ leibman/preprints/

[15] V. Bergelson, A. Leibman, T. Ziegler. The shifted primes and the multidimensional Szemerédi and polynomial van der Waerden Theorems. To appear in *Comptes Rendus Mathematique*, `arXiv:1007.1839`.

[16] V. Bergelson, R. McCutcheon. Idempotent ultrafilters, multiple weak mixing and Szemerédi's theorem for generalized polynomials. *J. Analyse Math.* **111** (2010), 77–130.

[17] Q. Chu. Convergence of weighted polynomial multiple ergodic averages. *Proc. Amer. Math. Soc.* **137** (2009), no. 4, 1363-1369.

[18] Q. Chu. Convergence of multiple ergodic averages along cubes for several commuting transformations. *Studia Math.* **196** (2010), no. 1, 13-22.

[19] Q. Chu. Multiple recurrence for two commuting transformations. To appear *Ergodic Theory Dynam. Systems*, `arXiv:0912.3381`.

[20] Q. Chu, N. Franzikinakis. Pointwise convergence for cubic and polynomial ergodic averages of non-commuting transformations. To appear in *Ergodic Theory Dynam. Systems*, `arXiv:1006.5239`.

[21] Q. Chu, N. Franzikinakis, B. Host. Ergodic averages of commuting transformations with distinct degree polynomial iterates. To appear in *Proc. Lond. Math. Soc.*, `arXiv:0912.2641`.

[22] J-P. Conze, E. Lesigne. Sur un théorème ergodique pour des mesures diagonales. *Probabilités, Publ. Inst. Rech. Math. Rennes*, 1987-1, Univ. Rennes I, Rennes, (1988), 1–31.

[23] N. Frantzikinakis. The structure of strongly stationary systems. *J. Analyse Math.* **93** (2004), 359–388.

[24] N. Frantzikinakis. Multiple ergodic averages for three polynomials and applications. *Trans. Amer. Math. Soc.* **360** (2008), no. 10, 5435–5475.

[25] N. Frantzikinakis. Equidistribution of sparse sequences on nilmanifolds. *J. Analyse Math.* **109** (2009), 1–43.

[26] N. Frantzikinakis. Multiple recurrence and convergence for Hardy sequences of polynomial growth. *J. Analyse Math.* **112** (2010), 79–135.

[27] N. Frantzikinakis, B. Host, B. Kra. The polynomial multidimensional Szemerédi Theorem along shifted primes. Preprint, `arXiv:1009.1484`.

[28] N. Frantzikinakis, M. Johnson, E. Lesigne, M. Wierdl. Powers of sequences and convergence. *Ergodic Theory Dynam. Systems* **30** (2010), no. 5, 1431–1456.

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

[30] N. Frantzikinakis, B. Kra. Convergence of multiple ergodic averages for some commuting transformations. *Ergodic Theory Dynam. Systems* **25** (2005), no. 3, 799–809.

[31] N. Frantzikinakis, B. Kra. Ergodic averages for independent polynomials and applications. *J. London Math. Soc.* **74** (2006), no. 1, 131–142.

[32] N. Frantzikinakis, E. Lesigne, M. Wierdl. Sets of k-recurrence but not $(k+1)$-recurrence. *Ann. Inst. Fourier (Grenoble)* **56** (2006), no. 4, 839–849.

[33] N. Frantzikinakis, E. Lesigne, M. Wierdl. Powers of sequences and recurrence. *Proc. Lond. Math. Soc.* (3) **98** (2009), no. 2, 504–530.

[34] N. Frantzikinakis, M. Wierdl. A Hardy field extension of Szemerédi's theorem. *Adv. Math.* **222**, (2009), 1–43.

[35] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Analyse Math.* **31** (1977), 204–256.

[36] H. Furstenberg. Recurrence in ergodic theory and combinatorial number theory. *Princeton University Press*, Princeton, 1981.

[37] H. Furstenberg, B. Weiss. A mean ergodic theorem for $(1/N)\sum_{n=1}^{N} f(T^n x) \quad g(T^{n^2}x)$. Convergence in ergodic theory and probability (Columbus, OH, 1993), Ohio State Univ. Math. Res. Inst. Publ., **5**, de Gruyter, Berlin, (1996), 193–227.

[38] W. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.* **11** (2001), 465–588.

[39] B. Green, T. Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. To appear in *Annals Math.*, `arXiv:0709.3562`.

[40] B. Green, T. Tao. The Mobious function is strongly orthogonal to nilsequences. To appear in *Annals. Math.*, `arXiv:0807.1736`.

[41] J. Griesmer. Abundant configurations in sumsets with one dense summand. *Preprint*, `arXiv:1011.4657`.

[42] B. Host. Ergodic seminorms for commuting transformations and applications. *S*tudia Math. **195** (1) (2009), 31–49.

[43] B. Host, B. Kra. Averaging along cubes. Modern dynamical systems and applications. *Cambridge Univ. Press*, Cambridge, (2004), 123–144.

[44] B. Host, B. Kra. Non-conventional ergodic averages and nilmanifolds. *Annals Math.* **161** (2005), 397–488.

[45] B. Host, B. Kra. Convergence of polynomial ergodic averages. *Isr. J. Math.* **149** (2005), 1–19.

[46] B. Host, B. Kra. Uniformity seminorms on l^∞ and applications. *J. Analyse Math.* **108** (2009), 219–276.

[47] B. Host, B. Kra. Nil-Bohr sets of integers. To appear in *Ergodic Theory Dynam. Systems.*, arXiv:0903.1642.

[48] M. Johnson. Convergence of polynomial ergodic averages for some commuting transformations. *Illinois J. Math.* **53** (2009), no. 3, 865-882.

[49] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of rotations of a nilmanifold. *Ergodic Theory Dynam. Systems* **25** (2005), no. 1, 201–213.

[50] A. Leibman. Pointwise convergence of ergodic averages for polynomial actions of $\mathbb{Z}^d$ by translations on a nilmanifold. *Ergodic Theory Dynam. Systems* **25** (2005), no. 1, 215–225.

[51] A. Leibman. Convergence of multiple ergodic averages along polynomials of several variables. *Isr. J. Math.* **146** (2005), 303–316.

[52] E. Lesigne. Théorèmes ergodiques pour une translation sur un nilvariété. *Ergodic Theory Dynam. Systems* **9** (1989), no. 1, 115-126.

[53] E. Lesigne. Sur une nil-variété, les parties minimales associées à une translation sont uniquement ergodiques. *Ergodic Theory Dynam. Systems* **11** (1991), no. 2, 379–391.

[54] E. Lesigne. Équations fonctionnelles, couplages de produits gauches et théorèmes ergodiques pour mesures diagonales. *Bull. Soc. Math. France* **121** (1993), no. 3, 315–351.

[55] A. Potts. Multiple ergodic averages for flows and an application. *Preprint*, `arXiv:0910.3687`.

[56] D. Rudolph. Eigenfunctions of $T \times S$ and the Conze-Lesigne algebra. Ergodic theory and its connections with harmonic analysis (Alexandria, 1993). *London Math. Soc. Lecture Note Ser.*, **205**, Cambridge Univ. Press, Cambridge, (1995), 369-432.

[57] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory Dynam. Systems* **28** (2008), no. 2, 657–688.

[58] T. Tao. Poincaré's legacies, pages from year two of a mathematical blog. Part I. *American Mathematical Society*, Providence, RI, 2009.

[59] H. Towsner. Convergence of diagonal ergodic averages. *Ergodic Theory Dynam. Systems* **29** (2009), 1309–1326.

[60] T. Wooley, T. Ziegler. Multiple recurrence and convergence along the primes. To appear in *Amer. J. of Math.*, `arXiv:1001.4081`.

[61] T. Ziegler. Nilfactors of $\mathbb{R}^m$-actions and configurations in sets of positive upper density in $\mathbb{R}^m$. *J. Analyse Math.* **99** (2006), 249–266.

[62] T. Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.* **20** (2007), 53–97.

(Nikos Frantzikinakis) University of Crete, Department of mathematics, Knossos Avenue, Heraklion 71409, Greece

E-mail address: `frantzikinakis@gmail.com`