SOME OPEN PROBLEMS ON MULTIPLE ERGODIC AVERAGES

NIKOS FRANTZIKINAKIS

Contents

1. Introduction

1.1. **What is this?** This is a first attempt to organize a list of some open problems on three closely related topics:

(1) The limiting behavior of single and multiple ergodic averages.

(2) Single and multiple recurrence properties of measure preserving systems.

(3) Universal patterns, meaning, patterns that can be found in every set of integers with positive upper density, and related problems on higher dimensions.

The list of problems is greatly influenced by my personal interests and is by no means meant to be a comprehensive list of open problems in the area widely known as ergodic Ramsey theory. Almost exclusively, problems related to actions of commuting measure preserving transformations are considered, and even within this confined class there are a few important topics not touched upon (for instance, the richness of the return times in various multiple recurrence results). For material and a list of problems that goes beyond the scope of this set of notes we refer the reader to the survey articles [24, 27, 28] and the references therein.

Whenever appropriate, I include the "original source" and a short history of the problem, as well as a hopefully accurate and up to date list of related work already done. I plan to update these notes from time to time, so you are welcome to contact me and help me improve them.

1.2. **The general theme.** A very general framework that can be used to describe the bulk of the problems listed below is the following: We are given a measure space $(X, \mathcal{X}, \mu)$ with $\mu(X) = 1$, invertible measure preserving transformations $T_1, \ldots, T_\ell \colon X \to X$, bounded measurable functions $f_1, \ldots, f_\ell \colon X \to \mathbb{C}$, and sequences $a_1, \ldots, a_\ell \colon \mathbb{N} \to \mathbb{Z}$. In some cases we also consider sequences that depend on several integer variables, but let's stick to the single variable case for the time being.

The first family of problems concerns the study of the limiting behavior of the so called multiple ergodic averages

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

where $Tf := f \circ T$ and $T^k := T \circ \cdots \circ T$. One would like to know whether these averages converge as $N \to \infty$ (in $L^2(\mu)$ or pointwise), find some structured factors that control their limiting behavior (called characteristic factors), and if possible, find a formula, or a usable way to extract information, for the limiting function. When $\ell = 1$, such problems have been studied extensively and in several cases solved (see the survey paper [170] for a variety of related results). Our main concern here is to study the averages (1) when $\ell \geq 2$. To get manageable problems, one typically restricts the class of eligible sequences, usually to be polynomial sequences, sequences arising from smooth functions, sequences related to the prime numbers, or random sequences, and also assumes that the transformations commute, or, to get started, that they are all equal. On the other hand, because of the nature of the implications in combinatorics that we are frequently interested in, it is not desirable to assume anything about the particular structure of each individual measure preserving transformation. Typically, the tools used to attack such problems include (i) elementary uniformity estimates, (ii) ergodic structure theorems (like the one in [115]), and (*iii*) equidistribution results on nilmanifolds. More on that on subsequent sections.

The second family of problems concerns the study of expressions of the form

$$\mu(A \cap T_1^{-a_1(n)} A \cap \cdots \cap T_\ell^{-a_\ell(n)} A) \tag{2}$$

where $A \in \mathcal{X}$ has positive measure. One wants to know whether such expressions are positive for some $n \in \mathbb{N}$, or even better, for lots of $n \in \mathbb{N}$ (for instance on the average), and if possible, get some explicit lower bound that depends only on the measure of the set A and on ℓ (optimally this is going to be of the form $(\mu(A))^{\ell+1}$). Such multiple recurrence results are typically obtained by carrying out an in depth analysis of the limiting behavior of the averages (1). Usually they are not hard if an explicit formula of the limiting function is known, but they can be very tricky in the absence of such a formula, even when we work with very special systems of algebraic nature.

Concerning the third family of problems, and restricting ourselves to subsets of $\mathbb{Z}$, one is interested to know, for example, whether every set of integers with positive upper density[1] contains patterns of the form

$$m,\ m + a_1(n), \dots,\ m + a_\ell(n)$$

for some (or lots of) $m \in \mathbb{Z}$ and $n \in \mathbb{N}$. A typical instance is the celebrated theorem of Szemerédi [177] stating that every set of integers with positive upper density contains arbitrarily long arithmetic progressions (this corresponds to the case $a_i(n) = in$ for $i = 1, \dots, \ell$, $\ell \in \mathbb{N}$). Using a correspondence principle of H. Furstenberg, one can translate such statements to multiple recurrence statements in ergodic theory; an equivalent problem is then to show that the expressions (2) are positive for some $n \in \mathbb{N}$ when all the measure preserving transformations $T_1, \dots, T_\ell$ are equal. Similar questions can be asked on higher dimensions, concerning patterns that can be found on subsets of $\mathbb{Z}^d$ with positive upper density. Such questions correspond to multiple recurrence statements when the transformations $T_1, \dots, T_\ell$ commute. This approach was originally used by H. Furstenberg in [91] to give an alternate proof of Szemerédi's theorem using ergodic theory. Subsequently, H. Furstenberg and Y. Katznelson gave the first proof of the multidimensional Szemerédi theorem [97] and the density Hales-Jewett theorem [100], and V. Bergelson and A. Leibman proved the polynomial extension of Szemerédi's theorem [35] (currently no proof that avoids ergodic theory is known for this result). And the story does not end there, in the last two decades new powerful tools in ergodic theory were developed and used, and are currently being used, to prove several other deep results in density Ramsey theory. The reader will find several such applications in subsequent sections and the extended bibliography section. Several additional applications can be found in the survey articles and the references therein [24, 27, 28].

Recently, an additional motivation for studying such problems has surfaced. It has to do with potential implications in number theory, in particular a connection to problems of finding patterns in the set of the prime numbers. Knowing that every set of integers with positive upper density contains patterns of a certain sort could be an important first step towards proving an analogous result for the set of primes. This idea originates from work of B. Green and T. Tao [104], where it was used to show that the primes contain arbitrarily long arithmetic

[1]The *upper density* $\bar{d}(E)$ of a set $E \subset \mathbb{Z}^d$ is defined by $\bar{d}(E) := \limsup_{N\to\infty} \frac{|E\cap[-N,N]^d|}{|[-N,N]^d|}$.

progressions. It was also subsequently used by T. Tao and T. Ziegler [183] to show that the primes contain arbitrarily long polynomial progressions.

1.3. **General conventions and notation.** We are going to use the following notation: $\mathbb{N} := \{1,2,\ldots\}$, $T^k := T \circ \cdots \circ T$, $Tf := f \circ T$, if $(X, \mathcal{X}, \mu, T)$ is a measure preserving system then $\mathcal{I}_T := \{A \in \mathcal{X} \colon T^{-1}A = A\}$, $\mathcal{K}_{\text{rat}} := \bigvee_{d\in\mathbb{N}} \mathcal{I}_{T^d}$. We use the symbol $\ll$ when some expression is majorized by a constant multiple of some other expression.

Acknowledgment. I would like to thank B. Kra and E. Lesigne for helpful comments.

2. Some useful tools and observations

2.1. **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 [91], but the term "characteristic factor" was coined in a paper of H. Furstenberg and B. Weiss [101].

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

$$A_N(f_1, \ldots, f_\ell) := \frac{1}{N}\sum_{n=1}^{N} T_1^{a_1(n)} f_1 \cdot \ldots \cdot T_\ell^{a_\ell(n)} f_\ell \tag{3}$$

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 $i = 1, \ldots, \ell$.[2]

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 (3).

2.2. **Gowers-Host-Kra seminorms.** When analyzing the limiting behavior of the averages (3), 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 (3) by the *Gowers-Host-Kra seminorms.* Similar seminorms were first introduced in combinatorics by T. Gowers [103] and their ergodic variant (that is more relevant for our study) was introduced by B. Host and B. Kra [115]. For an ergodic system $(X, \mathcal{X}, \mu, T)$ and function $f \in L^\infty(\mu)$, they are defined as follows:

$$|\!|\!| f |\!|\!|_1 := \left| \int f \ d\mu \right| ;$$

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

It is shown in [115] 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

[2]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*[3], then $\|\!| f \|\!|_k = |\int f \, d\mu|$ for every $k \in \mathbb{N}$.

2.3. **The factors $\mathcal{Z}_k$ and their structure.**

The seminorms $\|\!| \cdot \|\!|_k$ induce T-invariant sub-σ-algebras $\mathcal{Z}_{k-1}$ that satisfy

$$\text{(4)} \qquad \textit{for} \ \ f \in L^\infty(\mu), \ \ \mathbb{E}(f|\mathcal{Z}_{k-1}) = 0 \quad \textit{if and only if} \quad \|\!| f \|\!|_k = 0.$$

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

$$\text{(5)} \qquad \limsup_{N\to\infty} \|A_N(f_1, \ldots, f_\ell)\|_{L^2(\mu)} \ll \min_{i=1,\ldots,\ell} \|\!| f_i \|\!|_{k_i, T_i},$$ [4]

then one knows that the factors $\mathcal{Z}_{k_1-1,T_1}, \ldots, \mathcal{Z}_{k_\ell-1,T_\ell}$ are characteristic for mean convergence of the averages (3). 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 [115] stating that for ergodic systems the factor system $(X, \mathcal{Z}_k, \mu, T)$ is an inverse limit of k-step nilsystems[5]. 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

- $f = f_s + f_u + f_e$;
- $\|\!| f_u \|\!|_{k+1} = 0$; $\|f_e\|_{L^1(\mu)} \leq \varepsilon$; and
- $(f_s(T^n x))_{n\in\mathbb{N}}$ is a k-step nilsequence[6] for μ-almost every $x \in X$.

Such a decomposition also holds for non-ergodic systems (see Proposition 3.1 in [62]).

Combining the hypothetical seminorm estimates (5) with the aforementioned structure theorem (or the decomposition result), the problem of analyzing the limiting behavior of the averages (3) 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 [8, 78, 106, 107, 140, 141, 152, 153].

2.4. **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

---

[3]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 $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^N |\int \bar{f} \cdot T^n f \, d\mu|^2 = 0$.

[4]As a general principle, if one can show that $A_N(f_1, \ldots, f_\ell) \to^{L^2(\mu)} 0$ when all the transformations are weak mixing and one of the functions has mean 0, then one can adapt its proof to show (5).

[5]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.

[6]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 (3), very often a useful approach is to try to work out the following three steps:[7]

- Produce seminorm estimates like the ones in (5).
- 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: [7, 33, 40, 41, 58, 61, 62, 64, 76, 77, 79, 82, 83, 84, 85, 87, 90, 101, 108, 114, 115, 116, 119, 120, 124, 142, 154, 168, 173, 191, 195, 196]. 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[8] (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.

2.5. **The polynomial exhaustion technique.** We explain a technique that is often used to produce seminorm estimates of the type (5). It is based on an induction scheme (often called PET induction) introduced by V. Bergelson in [22]. 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

$$\limsup_{N\to\infty} \|A_N(f_1, \dots, f_\ell)\|_{L^2(\mu)} \ll \min_{i=1,\dots,\ell} |||f_i|||_{k_i} \tag{6}$$

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.

2.5.1. *The method.* The main idea is to use some variation of van der Corput's fundamental estimate and bound the left hand side in (6) 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

$$F_{a,r} := \{\text{non-constant polynomials of the form } a_i(n) - a(n),\ a_i(n+r) - a(n), i = 1, \dots, \ell\} \tag{7}$$

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 $r \in \mathbb{N}$.

[7] This rough plan is already implicit in the foundational paper of Furstenberg [91], the only difference is that in [91] 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.

[8] 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 (7)).

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 (6) 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 (6) for linear sequences [115], polynomial sequences [116, 142], block polynomials of fixed degree [90], and some sequences coming from smooth functions of polynomial growth [31, 79]. 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.[9]

2.5.2. *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 (6) 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 2.5.1. 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

[9] 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.

of the first vector is greater than the coordinate of the second vector. We order the (equivalence 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 (6) were established [116] 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 [142] (for alternate proofs see Lemma 4.7 in [79] or Theorem 1.4 in [61]).

2.6. **Equidistribution of polynomial sequences on nilmanifolds.** Let $X = G/\Gamma$ be a nilmanifold, $b_1, \ldots, b_\ell \in G$ and $x \in X$, and $a_1, \ldots, a_\ell \colon \mathbb{N} \to \mathbb{Z}$ 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 $g(n) := b_1^{a_1(n)} \cdot \ldots \cdot b_\ell^{a_\ell(n)}$,[10] is equidistributed on some subset Y of X.[11]

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[12] 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.[13] Luckily, the situation is much better understood when all the sequences $a_1, \ldots, a_\ell$ 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,[14] 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[15] (see Proposition 3.1 in [83]). Iterates of such

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

[11]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.

[12]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 $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \chi(g(n)) = 0$.

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

[14]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.

[15]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.

transformations can be computed explicitly[16], so one is much more comfortable to be dealing 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 [140]:

- 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 $x \in \mathbb{T}^d$.[17] Although this is somewhat harder to establish, it follows again by Weyl's equidistribution theorem modulo some straightforward computations.

In other cases[18] one needs quantitative variants of the previous qualitative equidistribution results. Such a result was proved by B. Green and T. Tao [106]; we are not going to give the precise statement here because this would require to introduce too much additional notation.

2.7. **Pleasant and magic extensions.** Motivated by the work of T. Tao [181], H. Towsner [186], T. Austin [9], and B. Host [111], introduced new tools that help us handle some multiple ergodic averages. In particular, a key conceptual breakthrough that first appeared in [9], is that in some instances by working with suitable extensions of a family of commuting measure preserving systems (called "pleasant extensions" in [9] and "magic extensions" in [111]), 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 [101]). 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

[16] 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.

[17] 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).

[18] 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 $f \in C(X)$.

mean convergence for the averages

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

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

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

(it is shown in [111] 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 [111] 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 [111] 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.[19] 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 [9, 111, 10, 59, 60] (see also [13] 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 [11, 12]).

2.8. **Furstenberg correspondence principle.** We frequently use the following correspondence principle of Furstenberg [91, 92] (the formulation given is from [23]) 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*

$$\bar{d}(E\cap(E-n_1\mathbf{v}_1)\cap\ldots\cap(E-n_\ell\mathbf{v}_\ell))\geq\mu(A\cap T_1^{-n_1}A\cap\cdots\cap T_\ell^{-n_\ell}A), \tag{8}$$

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

[19] 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 $\int f_1\otimes f_2\otimes f_3\otimes f_4\ d\mu^* := \lim_{M\to\infty}\frac{1}{M}\sum_{m=1}^M\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N\int f_1\cdot T^mf_2\cdot S^nf_3\cdot T^mS^nf_4\ d\mu$.

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

2.9. **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).

3. Some useful notions

To ease our exposition we collect some notions that are frequently used in subsequent sections.

3.1. **Recurrence.** The next notions are used to describe multiple recurrence properties of a single sequence:

Definition 3.1. We say that the sequence of integers $(a(n))$ is

- *good for ℓ-recurrence of powers* if for every $k_1, \ldots, k_\ell \in \mathbb{Z}$, every invertible measure preserving system $(X, \mathcal{B}, \mu, T)$, and set $A \in \mathcal{B}$ with $\mu(A) > 0$, we have
$$\mu(A \cap T^{-k_1 a(n)} A \cap \cdots \cap T^{-k_\ell a(n)} A) > 0$$
for some $n \in \mathbb{N}$ with $a(n) \neq 0$.[20]
- *good for multiple recurrence of powers* if it is good for ℓ-recurrence of powers for every $\ell \in \mathbb{N}$.
- *good for ℓ-recurrence of commuting transformations* if for every probability space $(X, \mathcal{X}, \mu)$, commuting invertible measure preserving transformations $T_1, \ldots, T_\ell \colon X \to X$, and set $A \in \mathcal{B}$ with $\mu(A) > 0$, we have
$$\mu(A \cap T_1^{-a(n)} A \cap \cdots \cap T_\ell^{-a(n)} A) > 0$$
for some $n \in \mathbb{N}$ with $a(n) \neq 0$.
- *good for multiple recurrence of commuting transformations* if it is good for ℓ-recurrence of commuting transformations for every $\ell \in \mathbb{N}$.

The fact that the sequence (n) is good for multiple recurrence of powers corresponds to the multiple recurrence result of H. Furstenberg [91], and the fact that it is good for multiple recurrence of commuting transformations corresponds to the multidimensional extension of this result of H. Furstenberg and Y. Katznelson [97]. Further examples of sequences that are good for multiple recurrence of commuting transformations include: integer polynomials with zero constant term [35], the shifted primes ([191] for powers and [42] in general), integer polynomials with zero constant term evaluated at the shifted primes ([191] for powers and [81] in general), several generalized polynomial sequences [44], and some random sequences of zero density [88]. The sequence $([n^c])$, where $c \in \mathbb{R}$ is positive, is known to be good for multiple recurrence of powers [90] (see also [79]), but if $c \in \mathbb{R} \setminus \mathbb{Q}$ is greater than 1, it is not known whether it is good for multiple recurrence of commuting transformations. Examples of sequences that are good for ℓ-recurrence of powers but are not good for $(\ell+1)$-recurrence of powers can be found in [86].

Let us also mention at this point some obstructions to recurrence. One can check that if the sequence $(a(n))$ is good for 1-recurrence, then the equation $a(n) \equiv 0 \mod r$ has a solution

[20]We remark, that for this and subsequent statements, the existence of a single $n \in \mathbb{N}$ for which the multiple intersection has positive measure, forces the existence of infinitely many $n \in \mathbb{N}$ with the same property (but not necessarily for a set of $n \in \mathbb{N}$ with positive upper density, although more often than not this is also true).

for every $r \in \mathbb{N}$; as a consequence, the sequences (p_n), (n^2+1), (2^n), are not good for 1-recurrence. More generally, if the sequence $(a(n))$ is good for 1-recurrence, then for every $\alpha \in \mathbb{R}$ and $\varepsilon > 0$ the inequality $\|a(n)\alpha\| \leq \varepsilon$ has a solution (letting $\alpha = 1/r$ gives the previous congruence condition); as a consequence, the sequence $([\sqrt{5}n+2])$ is not good for 1-recurrence (try $\alpha = 1/\sqrt{5}$ and $\varepsilon = 1/10$). This obstruction also implies that if a sequence $(a(n))$ is lacunary, meaning, $\liminf_{n\to\infty} a(n+1)/a(n) > 1$, then it is not good for 1-recurrence, since it is well known that in this case there exist an irrational α and a positive ε_0 such that $\|a(n)\alpha\| \geq \varepsilon_0$ for every $n \in \mathbb{N}$.

Furthermore, if a sequence $(a(n))$ is good for 2-recurrence of powers, then for every $\alpha \in \mathbb{R}$ and $\varepsilon > 0$ the inequality $\left\|a(n)^2\alpha\right\| \leq \varepsilon$ has a solution (this is implicit in [92] Section 9.1, and is proved in detail in [86]). It follows that if one forms a sequence by putting the elements of the set $\{n \in \mathbb{N} \colon \left\|n^2\sqrt{2}\right\| \in [1/4, 1/2]\}$ in increasing order, then this sequence is not going to be good for 2-recurrence of powers (one can show that it is going to be good for 1-recurrence [86]). And there are also other more restrictive obstructions, if the sequence $(a(n))$ is good for 2-recurrence of powers, then for every $\alpha, \beta \in \mathbb{R}$ and $\varepsilon > 0$ the inequality $\|[a(n)\alpha]a(n)\beta\| \leq \varepsilon$ has a solution. More generally, for ℓ-recurrence of powers, one can state further obstructions by using higher degree polynomials with zero constant term, or more complicated generalized polynomials.

The next notions are used to describe multiple recurrence properties of a collection of sequences:

Definition 3.2. We say that the collection of sequences of integers $\{(a_1(n)), \ldots, (a_\ell(n))\}$ is

- *good for ℓ-recurrence of a single transformation* if for every invertible measure preserving system $(X, \mathcal{B}, \mu, T)$, and set $A \in \mathcal{B}$ with $\mu(A) > 0$, we have
$$\mu(A \cap T^{-a_1(n)}A \cap \cdots \cap T^{-a_\ell(n)}A) > 0$$
for some $n \in \mathbb{N}$ with $a_i(n) \neq 0$ for $i = 1, \ldots, \ell$.
- *good for ℓ-recurrence of commuting transformations* if for every probability space $(X, \mathcal{X}, \mu)$, commuting invertible measure preserving transformations $T_1, \ldots, T_\ell \colon X \to X$, and set $A \in \mathcal{B}$ with $\mu(A) > 0$, we have
$$\mu(A \cap T_1^{-a_1(n)}A \cap \cdots \cap T_\ell^{-a_\ell(n)}A) > 0$$
for some $n \in \mathbb{N}$ with $a_i(n) \neq 0$ for $i = 1, \ldots, \ell$.

Examples of collections of ℓ sequences that are known to be good for ℓ-recurrence of commuting transformations include collections of: integer polynomials with zero constant term [35] (for arbitrary integer polynomials necessary and sufficient conditions where given in [40]) and integer polynomials with zero constant term evaluated at the shifted primes [81]. For every $\ell \in \mathbb{N}$, the collection $\{([n^{c_1}]), \ldots, ([n^{c_\ell}])\}$, where $c_1, \ldots, c_\ell \in \mathbb{R} \setminus \mathbb{Z}$, is known to be good for ℓ-recurrence of a single transformation [79], but it is not known whether it is always good for ℓ-recurrence of commuting transformations.

The correspondence principle of Furstenberg enables us to deduce statements in density Ramsey theory from multiple recurrence statements in ergodic theory. Using this principle and some elementary arguments one is able to reformulate the previous ergodic notions to purely combinatorial ones:

Theorem. *The sequence of integers $(a(n))$ is*

- *good for ℓ-recurrence of powers if and only if for every $k_1, \ldots, k_\ell \in \mathbb{Z}$, every set $E \subset \mathbb{Z}$ with $\bar{d}(E) > 0$ contains patterns of the form*

$$\{m, m + k_1 a(n), \ldots, m + k_\ell a(n)\}$$

for some $m \in \mathbb{Z}$ and $n \in \mathbb{N}$ with $a(n) \neq 0$.
- *good for ℓ-recurrence of commuting transformations if and only if for every $d \in \mathbb{N}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_\ell \in \mathbb{Z}^d$, every set $E \subset \mathbb{Z}^d$ with $\bar{d}(E) > 0$ contains patterns of the form*

$$\{\mathbf{m}, \mathbf{m} + a(n)\mathbf{v}_1, \ldots, \mathbf{m} + a(n)\mathbf{v}_\ell\}$$

for some $\mathbf{m} \in \mathbb{Z}^d$ and $n \in \mathbb{N}$ with $a(n) \neq 0$.

For instance, the fact that the sequence (n) is good for multiple recurrence of powers corresponds to the theorem of Szemerédi on arithmetic progressions.

Theorem. *The collection of sequences of integers $\{a_1(n), \ldots, a_\ell(n)\}$ is*

- *good for ℓ-recurrence of a single transformation if and only if every set $E \subset \mathbb{Z}$ with $\bar{d}(E) > 0$ contains patterns of the form*

$$\{m, m + a_1(n), \ldots, m + a_\ell(n)\}$$

for some $m \in \mathbb{Z}$ and $n \in \mathbb{N}$ with $a_i(n) \neq 0$ for $i = 1, \ldots, \ell$.
- *good for ℓ-recurrence of commuting transformations if and only if for every $\mathbf{v}_1, \ldots, \mathbf{v}_\ell \in \mathbb{Z}^d$, every set $E \subset \mathbb{Z}^d$ with $\bar{d}(E) > 0$ contains patterns of the form*

$$\{\mathbf{m}, \mathbf{m} + a_1(n)\mathbf{v}_1, \ldots, \mathbf{m} + a_\ell(n)\mathbf{v}_\ell\}$$

for some $\mathbf{m} \in \mathbb{Z}^d$ and $n \in \mathbb{N}$ with $a_i(n) \neq 0$ for $i = 1, \ldots, \ell$.

Let us also remark that the previous notions admit equivalent uniform versions that are often useful for applications. For instance, one can prove the following (see [34] for an argument that works for polynomials and [87] for an argument that works for general sequences):

Theorem. *The collection of sequences of integers $\{a_1(n), \ldots, a_\ell(n)\}$ is good for ℓ-recurrence of a single transformation if and only if*

(i) For every $\varepsilon > 0$ there exist $\delta = \delta(\varepsilon) > 0$ and $N_0 = N_0(\varepsilon)$, such that for every $N \geq N_0$ and integer set $E \subset [-N, N]$ with $|E| \geq \varepsilon N$, we have

$$|E \cap (E - a_1(n)) \cap \ldots \cap (E - a_\ell(n))| \geq \delta N$$

for some $n \in [1, N_0]$.

(ii) For every $\varepsilon > 0$ there exist $\gamma = \gamma(\varepsilon) > 0$ and $N_1 = N_1(\varepsilon)$, such that for every invertible measure preserving system $(X, \mathcal{B}, \mu, T)$ and $A \in \mathcal{B}$ with $\mu(A) \geq \varepsilon$, we have that

$$\mu(A \cap T^{-a_1(n)} A \cap \ldots \cap T^{-a_\ell(n)} A) \geq \gamma$$

for some $n \in [1, N_1]$.

3.2. **Convergence.** The next notion is used to describe multiple convergence properties of a single sequence:

Definition 3.3. We say that the sequence of integers $(a(n))$ is

- *good for ℓ-convergence of powers (of a single transformation)* if for every $k_1, \ldots, k_\ell \in \mathbb{Z}$, invertible measure preserving system $(X, \mathcal{X}, \mu, T)$, and functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$, the averages
$$\frac{1}{N}\sum_{n=1}^{N} T^{k_1 a(n)} f_1 \cdot \ldots \cdot T^{k_\ell a(n)} f_\ell$$
converge in the mean.
- *good for multiple convergence of powers (of a single transformation)* if it is good for ℓ-convergence of powers for every $\ell \in \mathbb{N}$.
- *good for ℓ-convergence of commuting transformations* if for every probability space $(X, \mathcal{X}, \mu)$, commuting invertible measure preserving transformations $T_1, \ldots, T_\ell \colon X \to X$, and functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$, the averages
$$\frac{1}{N}\sum_{n=1}^{N} T_1^{a(n)} f_1 \cdot \ldots \cdot T_\ell^{a(n)} f_\ell$$
converge in the mean.
- *good for multiple convergence of commuting transformations* if it is good for ℓ-convergence of commuting transformations for every $\ell \in \mathbb{N}$.

Examples of sequences that are good for multiple convergence of commuting transformations include: the sequence (n) [181] (see also [186, 9, 111]), the sequence of primes [81], and some random sequences of zero density [88]. Additional examples of sequences that are known to be good for multiple convergence of powers include: sequences given by integer polynomials [116, 141], integer polynomials evaluated at the primes [191] (see also [81]), and sequences of the form $([n^a])$ where $a > 0$ [79]. Examples of sequences that are good for ℓ-convergence of powers but are not good for $(\ell+1)$-convergence of powers can be found in [86].

The next notion is used to describe multiple convergence properties of a collection of sequences:

Definition 3.4. We say that the collection of sequences of integers $\{(a_1(n)), \ldots, (a_\ell(n))\}$ is

- *good for ℓ-convergence of a single transformation* if for every invertible measure preserving system $(X, \mathcal{X}, \mu, T)$, and functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$, the averages
$$\frac{1}{N}\sum_{n=1}^{N} T^{a_1(n)} f_1 \cdot \ldots \cdot T^{a_\ell(n)} f_\ell$$
converge in the mean.
- *good for ℓ-convergence of commuting transformations* if for every probability space $(X, \mathcal{X}, \mu)$, commuting invertible measure preserving transformations $T_1, \ldots, T_\ell \colon X \to X$, and functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$, the averages
$$\frac{1}{N}\sum_{n=1}^{N} T_1^{a_1(n)} f_1 \cdot \ldots \cdot T_\ell^{a_\ell(n)} f_\ell$$

converge in the mean.

Examples of collections of sequences (not coming from multiples of the same sequence) that are known to be good for ℓ-convergence of commuting transformations include: collections of integer polynomials with distinct degrees [62] or any such collection of polynomials evaluated at the primes [81], and collections of sequences of the form $\{(n), (a_n(\omega))\}$ where $(a_n(\omega))$ is a well chosen random sequence of zero density [88]. Additional examples of collections of sequences that are good for ℓ-convergence of a single transformation include: arbitrary collections of integer polynomials, or integer polynomials evaluated at the primes [191] (see also [81]), and the collection $\{([n^{c_1}]), \ldots, ([n^{c_\ell}])\}$, where $c_1, \ldots, c_\ell \in \mathbb{R} \setminus \mathbb{Z}$ are positive [79].

4. Problems related to general sequences

In this section we give a list of problems related to the study of multiple ergodic averages involving iterates given by general sequences and related applications to multiple recurrence results.

4.1. **The structure of multicorrelation sequences.** Given a probability space $(X, \mathcal{X}, \mu)$, commuting invertible measure preserving transformations $T_1, \ldots, T_\ell \colon X \to X$, and functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$, we are interested in determining the structure of the multiple correlation sequences, meaning sequences $\mathcal{C} \colon \mathbb{Z}^\ell \to \mathbb{C}$ defined by the formula

$$\mathcal{C}(n_1, \ldots, n_\ell) := \int f_0 \cdot T_1^{n_1} f_1 \cdot \ldots \cdot T_\ell^{n_\ell} f_\ell \, d\mu. \tag{9}$$

The next result gives a very satisfactory solution to this problem for $\ell = 1$, and is going to serve as our model for possible generalizations. It can be deduced from Herglotz's theorem on positive definite sequences (the sequence $n \mapsto \int \bar{f} \cdot T^n f \, d\mu$ is positive definite) and a polarization identity.

Theorem. *Let $(X, \mathcal{X}, \mu, T)$ be an invertible measure preserving system and $f, g \in L^\infty(\mu)$. Then there exists a complex Borel measure σ on $[0, 1)$, with bounded variation, such that for every $n \in \mathbb{N}$ we have*

$$\int f \cdot T^n g \, d\mu = \int_0^1 e^{2\pi i n t} \, d\sigma(t). \tag{10}$$

Finding a formula analogous to (10), with the multiple correlation sequences (9) in place of the single correlation sequences, is a very natural and important open problem;[21] a satisfactory solution is going to give us new insights and significantly improve our ability to deal with multiple ergodic averages. There are indications that sequences of polynomial nature should replace the "linear" sequences $(e^{2\pi i n t})_{n \in \mathbb{N}}$. The most reasonable candidates at this point seem to be some collection of multivariable nilsequences.[22] For instance, examples of 2-step nilsequences

[21] I am not aware of a place in the literature where this problem is stated explicitly, but it has definitely been in the mind of experts for several years.

[22] If $X = G/\Gamma$ is a k-step nilmanifold, $b_1, \ldots, b_\ell \in G$, $x \in X$, and F is Riemann integrable on X, we call the sequence $(F(b_1^{n_1} \cdots b_\ell^{n_\ell} x))$ a *basic k-step nilsequence with ℓ-variables*. A *k-step nilsequence with ℓ-variables*, is a uniform limit of basic k-step nilsequences with ℓ-variables. As is easily verified, the collection of k-step nilsequences, with the topology of uniform convergence, forms a closed algebra.

in 2 variables are the sequences $\big(e^{i([m\alpha]n\beta+m\gamma+n\delta)}\big)$ ($[x]$ denotes the integer part of x), where $\alpha,\beta,\gamma,\delta\in\mathbb{R}$.

Problem 1. *Determine the structure of the multiple correlation sequences $\mathcal{C}(n_1,\ldots,n_\ell)$ defined by (9). Is it true that the building blocks are multi-variable nilsequences? More explicitly, is it true that there exists a complex Borel measure σ on a compact metric space X, of bounded variation, and multi-variable nilsequences $\mathcal{N}_x(n_1,n_2,\ldots,n_\ell)$, $x\in X$, such that $x\mapsto N_x$ is integrable with respect to σ and*

$$\mathcal{C}(n_1,n_2,\ldots,n_\ell)=\int_X \mathcal{N}_x(n_1,n_2,\ldots,n_\ell)\ d\sigma(x)?$$

Some commutativity assumption on the transformations is needed, otherwise simple examples show that nilsequences cannot be the only building blocks[23]. If the answer turns out to be positive, then several well known problems in the area will be solved immediately using some known equidistribution results on nilmanifolds; for instance, mean convergence for the averages $\frac{1}{N}\sum_{n=1}^N T_1^{p_1(n)}f_1\cdot\ldots\cdot T_\ell^{p_\ell(n)}f_\ell$ where $p_1,\ldots,p_\ell\in\mathbb{Z}[t]$, and mean convergence for the averages $\frac{1}{N}\sum_{n=1}^N T_1^{a_n(\omega)}f_1\cdot\ldots\cdot T_\ell^{a_n(\omega)}f_\ell$ where $(a_n(\omega))$ is a random non-lacunary sequence (both problems are stated carefully in subsequent sections).

Even resolving special cases of this problem would be extremely interesting. One particular instance is the following:

Special Case of Problem 1. *Let $(X,\mathcal{X},\mu,T)$ be an invertible ergodic measure preserving system and $f,g,h\in L^\infty(\mu)$. Determine the structure of the multiple correlation sequences $(\mathcal{C}(n))$ defined by*

$$\mathcal{C}(n):=\int f\cdot T^n g\cdot T^{2n}h\ d\mu.$$

Is it true that the building blocks are 2-step nilsequences?

In [33] it is shown that for ergodic systems one has the decomposition

$$\mathcal{C}(n)=\mathcal{N}(n)+e(n)$$

where $\mathcal{N}(n)$ is a (single variable) two step nilsequence and $\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N|e(n)|=0$. Unfortunately, this result does not provide information on the error term $e(n)$ (other than it converges to 0 in density), and as a consequence it is of little use when one studies sparse subsequences of the sequence $\mathcal{C}(n)$.

4.2. **Necessary and sufficient conditions for ℓ-convergence.** The next result can be deduced from formula (10) and serves as our model for giving usable necessary and sufficient conditions for ℓ-convergence:

Theorem. *If $(a(n))$ is a sequence of integers, then the following statements are equivalent:*

- *The sequence $(a(n))$ is good for 1-convergence.*
- *The sequence $(a(n))$ is good for 1-convergence for rotations on the circle.*
- *The sequence $\big(\frac{1}{N}\sum_{n=1}^N e^{ia(n)t}\big)$ converges for every $t\in\mathbb{R}$.*

[23] Let $T,S\colon\mathbb{T}\to\mathbb{T}$ be given by $Tx:=2x$, $Sx:=2x+\alpha$, and $f(x):=e^{-ix}$, $g(x):=e^{ix}$. Then $\int T^nf\cdot S^ng\ dx=e^{i\cdot(2^n-1)\alpha}$ and one can show that $(e^{i\cdot 2^n\alpha})$ is not a nilsequence for $\alpha\in\mathbb{R}\setminus\mathbb{Q}$.

Since a formula that generalizes (10) to multiple correlation sequences is not available, we are unable to prove analogous necessary and sufficient conditions for ℓ-convergence. Nevertheless, inspired by Problem 1 we make the following natural guess:

Problem 2. *If $(a_1(n)),\dots,(a_\ell(n))$ are sequences of integers show that the following statements are equivalent:*

- *The sequences $(a_1(n)),\dots,(a_\ell(n))$ are good for ℓ-convergence of commuting transformations.*
- *The sequences $(a_1(n)),\dots,(a_\ell(n))$ are good for ℓ-convergence of ℓ-step nilsystems.*
- *The sequence $\big(\frac{1}{N}\sum_{n=1}^N \mathcal{N}(a_1(n),\dots,a_\ell(n))\big)$ converges for every ℓ-variable ℓ-step nilsequence $\mathcal{N}$.*

A similar problem was formulated in [82]. Although the equivalence of the second and third condition does not appear to be obvious, we are mainly interested in knowing if the third (or second) conditions implies the first. Even the following very special case is open:

Special Case of Problem 2. *Let $(a(n))$ be a sequence of integers such that the averages $\frac{1}{N}\sum_{n=1}^N \mathcal{N}(a(n))$ converge for every 2-step nilsequence $\mathcal{N}$. Show that for every invertible measure preserving system $(X,\mathcal{X},\mu,T)$ and functions $f,g\in L^\infty(\mu)$, the averages*

$$\frac{1}{N}\sum_{n=1}^N T^{a(n)}f\cdot T^{2a(n)}g$$

converge in the mean.

This is known when $(a(n))$ is strictly increasing and its range has positive density; it follows from Theorem 1.9 in [33].

4.3. **Sufficient conditions for ℓ-recurrence.** Our model result is the next theorem of T. Kamae and M. Mendès-France [127] that gives usable conditions for checking that a sequence is good for 1-recurrence:

Theorem. *Let $(a(n))$ be sequence of integers that satisfies*

- *the sequence $(a(n)\alpha)_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$ for every irrational α, and*
- *the set $\{n\in\mathbb{N}\colon r|a(n)\}$ has positive upper density for every $r\in\mathbb{N}$.*

Then the sequence $(a(n))$ is good for 1-recurrence.

Once again, the proof of this result depends upon knowing identity (10). Since appropriate generalizations are not known for multiple correlation sequences, we are unable to give a similar criterion for ℓ-recurrence when $\ell\geq 2$. To state a conjectural criterion we extend the notion of an irrational rotation on the circle to general connected nilmanifolds: Given a connected nilmanifold $X=G/\Gamma$, an *irrational nilrotation in X* is an element $b\in G$ such that the sequence $(b^n\Gamma)_{n\in\mathbb{N}}$ is equidistributed on X.

Problem 3. *Let $(a(n))$ be a sequence that satisfies:*

- *for every connected ℓ-step nilmanifold X and every irrational nilrotation b in X the sequence $(b^{a(n)}\Gamma)_{n\in\mathbb{N}}$ is equidistributed in X, and*
- *the set $\{n\in\mathbb{N}\colon r|a(n)\}$ has positive upper density for every $r\in\mathbb{N}$.*

Show that the sequence $(a(n))$ is good for ℓ-recurrence of powers.

This problem was first formulated in [87] and in the same article a positive answer was given for sequences with range a set of integers with positive density. The stated conditions are satisfied by any integer polynomial sequence with zero constant term (follows from results in [140]), the sequence $([n^c])$ for every $c > 0$ (it follows from results in [78]), and the sequence of shifted primes $(p_n - 1)$ (it follows from results in [107]).

4.4. **Powers of sequences and recurrence.** It is known that if a sequence is good for ℓ-convergence of powers, then its first ℓ powers are good for 1-convergence. More precisely, the following holds (this is implicit in [92] Section 9.1, and is proved in detail in [86]):

Theorem. *If $(a(n))$ is good for ℓ-convergence of powers, then $(a(n)^k)$ is good for 1-convergence for $k = 1, \dots, \ell$.*

It is unclear whether a similar property holds for recurrence.

Problem 4. *If $(a(n))$ is good for ℓ-recurrence of powers, is then $(a(n)^k)$ good for 1-recurrence for $k = 1, \dots, \ell$?*

This problem was first stated in [86] and it is open even when $\ell = 2$. It is known that if $(a(n))$ is good for 2-recurrence of powers, then the sequence $(a(n)^2)$ is good for Bohr recurrence, meaning it is good for 1-recurrence for all rotations on tori (see [92] Section 9.1, or [86]). A well known question of Y. Katznelson asks whether a set of Bohr recurrence is necessarily a set of topological recurrence (for background on this question see [49, 129, 188]). Although there exist examples of sets of topological recurrence that are not sets of 1-recurrence [134], all known examples are rather complicated. All these lead one to believe that a possible example showing that the answer to Problem 4 is negative should be complicated.

4.5. **Commuting vs powers of a single transformation.** If a sequence is good for 2-convergence of commuting transformations, then, of course, it is also good for 2-convergence of powers. Interestingly, no example that distinguishes the two notions is known and in fact there may be none:

Problem 5. *Is there a sequence that is good for 2-convergence of powers but it is not good for 2-convergence of commuting transformations?*

The corresponding question for recurrence is also open:

Problem 6. *Is there a sequence that is good for 2-recurrence of powers but it is not good for 2-recurrence of commuting transformations?*

This question was first stated in [24] (Question 8). In [24], V. Bergelson states that the answer is very likely yes.

4.6. **Fast growing sequences.** Despite the successes in dealing with multiple recurrence and convergence problems of sequences that do not grow faster than polynomials, when it comes down to fast growing sequences our knowledge is very limited. Say that a sequence $(a(n))$ of positive integers is *fast growing* if $\lim_{n\to\infty} \log(a(n))/\log n = \infty$ (equivalently, if it is of the form $(n^{b(n)})$ with $b(n) \to \infty$).

Problem 7. *Give an explicit example of a fast growing sequence that is good for multiple recurrence and convergence of powers.*

Even for $\ell = 2$ no such example is known. Several explicit natural examples of fast growing sequences that do not grow exponentially fast should work, for instance, the sequences $([n^{(\log n)^a}])$, $(n^{[(\log n)^b]})$, $([e^{n^c}])$ for $a, b > 0$ and $c \in (0,1)$. Unfortunately though, it is very hard to work with these sequences, even for issues related to 1-recurrence and 1-convergence. Probably a sequence like $(n^{[\log\log n]})$ is easier to work with. One could also try to construct a (not so explicit) example using random sequences (more on this in another section).

5. Problems related to polynomial sequences

In this section we give a list of problems related to the study of multiple ergodic averages involving iterates given by polynomial sequences, and related applications to multiple recurrence.

5.1. Powers of a single transformation.

Let $\mathcal{P} := \{p_1, \ldots, p_\ell\}$ be a family of integer polynomials that are *essentially distinct*, meaning, all polynomials and their differences are nonconstant. First, we consider averages of the form

$$\frac{1}{N}\sum_{n=1}^{N} T^{p_1(n)} f_1 \cdot \ldots \cdot T^{p_\ell(n)} f_\ell, \tag{11}$$

where $(X, \mathcal{X}, \mu, T)$ is an invertible measure preserving system and $f_1, \ldots, f_\ell \in L^\infty(\mu)$. We remark that all the mean convergence results stated in this section work equally well for averages of the form $\frac{1}{\Phi_N}\sum_{n\in\Phi_N}$ in place of the averages $\frac{1}{N}\sum_{n\in N}$ where $(\Phi_N)_{N\in\mathbb{N}}$ is any Følner sequence of subsets of $\mathbb{N}$.

Before discussing some problems related to the characteristic factors of the averages (11), we state a result of B. Host and B. Kra [116] and A. Leibman [142] that gives useful information about their structure.

Theorem. *There exists $d = d(\mathcal{P})$ such that the factor $\mathcal{Z}_{d,T}$ is characteristic for mean convergence of the averages* (11)*.*

We emphasize that the value of $d(\mathcal{P})$ in the previous statement does not depend on the system or the functions involved. Given a family of polynomials $\mathcal{P}$, we denote by $d_{min}(\mathcal{P})$ the minimal value of $d(\mathcal{P})$ that works in the previous theorem. This value is in general hard to pin down and depends on the algebraic relations that the polynomials satisfy. For instance, we know that $d_{min}(\{n, 2n, \ldots, \ell n\}) = \ell - 1$ ([115, 196]), and $d_{min}(\mathcal{P}) = 1$ when $\mathcal{P}$ consists of at least two rationally independent polynomials ([83, 85]). But it is not only linear relations between the polynomials that matter, for instance, we know that $d_{min}(\{n, 2n, n^2\}) = 2$ while $d_{min}(\{n, 2n, n^3\}) = 1$ ([77, 146]). More examples of families $\mathcal{P}$ where $d_{min}(\mathcal{P})$ has been computed can be found in [77, 146]. Furthermore, in [146] a (rather complicated) algorithm is given for computing this value. Despite such progress, the following is still open (the problem is implicit in [39] and was stated explicitly in [146]):

Problem 8. *If $|\mathcal{P}| \geq 2$, show that $d_{min}(\mathcal{P}) \leq |\mathcal{P}| - 1$.*

The estimate is known when $|\mathcal{P}| = 2, 3$ ([77]) and it is open for $|\mathcal{P}| = 4$. The problem is even open when one is restricted to the class of Weyl systems, meaning, systems of the form $(\mathbb{T}^d, \mathcal{B}_{\mathbb{T}^d}, m_{\mathbb{T}^d}, T)$ where $T\colon \mathbb{T}^d \to \mathbb{T}^d$ is a unipotent affine transformation. We denote with $d^{\mathrm{W}}_{min}(\mathcal{P})$ the minimum value of $d(\mathcal{P})$ such that the factor $\mathcal{Z}_{d(\mathcal{P}),T}$ is characteristic for mean convergence of the averages (11) for all Weyl systems (properties of $d^{\mathrm{W}}_{min}(\mathcal{P})$ were studied in [39]).

Special Case of Problem 8. *If* $|\mathcal{P}| \geq 2$, *show that* $d^{W}_{min}(\mathcal{P}) \leq |\mathcal{P}| - 1$.

This problem was first stated in [39] (set $W(P) := d^{W}_{min}(\mathcal{P}) + 1$ in the remark after Proposition 5.3). The estimate is known when $|\mathcal{P}| = 2, 3, 4$ ([77, 158]) and it is open when $|\mathcal{P}| = 5$.

Interestingly, no example is known where $d_{min}(\mathcal{P}) \neq d^{\mathrm{W}}_{min}(\mathcal{P})$, so it is natural to suspect that these two values are always equal.

Problem 9. *Show that* $d_{min}(\mathcal{P}) = d^{W}_{min}(\mathcal{P})$.

This problem was first stated in [39]. The identity is known when $|P| = 3$ ([77]) and is open when $|P| = 4$. Obviously one has $d^{\mathrm{W}}_{min}(\mathcal{P}) \leq d_{min}(\mathcal{P})$. Some bounds in the other direction are given in [146].

Mean convergence of the averages (11) was established after a long series of intermediate results; the papers [91, 63, 64, 65, 101, 173, 112, 115, 196] dealt with the important case of linear polynomials, and using the machinery introduced in [115], convergence for arbitrary polynomials was finally obtained by B. Host and B. Kra in [116] except for a few cases that were treated by A. Leibman in [142].

Theorem. *Let* $(X, \mathcal{X}, \mu, T)$ *be an invertible measure preserving system,* $f_1, \dots, f_\ell \in L^\infty(\mu)$ *be functions, and* $p_1, \dots, p_\ell$ *be integer polynomials. Then the averages* (11) *converge in the mean as* $N \to \infty$.

Furthermore, explicit formulas for the limit can be given for special families of polynomials [194, 83, 85, 77, 146], but no such formula is known for general families of polynomials.

In most cases, it is still unknown whether mean convergence can be boosted to pointwise convergence. We mention two particular cases that are open:

Problem 10. *Show that the averages*

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^n x) \cdot f_2(T^{2n} x) \cdot f_3(T^{3n} x), \quad \text{or the averages} \quad \frac{1}{N}\sum_{n=1}^{N} f_1(T^n x) \cdot f_2(T^{n^2} x),$$

converge pointwise.

Pointwise convergence of the averages (11) is known when $\ell = 1$ [52] and is also known when $\ell = 2$ and both polynomials are linear [55] (see also [68] for an alternate proof). In all other cases the problem is open even for weak mixing systems. Partial results that deal with special classes of transformations can be found in [3, 5, 15, 16, 70, 140, 154, 155].

5.2. **Commuting transformations.** Throughout this section $(X, \mathcal{X}, \mu)$ is a probability space, $T_1, \dots, T_\ell \colon X \to X$ are commuting, invertible measure preserving transformations, $f_1, \dots, f_\ell \in L^\infty(\mu)$ are functions, and $p_1, \dots, p_\ell$ are polynomials with integer coefficients.

We start with the following result of V. Bergelson, B. Host, and B. Kra [33]:

Theorem. *For ergodic systems one has the decomposition*

$$\int f_0 \cdot T^n f_1 \cdot \ldots \cdot T^{\ell n} f_\ell \; d\mu = N(n) + e(n)$$

where $(N(n))$ *is an* ℓ*-step nilsequence and* $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} |e(n)| = 0$.

A more general result that uses polynomial iterates in place of the linear iterates was proved recently in [147]. A key ingredient in the proof of the previous theorem is the fact that the factor $\mathcal{Z}_{\ell,T}$ is characteristic for convergence of the averages $\frac{1}{N}\sum_{n=1}^N |\int f_0\cdot T^nf_1\cdot\ldots\cdot T^{\ell n}f_\ell\, d\mu|$. When one uses commuting transformations in place of powers of the same transformation an analogous property fails, nevertheless, there are no known examples of multicorrelation sequences of commuting transformations that are genuinely different than nilsequences.

Problem 11. *Is it true that one always has the decomposition*

$$\int f_0\cdot T_1^nf_1\cdot\ldots\cdot T_\ell^nf_\ell\, d\mu = N(n)+e(n)$$

where $(N(n))$ *is an* ℓ*-step nilsequence and* $\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N|e(n)|=0$*?*

The question is open even when $\ell=2$. Notice that we make no ergodicity assumptions, so in particular this problem is open even when $\ell=2$ and $T_2=T_1^2$.

We move to some problems related to convergence properties of multiple ergodic averages. A very natural problem (stated explicitly in [24] but was advertised long before 1996 by H. Furstenberg and others) is to extend the mean convergence result involving polynomial iterates of a single transformation to several commuting transformations:

Problem 12. *Show that the averages*

$$\frac{1}{N}\sum_{n=1}^N T_1^{p_1(n)}f_1\cdot\ldots\cdot T_\ell^{p_\ell(n)}f_\ell \tag{12}$$

converge in the mean as $N\to\infty$*.*

Mean convergence is known when the transformations $T_1,\ldots,T_\ell$ are powers of the same transformation ([116, 142]), when the polynomials are linear [181] (with alternate proofs given in [186, 9, 111]), and when the polynomials have distinct degrees [62]. Convergence is also known for general families of polynomials if one imposes very strong ergodicity assumptions on the transformations [124]. See also [11, 12] where techniques from [9] have been refined and extended, aiming to eventually handle the case of general polynomial iterates. Despite such intense efforts, convergence is still not known for some simple families of polynomials, for instance, when $\ell=2$ and $p_1(n)=p_2(n)=n^2$, or when $p_1(n)=n^2$, $p_2(n)=n^2+n$.

As mentioned previously, when all transformations are equal, and the polynomials are essentially distinct, then characteristic factors of the averages (12) can be chosen to have very special algebraic structure. For general commuting transformations this is no longer the case; if one chooses $p_2=p_1=n$, $T_1=T_2$, and $f_2=\bar{f}_1$, then the averages (12) do not converge to 0 unless $f_1=f_2=0$. The same problem persists when two of the polynomials are *pairwise dependent*, meaning, some non-trivial linear combination of two of the polynomials is constant. But in all other cases, there is no obvious obstruction to having "simple" characteristic factors.

Problem 13. *Suppose that the polynomials* $p_1,\ldots,p_\ell\in\mathbb{Z}[t]$ *are* **pairwise** *independent. Show that there exists* $d\in\mathbb{N}$ *such that the factors* $\mathcal{Z}_{d,T_1},\ldots,\mathcal{Z}_{d,T_\ell}$ *are characteristic factors for the averages* (12)*.*

This is known to be the case when the polynomials have distinct degrees [62]. But it is not known for some simple families of integer polynomials, for instance, for the family $\{n^3,n^3+n\}$ or the family $\{n,n^2,n^2+n\}$. Even for weak mixing transformations the problem is open:

Special Case of Problem 13. *Suppose that the transformations $T_1,\ldots,T_\ell\colon X\to X$ are weak mixing and the polynomials $p_1,\ldots,p_\ell\in\mathbb{Z}[t]$ are* **pairwise** *independent. Show that*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}T_1^{p_1(n)}f_1\cdot\ldots\cdot T_\ell^{p_\ell(n)}f_\ell=\int f_1\;d\mu\cdot\ldots\cdot\int f_\ell\;d\mu.$$

When all transformations are equal and the polynomials are in general position, characteristic factors for the averages (12) turn out to be extremely simple [83, 85]:

Theorem. *Suppose that the polynomials $p_1,\ldots,p_\ell\in\mathbb{Z}[t]$ are rationally independent. Then the rational Kronecker factor $\mathcal{K}_{rat}(T)$[24] is a characteristic factor for the averages* (11).

It is very likely that this result generalizes to the case of several commuting transformations:

Problem 14. *Suppose that the polynomials $p_1,\ldots,p_\ell\in\mathbb{Z}[t]$ are rationally independent. Show that the factors $\mathcal{K}_{rat}(T_1),\ldots,\mathcal{K}_{rat}(T_\ell)$ are characteristic factors for the averages* (12).

This was proved in [62] when $\ell=2$ and $p_1(n)=n$. In the same article a somewhat weaker property was proved for all monomials with distinct degrees. We mention also a closely related multiple recurrence problem:

Problem 15. *Suppose that the polynomials $p_1,\ldots,p_\ell\in\mathbb{Z}[t]$ are rationally independent and have zero constant term. Show that for every $A\in\mathcal{X}$ and every $\varepsilon>0$, there exists $n\in\mathbb{N}$ such that*

$$\mu(A\cap T_1^{-p_1(n)}A\cap\cdots\cap T_\ell^{-p_\ell(n)}A)\geq\mu(A)^{\ell+1}-\varepsilon. \tag{13}$$

In fact, the set of integers n for which (13) holds is expected to have bounded gaps. The lower bounds are known when all transformations are equal [84] and they are also known for general commuting transformations when the polynomials are monomials with distinct degrees [62]. The result fails if the polynomials are distinct and pairwise dependent; in this case no fixed power of $\mu(A)$ works as a lower bound in (13) for every system and set [33]. On the other hand, the assumption that the polynomials are rationally independent is not necessary, for instance, the result is expected to work for the family of polynomials $\{n,n^2,n^2+n\}$ (this is known to be the case when all transformations are equal [77]). We remark that Problem 14 is solved, then the conjectured lower bounds of Problem 15 will follow rather easily.

Regarding pointwise convergence of multiple ergodic averages of commuting transformations, progress has been extremely scarce. Even when one uses two commuting transformations and linear iterates convergence is not known in general. The following is a well known open problem:

Problem 16. *Let $(X,\mathcal{X},\mu)$ be a probability space, $T,S\colon X\to X$ be commuting invertible measure preserving transformations, and $f,g\in L^\infty(\mu)$ be functions. Show that the averages*

$$\frac{1}{N}\sum_{n=1}^{N}f(T^nx)\cdot g(S^nx)$$

converge pointwise.

For a list of partial results that apply to special classes of transformations see the list after Problem 10.

[24]Given a measure preserving system $(X,\mathcal{X},\mu,T)$ we define $\mathcal{K}_{rat}(T)=\bigvee_{d\in\mathbb{N}}\mathcal{I}_{T^d}$.

5.3. **Not necessarily commuting transformations.** All problems in the previous sections were stated for families of transformations that commute. It is very likely that all positive results extend to the case where the transformations generate a nilpotent group. For instance, we mention a problem from [36]:

Problem 17. *Let $(X, \mathcal{X}, \mu)$ be a probability space, $T_1, \ldots, T_\ell \colon X \to X$ be invertible measure preserving transformations that generate a nilpotent group, and $f_1, \ldots, f_\ell \in L^\infty(\mu)$ be functions. Show that the averages*

$$\frac{1}{N}\sum_{n=1}^{N} T_1^n f_1 \cdot \ldots \cdot T_\ell^n f_\ell$$

converge in $L^2(\mu)$.

Convergence is known when $\ell = 2$ [36] and is open for $\ell = 3$. The interested reader should look in [36] for a list of other closely related open problems. See also [137] for a related multiple recurrence result.

When one works with arbitrary families of invertible measure preserving transformations the next result shows that one cannot expect to have similar convergence results:

Theorem. *Let $a, b \colon \mathbb{N} \to \mathbb{Z} \setminus \{0\}$ be $1-1$ sequences. Then there exist invertible Bernoulli measure preserving transformations T and S acting on the same probability space $(X, \mathcal{X}, \mu)$ such that*

- *for some $f, g \in L^\infty(\mu)$ the averages $\frac{1}{N}\sum_{n=1}^{N}\int T^{a(n)} f \cdot S^{b(n)} g \, d\mu$ diverge;*
- *for some $A \in \mathcal{X}$ with $\mu(A) > 0$ we have $T^{-a(n)}A \cap S^{-b(n)}A = \emptyset$ for every $n \in \mathbb{N}$.*

To construct such examples it suffices to modify examples of D. Berend (see Ex 7.1 in [15]) and H. Furstenberg (page 40 in [92]) that cover the case $a(n) = b(n) = n$ (the details will appear in [88]). When $a(n) = b(n)$, it is also known that given any finitely generated solvable group G of exponential growth, there exist invertible measure preserving transformations T, S, with $< T, S > \subset G$, and such that and the conclusion of the previous theorem holds for those T and S. It is interesting that despite such negative news, once one introduces an extra variable, several convergence (and very likely recurrence) results can be extended to arbitrary families of measure preserving transformations. We mention an example from [61]:

Theorem. *Let $(X, \mathcal{X}, \mu)$ be a probability space, $T_1, \ldots, T_\ell \colon X \to X$ be invertible measure preserving transformations, $f_1, \ldots, f_\ell \in L^\infty(\mu)$ be functions, $p_1, \ldots, p_\ell$ be essentially distinct polynomials, and $a \in (0, 1/d)$. Then the averages*

$$\frac{1}{N^{1+a}} \sum_{1 \leq m \leq N, 1 \leq n \leq N^a} f_1(T_1^{m+p_1(n)}x) \cdot \ldots \cdot f_\ell(T_\ell^{m+p_\ell(n)}x) \tag{14}$$

converge pointwise as $N \to \infty$.

One can show that the assumption that the polynomials are essentially distinct is necessary. It was also shown in [61] that there exists $d \in \mathbb{N}$ such that the factors $\mathcal{Z}_{d,T_1}, \ldots, \mathcal{Z}_{d,T_\ell}$ are characteristic for pointwise convergence of the averages (14). Interestingly, the corresponding multiple recurrence result (that would generalize the polynomial Szemerédi theorem) remains open:

Problem 18. *Let $(X, \mathcal{X}, \mu)$ be a probability space, $T_1, \ldots, T_\ell \colon X \to X$ be invertible measure preserving transformations, and $p_1, \ldots, p_\ell$ be distinct polynomials with zero constant term. Show that for every $A \in \mathcal{X}$ with $\mu(A) > 0$ we have*

$$\mu(A \cap T_1^{-m-p_1(n)} A \cap \cdots \cap T_\ell^{-m-p_\ell(n)} A) > 0$$

for some $m, n \in \mathbb{N}$.[25]

The assumption that the polynomials are distinct is necessary since as mentioned before there exist (non-commuting) transformations T, S, acting on the same probability space $(X, \mathcal{X}, \mu)$, and a set $A \in \mathcal{X}$ with $\mu(A) > 0$ such that $\mu(T^n A \cap S^n A) = 0$ for every $n \in \mathbb{N}$. The multiple recurrence property is known to hold when all the transformations are weak mixing [61], but for general measure preserving systems even some of the simplest cases are open:

Special Case of Problem 18. *Let $(X, \mathcal{X}, \mu)$ be a probability space and $T, S, R \colon X \to X$ be invertible measure preserving transformations. Show that for every $A \in \mathcal{X}$ with $\mu(A) > 0$ there exist $m, n \in \mathbb{N}$ such that*

$$\mu(A \cap T^{-m} A \cap S^{-m-n} A \cap R^{-m-2n} A) > 0.$$

6. Problems related to sequences arising from smooth functions

In this section we give a list of problems related to the study of multiple ergodic averages involving iterates given by sequences arising from smooth functions, and related applications to multiple recurrence and number theory.

We are going to restrict ourselves, almost entirely, to a class of non-oscillatory functions that is rich enough to contain several interesting examples. Its formal definition is the following: Let B be the collection of equivalence classes of real valued functions defined on some half-line (c, ∞), where we identify two functions if they agree eventually. A *Hardy field* is a subfield of the ring $(B, +, \cdot)$ that is closed under differentiation. With $\mathcal{H}$ we denote the *union of all Hardy fields*. It is easy to check that if a function belongs in $\mathcal{H}$, then it is eventually monotonic and the same holds for its derivatives, so if $a \in \mathcal{H}$, all limits $\lim_{t\to\infty} a^{(k)}(t)$ exist (they may be infinite). We call a *Hardy sequence* any sequence of the form $([a(n)])$ where $a \in \mathcal{H}$.

An explicit example of a Hardy field to keep in mind is the set $\mathcal{LE}$ that consists of all *logarithmico-exponential functions* (introduced by Hardy in [109]), meaning all functions defined on some half-line (c, ∞) using a finite combination of the symbols $+, -, \times, :, \log, \exp$, operating on the real variable t and on real constants. For example, all rational functions and the functions $t^{\sqrt{2}}$, $t \log t$, $t^{\sqrt{\log\log t}}/\log(t^2+1)$ belong in $\mathcal{LE}$. Let us stress though that the set $\mathcal{H}$ is much more extensive than the set $\mathcal{LE}$; it contains all antiderivatives of elements of $\mathcal{LE}$, the Riemann zeta function ζ, the Euler Gamma function Γ, etc.

The main advantage we get by working with elements of $\mathcal{H}$ is that it is possible to relate their growth rates with the growth rates of their derivatives.[26] As a consequence, a single growth

[25] This would imply that given a countable amenable group G and arbitrary elements $a_1, \ldots, a_\ell \in G$, for every $E \subset G$ that has positive upper density with respect to some Følner sequence in G, there exist $g \in M$ and $m, n \in \mathbb{N}$ such that g, $a_1^{m+p_1(n)} g, \ldots, a_\ell^{m+p_\ell(n)} g \in E$.

[26] If $a \in \mathcal{H}$ and $b \in \mathcal{LE}$, then there exists a Hardy field that contains both a and b. As a consequence, the limit $\lim_{t\to\infty} a'(t)/b'(t)$ exists (it may be infinite), and so assuming that $a(t), b(t) \to \infty$, we get (using L'Hospital's rule) that the quotients $a(t)/b(t)$ and $a'(t)/b'(t)$ have the same limit as $t \to \infty$. We deduce, for instance, that if $a \in \mathcal{H}$ satisfies $a(t)/t^2 \to \infty$, then $a'(t)/t \to \infty$ and $a''(t) \to \infty$.

condition encodes a lot of useful information and this enables us to give more transparent and appetizing statements.

Background material on Hardy fields can be found in [46, 47, 48, 109, 110, 171].

6.1. **Powers of a single transformation.**

6.1.1. *Hardy sequences of polynomial growth.* To avoid repetition we remark that in this subsection we always work with a family $\mathcal{F} := \{a_1(t), \dots, a_\ell(t)\}$ of functions of *polynomial growth* (meaning $a_i(t)/t^k \to 0$ for some $k \in \mathbb{N}$) that belong to the same Hardy field. With $\text{span}^*(\mathcal{F})$ we denote the set of all *non-trivial* linear combinations of elements of $\mathcal{F}$.

We first state two problems from [79] related to the mean convergence of multiple ergodic averages involving iterates given by Hardy sequences. The following result was proved in [79] (the case $\ell = 1$ was first handled in [50]):

Theorem. *Let $a \in \mathcal{H}$ have polynomial growth. Then the sequence $([a(n)])$ is good for multiple convergence of powers if and only if one of the following conditions is satisfied:*

- $|a(t) - cp(t)|/\log t \to \infty$ *for every* $c \in \mathbb{R}$ *and* $p \in \mathbb{Z}[t]$; *or*
- $a(t) - cp(t) \to d$ *for some* $c, d \in \mathbb{R}$; *or*
- $|a(t) - t/m| \ll \log t$ *for some* $m \in \mathbb{Z}$.

For instance, the sequences (n^2), $([n^{3/2}])$, $([n \log n])$, $([n^2 + (\log n)^2])$, $([n^2 + n\sqrt{2} + \log\log n])$ are all good for multiple convergence of powers, but the sequences $([n^2 + \log n])$, $([n^2\sqrt{2} + \log\log n])$ are not good for 1-convergence. Unlike the case of polynomial sequences, if $a \in \mathcal{H}$ satisfies $a(t)/t^{k-1} \to \infty, a(t)/t^k \to 0$ for some $k \in \mathbb{N}$, then the sequence $([a(n)])$ takes odd (respectively even) values in arbitrarily large intervals. As a consequence, when T is the rotation by $1/2$ on the circle and $f := \mathbf{1}_{[0,1/2]}$, the $L^2(\mu)$-limit $\lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} T^{[a(n)]} f$ does not exist for some appropriately chosen Følner sequence $(\Phi_N)_{N\in\mathbb{N}}$ of subsets of $\mathbb{N}$.

The next problem seeks to give similar necessary and sufficient conditions for ℓ-convergence of arbitrary collections of sequences arising from functions of polynomial growth that belong to the same Hardy field. We remind the reader that in such circumstances one is seeking to prove mean convergence for averages of the form

$$\frac{1}{N}\sum_{n=1}^{N} T^{[a_1(n)]} f_1 \cdot \ldots \cdot T^{[a_\ell(n)]} f_\ell. \tag{15}$$

Problem 19. *Let $\mathcal{F}$ be as above. Show that the collection of sequences $\{([a_1(n)]), \dots, ([a_\ell(n)])\}$ is good for ℓ-convergence of a single transformation if and only if every function $a \in \text{span}^*(\mathcal{F})$ satisfies one of the following conditions:*

- $|a(t) - cp(t)|/\log t \to \infty$ *for every* $c \in \mathbb{R}$ *and* $p \in \mathbb{Z}[t]$; *or*
- $a(t) - cp(t) \to d$ *for some* $c, d \in \mathbb{R}$; *or*
- $|a(t) - t/m| \ll \log t$ *for some* $m \in \mathbb{Z}$.

Convergence was proved in [79] under much more restrictive conditions than those advertised here. The collection of sequences $\{([n \log n]), ([n^2 \log n]), \dots, ([n^\ell \log n])\}$ is an explicit example that is expected to be good for ℓ-convergence of a single transformation but this is not known yet (not even for all weak mixing systems, or all nilsystems).

When the multiple ergodic averages of a collection of Hardy sequences of polynomial growth converge in the mean one would like to have an explicit formula for their limit. In general, such a limit formula can be extremely complicated but when the sequences are in "general position" the limit is expected to be very simple:

Problem 20. *Let $\mathcal{F}$ be as above and suppose that for every function $a \in span^*(\mathcal{F})$ we have $|a(t) - cp(t)|/\log t \to \infty$ for every $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$. Show that for every invertible measure preserving system $(X, \mathcal{B}, \mu, T)$ and functions $f_1, \dots, f_\ell \in L^\infty(\mu)$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^{[a_1(n)]} f_1 \cdot \ldots \cdot T^{[a_\ell(n)]} f_\ell = \mathbb{E}(f_1|\mathcal{I}_T) \cdot \ldots \cdot \mathbb{E}(f_\ell|\mathcal{I}_T) \tag{16}$$

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

The identity is known when $a_i(t) = t^{c_i}$, $i = 1, \dots, \ell$, where $c_1, \dots, c_\ell \in \mathbb{R} \setminus \mathbb{Z}$ are different and positive [79] (this was established first in [31] when $c_i \in (0, 1)$, or when the system is weak mixing).[27] A particular collection of sequences for which the identity is not known is the one mentioned before: $\{([n \log n]), ([n^2 \log n]), \dots, ([n^\ell \log n])\}$. If some function $a \in \text{span}(\mathcal{F})$ satisfies $|a(t) - cp(t)| \ll \log t$ for some $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$ with $\deg(p) \geq 2$, then one easily sees that (16) fails for T given by an appropriate rotation on $\mathbb{T}^\ell$.

An intermediate step that would help solve the previous two problems is to find suitable characteristic factors for the relevant multiple ergodic averages. We state a problem from [79] of this sort that we find of independent interest:

Problem 21. *Let $\mathcal{F}$ be as above and suppose that $a_i(t)/\log t \to \infty$ and $(a_i(t)-a_j(t))/\log t \to \infty$ whenever $i \neq j$. Show that for every invertible measure preserving system $(X, \mathcal{X}, \mu, T)$ the factor $\mathcal{Z}_T := \bigvee_{d \in \mathbb{N}} \mathcal{Z}_{d,T}$ is characteristic for mean convergence of the averages* (15)*.*

This is known when for some $\varepsilon > 0$ we have $a_i(t)/t^\varepsilon \to \infty$ and $(a_i(t) - a_j(t))/t^\varepsilon \to \infty$ whenever $i \neq j$ [79], and the methods of [79] (see the proof of Theorem 2.4 there) can be used to show that it also holds when $a_i(t) = ia(t)$ for $i = 1, \dots, \ell$ and $a(t)/\log t \to \infty$. On the other hand, in the generality stated, the problem is open even for weak mixing systems:

Special Case of Problem 21. *Let $\mathcal{F}$ be as above and suppose that $a_i(t)/\log t \to \infty$ and $(a_i(t)-a_j(t))/\log t \to \infty$ whenever $i \neq j$. Show that for every weak mixing system the averages* (15) *converge in the mean to the product of the integrals of the individual functions.*

One can check that the stated assumptions are necessary.

Next we state some problems related to multiple recurrence. The following result was proved in [79] (see also [90] for a special case):

Theorem. *Let $a \in \mathcal{H}$ have polynomial growth and suppose that $|a(t) - cp(t)| \to \infty$ for every $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$. Then the sequence $([a(n)])$ is good for multiple recurrence of powers.*

It follows that the sequences $([n^{\sqrt{2}}])$, $([n \log n])$, $([n^2 + (\log n)^2])$, $([n^2 + \log n])$, $([n^2\sqrt{2} + \log \log n])$ $([n^2 + n\sqrt{2}])$ are all good for multiple recurrence of powers. The previous result does not handle the case $a(t) = cp(t) + d + e(t)$ where p is an integer polynomial with zero constant

[27] More generally, the identity was shown in [79] when the functions $a_1, \dots, a_\ell$ and their pairwise differences belong to the set $\mathcal{LE} \cap \{a\colon a(t)/t^{k+\varepsilon} \to \infty, a(t)/t^{k+1} \to 0, \text{ for some } k \geq 0 \text{ and } \varepsilon > 0\}$.

term, $e \in \mathcal{H}$ is non-negative and converges to zero, and $d \in \mathbb{R}$. If $d = 0$, then one can show that the sequence $([a(n)])$ is good for multiple recurrence of powers. The case where $d \neq 0$ is trickier. For instance, the sequence $([\sqrt{5n+1}])$ is good for multiple recurrence of powers but the sequence $([\sqrt{5n+2}])$ is not good for 1-recurrence for the rotation on the circle by $1/\sqrt{5}$.

Next we state a problem from [79] that seeks to give necessary conditions for ℓ-recurrence of arbitrary collections of sequences arising from functions of polynomial growth that belong to the same Hardy field. We remind the reader that in such circumstances one is seeking to prove that whenever $\mu(A) > 0$ we have $\mu(A \cap T^{-[a_1(n)]}A \cap \cdots \cap T^{-[a_\ell(n)]}A) > 0$ for some $n \in \mathbb{N}$.

Problem 22. *Let $\mathcal{F}$ be as above and suppose that for every function $a \in span^*(\mathcal{F})$ we have $|a(t) - cp(t)| \to \infty$ for every $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$. Show that the collection of sequences $\{([a_1(n)]), \ldots, ([a_\ell(n)])\}$ is good for ℓ-recurrence of a single transformation.*

As stated above, this is known for $\ell = 1$ [90].

Lastly, we mention an interesting multiple recurrence problem involving a collection of sequences not covered by the previous problem. Using consecutive values of positive fractional powers is expected to produce good multiple return times, a result that fails for integral powers:

Problem 23. *Let c be a positive real number that is not an integer. Show that for every $\ell \in \mathbb{N}$, the collection of sequences $\{([n^c]), ([(n+1)^c]), \ldots, ([(n+\ell)^c])\}$ is good for $(\ell+1)$-recurrence of a single transformation.*

If the integer ℓ is greater than $[c]$, then some non-trivial linear combination of the functions $t^c, (t+1)^c, \ldots, (t+\ell)^c$ converges to 0; so in this particular instance the assumptions of Problem 22 are not be satisfied. The conclusion fails trivially when $c = 1$, to see that it fails for $c = 2, 3, \ldots$ it suffices to consider appropriate rotations on the circle.

6.1.2. *Hardy sequences of super-polynomial growth.* Despite the fact that multiple recurrence and convergence properties of Hardy sequences of polynomial growth are relatively well understood, when it comes down to sequences that grow faster than polynomials, even the most basic problems are open.

Problem 24. *Find an example of a function $a \in \mathcal{H}$ that grows faster than polynomials, meaning $a(t)/t^k \to \infty$ for every $k \in \mathbb{N}$, such that the sequence $[a(n)]$ is good for multiple recurrence and convergence of powers.*

The sequences $([n^{(\log n)^a}])$, $([e^{n^b}])$, where $a > 0$ and $b \in (0,1)$, seem to be natural candidates; unfortunately they are extremely hard to work with. Even when $\ell = 1$ the relevant exponential sum estimates needed to prove convergence appear to be out of reach in most cases; for the first sequence such estimates are available only when $a \in (0, 1/2)$ [128], and no estimates are available for the second sequence. On the other hand, a slower growing sequence, like the sequence $([n^{\log\log n}])$ may be easier to handle. But even for this sequence, 2-recurrence and 2-convergence is not known for all weak mixing systems or all nilsystems.

6.1.3. *Hardy sequences evaluated at the primes.* With p_n we denote the n-th prime.

Problem 25. *Let $c \in \mathbb{R} \setminus \mathbb{Z}$ be a positive real number. Show that the sequence $([p_n^c])$ is good for multiple recurrence and convergence of powers.*

Proving multiple recurrence is trivial when $c<1$ since in this case the sequence $([p_n^c])$ misses at most finitely many positive integer values. It is known that if $c\in\mathbb{R}\setminus\mathbb{Z}$ is positive, then the sequence of fractional parts $(\{p_n^c\})$ is equidistributed in the unit interval (see [176] or [190] for $c<1$ and [149] for $c>1$). Probably the techniques used to prove these equidistribution results suffice to prove 1-recurrence and 1-convergence (it suffices to show that the sequence $(\{p_n^c\alpha\})$ is equidistributed in the unit interval for every non-zero $\alpha\in\mathbb{R}$), but the problem is open for ℓ-recurrence and ℓ-convergence when $\ell\geq 2$.

6.1.4. *Oscillatory sequences.* All the previous problems deal with sequences that do not oscillate. Multiple recurrence and convergence properties of oscillatory sequences are not well studied and analyzing some simple looking sequences leads to interesting problems:

Problem 26. *Show that the sequence* $([n\sin n])$ *is good for multiple convergence of powers.*

Quite likely one can say more; the averages $\frac{1}{N}\sum_{n=1}^N T^{a(n)}f_1\cdot T^{2a(n)}f_2\cdot\ldots\cdot T^{\ell a(n)}f_\ell$ have the same limiting value when $a(n)=[n\sin n]$ and $a(n)=n$. This is known for $\ell=1$, it follows from equidistribution results in [19] (see also related results in [20, 21]). As far as I know the problem has not been studied when $\ell\geq 2$, even for particular classes of measure preserving systems, like nilsystems or weakly mixing systems.

6.2. **Commuting transformations.** As mentioned before if a Hardy sequence has polynomial growth and stays away from constant multiples of integer polynomials, then it is going to be good for multiple recurrence of powers. The next problem seeks to extend this to the case of commuting transformations. We remind the reader that in such circumstances one is seeking to prove that whenever $\mu(A)>0$ we have $\mu(A\cap T_1^{-[a(n)]}A\cap\cdots\cap T_\ell^{-[a(n)]}A)>0$ for some $n\in\mathbb{N}$.

Problem 27. *Let* $a\in\mathcal{H}$ *have polynomial growth and suppose that* $|a(t)-cp(t)|\to\infty$ *for every* $c\in\mathbb{R}$ *and* $p\in\mathbb{Z}[t]$. *Then the sequence* $([a(n)])$ *is good for multiple recurrence of commuting transformations.*

The proof that any such $([a(n)])$ is good for multiple recurrence of powers relies crucially on the precise algebraic structure of suitable characteristic factors for the corresponding multiple ergodic averages; an advantage that is lost when one works with commuting transformations.

Problem 28. *Let* $(X,\mathcal{X},\mu)$ *be a probability space,* $T_1,\ldots,T_\ell\colon X\to X$ *be commuting invertible measure preserving transformations, and* $f_1,\ldots,f_\ell\in L^\infty(\mu)$. *Show that for every positive real number* c *the following limit exists in* $L^2(\mu)$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}T_1^{[n^c]}f_1\cdot\ldots\cdot T_\ell^{[n^c]}f_\ell, \tag{17}$$

and if c *is not an integer, then it is equal to* $\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N T_1^n f_1\cdot\ldots\cdot T_\ell^n f_\ell$.

For $c=1$ the existence of the limit (17) is known [181]. The case $0<c<1$ can be easily reduced to the case $c=1$. So the interesting case is when $c>1$ in which case the problem is open even when $\ell=2$ and all transformations are assumed to be weak mixing.

Problem 29. *Let $(X, \mathcal{X}, \mu)$ be a probability space, $T_1, \ldots, T_\ell \colon X \to X$ be commuting invertible measure preserving transformations, and $f_1, \ldots, f_\ell \in L^\infty(\mu)$. Let $c_1, \ldots, c_\ell \in \mathbb{R} \setminus \mathbb{Z}$ be positive and distinct. Show that*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T_1^{[n^{c_1}]} f_1 \cdot \ldots \cdot T_\ell^{[n^{c_\ell}]} f_\ell = \mathbb{E}(f_1|\mathcal{I}_{T_1}) \cdot \ldots \cdot \mathbb{E}(f_\ell|\mathcal{I}_{T_\ell}),$$

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

The identity is known when all the transformations are equal [79] and is also known when all the exponents c_i are smaller than 1 [79] (in which case the assumption that the transformations commute is not needed). The interesting case is when $\ell \geq 2$ and all the exponents are greater than 1. Easy examples show that the limit formula fails if one of the powers is an integer different than 1.

6.3. **Configurations in the primes.** As we mentioned in the introduction, the theorem of Szemerédi on arithmetic progressions [177], and its polynomial extension [35], have been instrumental in proving that the primes contain arbitrarily long arithmetic progressions [104] and polynomial progressions [183]. It is then natural to expect that the various available Hardy field extensions of the theorem of Szemerédi [79, 90] can be used to prove that the primes contain the corresponding Hardy field patterns. For instance:

Problem 30. *Let $\ell \in \mathbb{N}$ and $c, c_1, \ldots, c_\ell$ be positive real numbers. Show that the prime numbers contain patterns of the form*

$$\{m, m + [n^c], m + 2[n^c], \ldots, m + \ell[n^c]\} \quad \text{and} \quad \{m, m + [n^{c_1}], \ldots, m + [n^{c_\ell}]\}.$$

When all exponents are rational the existence of such patterns follows immediately from [183].

7. Problems related to random sequences

In this section we give a list of problems related to the study of multiple ergodic averages involving iterates given by random sequences of integers.

The random sequences that we work with are constructed by selecting a positive integer n to be a member of our sequence with probability $\sigma_n \in [0,1]$. More precisely, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent $0-1$ valued random variables with

$$\mathbb{P}(X_n = 1) := \sigma_n \ \text{ and } \ \mathbb{P}(X_n = 0) := 1 - \sigma_n$$

where σ_n is a decreasing sequence of positive real numbers that satisfies $\sum_{n=1}^\infty \sigma_n = \infty$ (in which case $\sum_{n=1}^\infty X_n(\omega) = +\infty$ almost surely). The random sequence $(a_n(\omega))_{n\in\mathbb{N}}$ is constructed by taking the positive integers n for which $X_n(\omega) = 1$ in increasing order. Equivalently, $a_n(\omega)$ is the smallest $k \in \mathbb{N}$ such that $X_1(\omega) + \cdots + X_k(\omega) = n$. If $\sigma_n = n^{-a}$ for some $a \in (0,1)$, then one can show that almost surely $a_n(\omega)/n^{1/(1-a)}$ converges to a non-zero constant. On the other hand, if $\sigma_n = 1/n$, then almost surely there exists a subsequence (n_k) of the integers, of density arbitrarily close to one, such that the sequence $(a_{n_k}(\omega))$ is lacunary [125] (this is no longer the case if $n\sigma_n \to \infty$). So it makes some sense to call *random non-lacunary sequences* the random sequences one gets when σ_n satisfies $n\sigma_n \to \infty$.

We say that a certain property holds almost surely for the sequences $(a_n(\omega))_{n\in\mathbb{N}}$, if there exists a universal set $\Omega_0 \in \mathcal{F}$, such that $\mathbb{P}(\Omega_0) = 1$, and for every $\omega \in \Omega_0$ the sequence $(a_n(\omega))_{n\in\mathbb{N}}$ satisfies the given property.

The next result was proved by M. Boshernitzan [47] for mean convergence and by J. Bourgain [52] for pointwise convergence (see also [170] for a nice exposition of these results).

Theorem. *If $n\sigma_n \to \infty$, then almost surely the following holds: for every invertible measure preserving system $(X, \mathcal{X}, \mu, T)$ and function $f \in L^\infty(\mu)$, the averages*

$$\frac{1}{N}\sum_{n=1}^{N} T^{a_n(\omega)} f \tag{18}$$

converge in the mean and their limit equals the $L^2(\mu)$-limit of the averages $\frac{1}{N}\sum_{n=1}^{N} T^n f$. Furthermore, if $n\sigma_n/(\log\log n)^{1+\delta} \to \infty$ for some $\delta > 0$, then the conclusion also holds pointwise.

It is known that the mean convergence result fails if $\sigma_n = 1/n$ and the pointwise convergence result fails if $\sigma_n = (\log\log n)^{1/3}/n$ (for both results see [125]). It is unclear whether the pointwise convergence result fails when say $\sigma_n = \log\log n/n$.

One would naturally like to extend the previous convergence result to multiple ergodic averages:

Problem 31. *Suppose that $n\sigma_n \to \infty$. Show that almost surely the following holds: For every probability space $(X, \mathcal{X}, \mu)$, commuting invertible measure preserving transformations $T_1, \ldots, T_\ell \colon X \to X$, and functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$, the averages*

$$\frac{1}{N}\sum_{n=1}^{N} T_1^{a_n(\omega)} f_1 \cdot \ldots \cdot T_\ell^{a_n(\omega)} f_\ell$$

converge in the mean and their limit equals the $L^2(\mu)$-limit of the averages $\frac{1}{N}\sum_{n=1}^{N} T_1^n f_1 \cdot \ldots \cdot T_\ell^n f_\ell$.

A special case of this problem was mentioned in 2004 by M. Wierdl, it can be found here http://math.stanford.edu/yk70/ik70-op.pdf. Mean convergence is known when $\ell = 2$ and $\sigma_n = n^{-a}$ where $a \in (0, 1/2)$ [88]. The methods of [88] also work for mean convergence when $\ell > 2$, but for a smaller range of eligible exponents a that depends on ℓ. Under the assumption that $n\sigma_n \to \infty$, pointwise convergence is known when the transformations are given by powers of the same nilrotation [78] and the functions are continuous. On the other hand, when $\sigma_n = n^{-a}$ for some $a > 1/2$, $\ell = 2$, and $T_2 = T_1^2$, mean convergence is not even known when we restrict ourselves to the class of weak mixing systems.

Next we mention a problem where one combines deterministic and random iterates.

Problem 32. *Suppose that $n\sigma_n \to \infty$. Show that almost surely the following holds: For every probability space $(X, \mathcal{X}, \mu)$, commuting invertible measure preserving transformations $T, S \colon X \to X$, and functions $f, g \in L^\infty(\mu)$, we have*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} T^n f \cdot S^{a_n(\omega)} g = \mathbb{E}(f|\mathcal{I}_T) \cdot \mathbb{E}(g|\mathcal{I}_S)$$

where the limit is taken in $L^2(\mu)$. Furthermore, if $\sigma_n = n^{-a}$ for some $a \in (0,1)$, show that the convergence also holds pointwise.

The desired convergence (in the mean and pointwise) is only known when $a \in (0, /14)$ [88]. The problem does not appear to be much easier to solve when $T = S$ is a weak mixing transformation and we are only interested in mean convergence.

In the next problem we are going to work with two random sequences with different growth rates that are chosen independently of each other. More precisely, let $X_1, Y_1, X_2, Y_2, \ldots$ be a sequence of independent 0-1 valued random variables with $\mathbb{P}(X_n = 1) := n^{-a}$ and $\mathbb{P}(Y_n = 1) := n^{-b}$ for some $a, b \in (0,1)$. We construct the random sequence $(a_n(\omega))$ by taking the positive integers n for which $X_n(\omega) = 1$ in increasing order, and the random sequence $(b_n(\omega))$ by taking the positive integers n for which $Y_n(\omega) = 1$ in increasing order.

Problem 33. *Suppose that $a \neq b$. Show that almost surely the following holds: For every probability space $(X, \mathcal{X}, \mu)$, commuting invertible measure preserving transformations $T, S\colon X \to X$, and functions $f, g \in L^\infty(\mu)$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^{a_n(\omega)} f \cdot S^{b_n(\omega)} g = \mathbb{E}(f|\mathcal{I}_T) \cdot \mathbb{E}(g|\mathcal{I}_S)$$

where the limit is taken in $L^2(\mu)$ or pointwise.

The problem seems non-trivial even when $T = S$ is a weak mixing transformation.

8. Extended bibliography organized by theme

Below we give a rather extensive bibliography with material that is directly related to the problems discussed before, organized by topic. We caution the reader that this is *not* a comprehensive list of articles in ergodic Ramsey theory, and in fact articles on several important topics are missing from this list. For instance, the reader will find very few articles related to actions of non-commuting measure preserving transformations, the richness of return times in various multiple recurrence results, and no articles related to problems in topological dynamics and applications in partition Ramsey theory. Luckily, there are several excellent places to look for such topics, for instance, the survey articles of V. Bergelson [24, 27, 28] cover a lot of related material and contain an extensive bibliography up to 2006.

Linear sequences: [2, 3, 5, 6, 7, 9, 10, 14, 15, 16, 17, 18, 25, 56, 33, 36, 37, 59, 60, 63, 64, 65, 68, 69, 73, 75, 76, 84, 91, 97, 98, 99, 111, 112, 113, 114, 115, 119, 123, 139, 151, 154, 155, 162, 173, 178, 181, 184, 186, 192, 193, 195, 196].

Polynomial sequences: [11, 12, 13, 22, 29, 30, 34, 35, 39, 40, 41, 43, 44, 52, 53, 54, 55, 58, 61, 62, 70, 74, 77, 79, 82, 83, 85, 87, 101, 108, 116, 124, 137, 142, 150, 156, 158, 168, 185].

Other sequences (smooth, random, prime numbers, generalized polynomials): [30, 31, 38, 42, 47, 50, 51, 72, 79, 80, 81, 88, 90, 104, 105, 125, 136, 148, 160, 161, 183, 189, 191].

Equidistribution on nilmanifolds and other nil-stuff: [8, 57, 78, 106, 107, 117, 118, 120, 121, 122, 138, 140, 141, 143, 144, 145, 146, 147, 148, 152, 153, 157, 163, 164, 165, 166, 174, 175, 194].

Books and survey articles on related topics: [1, 4, 8, 23, 24, 26, 27, 28, 66, 67, 71, 89, 92, 93, 94, 95, 96, 102, 110, 126, 130, 131, 132, 133, 135, 157, 159, 167, 169, 170, 172, 179, 180, 182, 187, 188].

References

[1] J. Aaronson. An introduction to infinite ergodic theory. *Mathematical Surveys and Monographs* **50**, American Mathematical Society, Providence, RI, 1997.

[2] J. Aaronson, H. Nakada. Multiple recurrence of Markov shifts and other infinite measure preserving transformations. *Israel J. Math.* **117** (2000), 285–310.

[3] I. Assani. Multiple recurrence and almost sure convergence for weakly mixing dynamical systems. *Israel J. Math.* **103** (1998), 111-124.

[4] I. Assani. Wiener Wintner ergodic theorems. *World Scientific Publishing Co., Inc.*, River Edge, NJ, 2003.

[5] I. Assani. Pointwise convergence of nonconventional averages. *Colloq. Math.* **102** (2005), no. 2, 245-262.

[6] I. Assani. Averages along cubes for not necessarily commuting m.p.t. Ergodic theory and related fields. *Contemp. Math.* **430**, Amer. Math. Soc., Providence, RI, (2007), 1–19.

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

[8] 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).

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

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

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

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

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

[14] T. Austin, T. Eisner, T. Tao. Nonconventional ergodic averages and multiple recurrence for von Neumann dynamical systems. To appear in *Pacific J. Math.*, `arXiv:0912.5093`

[15] B. Berend. Joint ergodicity and mixing. *J. Analyse Math.* **45** (1985), 255–284.

[16] B. Berend. Multiple ergodic theorems. *J. Analyse Math.* **50** (1988), 123–142.

[17] D. Berend, V. Bergelson. Jointly ergodic measure-preserving transformations. *Israel J. Math.* **49** (1984), no. 4, 307-314.

[18] D. Berend, V. Bergelson. Characterization of joint ergodicity for noncommuting transformations. *Israel J. Math.* **56** (1986), no. 1, 123-128.

[19] D. Berend, G. Kolesnik. Distribution modulo 1 of some oscillating sequences. *Israel J. Math.* **71** (1990), no. 2, 161-179.

[20] D. Berend, M. Boshernitzan, G. Kolesnik. Distribution modulo 1 of some oscillating sequences II. *Israel J. Math.* **92** (1995), no. 1-3, 125-147.

[21] D. Berend, M. Boshernitzan, G. Kolesnik. Distribution modulo 1 of some oscillating sequences III. *Acta Math. Hungar.* **95** (2002), no. 1-2, 1-20.

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

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

[24] V. Bergelson. Ergodic Ramsey Theory – an update, Ergodic Theory of $\mathbb{Z}^d$-actions (edited by M. Pollicott and K. Schmidt). *London Math. Soc. Lecture Note Series* **228** (1996), 1–61.

[25] V. Bergelson. The multifarious Poincare recurrence theorem. Descriptive set theory and dynamical systems. *London Math. Soc. Lecture Note Ser.* **277**, Cambridge Univ. press, Cambridge, (2000), 31–57.

[26] V. Bergelson. Ergodic theory and Diophantine problems. Topics in symbolic dynamics and applications. *London Math. Soc. Lecture Note Ser.* **279**, Cambridge Univ. Press, Cambridge, (2000), 167-205.

[27] V. Bergelson. Combinatorial and Diophantine Applications of Ergodic Theory (with appendices by A. Leibman and by A. Quas and M. Wierdl). *Handbook of Dynamical Systems*, Vol. 1B, B. Hasselblatt and A. Katok, eds., Elsevier, (2006), 745–841.

[28] V. Bergelson. Ergodic Ramsey Theory: a dynamical approach to static theorems. *Proceedings of the International Congress of Mathematicians*, Madrid 2006, Vol. II, 1655–1678.
[29] V. Bergelson, M. Boshernitzan, J. Bourgain. Some results on nonlinear recurrence. *J. Analyse Math.* **62** (1994), 29–46.
[30] V. Bergelson, I. Håland-Knutson. Sets of recurrence and generalized polynomials. Convergence in ergodic theory and probability (Columbus, OH, 1993), *Ohio State Univ. Math. Res. Inst. Publ.*, **5**, de Gruyter, Berlin, (1996), 91–110.
[31] 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
[32] V. Bergelson, I. Håland-Knutson, R. McCutcheon. IP Systems, generalized polynomials and recurrence. *Ergodic Theory Dynam. Systems* **26** (2006), no. 4, 999–1019.
[33] V. Bergelson, B. Host, B. Kra, with an appendix by I. Ruzsa. Multiple recurrence and nilsequences. *Inventiones Math.* **160** (2005), no. 2, 261–303.
[34] V. Bergelson, B. Host, R. McCutcheon, F. Parreau. Aspects of uniformity in recurrence. *Colloq. Math.* **84/85** (2000), no. 2, 549–576.
[35] V. Bergelson, A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.* **9** (1996), 725–753.
[36] V. Bergelson, A. Leibman. A nilpotent Roth theorem. *Inventiones Mathematicae* **147** (2002), 429–470
[37] V. Bergelson, A. Leibman. Failure of Roth theorem for solvable groups of exponential growth. *Ergodic Theory Dynam. Systems* **24** (2004), no. 1, 45–53.
[38] V. Bergelson, A. Leibman. Distribution of values of bounded generalized polynomials. *Acta Math.* **198** (2007), 155–230.
[39] V. Bergelson, A. Leibman, E. Lesigne. Complexities of finite families of polynomials, Weyl systems, and constructions in combinatorial number theory. *J. Analyse Math.* **103** (2007), 47–92.
[40] V. Bergelson, A. Leibman, E. Lesigne. Intersective polynomials and the polynomial Szemerédi theorem. *Adv. Math.* **219** (2008), no. 1, 369–388.
[41] V. Bergelson, A. Leibman, C. Moreira. From discrete- to continuous-time ergodic theorems. *Preprint.* Available at http://www.math.osu.edu/ leibman/preprints/
[42] 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`.
[43] V. Bergelson, R. McCutcheon. Uniformity in polynomial Szemerédi theorem, Ergodic Theory of $\mathbb{Z}^d$-actions (edited by M. Pollicott and K. Schmidt). *London Math. Soc. Lecture Note Series* **228** (1996), 273–296.
[44] V. Bergelson, R. McCutcheon. Idempotent ultrafilters, multiple weak mixing and Szemerédi's theorem for generalized polynomials. *J. Analyse Math.* **111** (2010), 77–130.
[45] M. Boshernitzan. An extension of Hardy's class L of "Orders of Infinity". *J. Analyse Math.* **39** (1981), 235–255.
[46] M. Boshernitzan. New "Orders of Infinity". *J. Analyse Math.* **41** (1982), 130–167.
[47] M. Boshernitzan. Homogeneously distributed sequences and Poincaré sequences of integers of sublacunary growth. *Monatsh. Math.* **96** (1983), no. 3, 173–181.
[48] M. Boshernitzan. Uniform distribution and Hardy fields. *J. Analyse Math.* **62** (1994), 225–240.
[49] M. Boshernitzan, E. Glasner. On two recurrence problems. *Fund. Math.* **206** (2009), 113–138.
[50] M. Boshernitzan, G. Kolesnik, A. Quas, M. Wierdl. Ergodic averaging sequences. *J. Analyse Math.* **95** (2005), 63–103.
[51] M. Boshernitzan, M. Wierdl. Ergodic theorems along sequences and Hardy fields. *Proc. Nat. Acad. Sci. U.S.A.* **93** (1996), no. 16, 8205-8207.
[52] J. Bourgain. On the maximal ergodic theorem for certain subsets of the positive integers. *Israel J. Math.* **61** (1988), 39–72.
[53] J. Bourgain. An approach to pointwise ergodic theorems. *Lecture Notes in Math.* **1317**, Springer, Berlin, (1988), 204-223.
[54] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. With an appendix by the author, H. Furstenberg, Y. Katznelson and D. Ornstein. *Inst. Hautes tudes Sci. Publ. Math.* **69** (1989), 5-45.
[55] J. Bourgain. Double recurrence and almost sure convergence. *J. Reine Angew. Math.* **404** (1990), 140–161.

[56] T. Brown, R. Graham, B. Landman. On the set of common differences in van der Waerden's theorem on arithmetic progressions. *Canad. Math. Bull.* **42** (1999), 25–36.
[57] O. Camarena, B. Szegedy. Nilspaces, nilmanifolds and their morphisms. *Preprint*, `arXiv:1009.3825`.
[58] Q. Chu. Convergence of weighted polynomial multiple ergodic averages. *Proc. Amer. Math. Soc.* **137** (2009), no. 4, 1363-1369.
[59] Q. Chu. Convergence of multiple ergodic averages along cubes for several commuting transformations. *Studia Math.* **196** (2010), no. 1, 13-22.
[60] Q. Chu. Multiple recurrence for two commuting transformations. To appear *Ergodic Theory Dynam. Systems*, `arXiv:0912.3381`.
[61] 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`.
[62] 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`.
[63] J-P. Conze, E. Lesigne. Théorèmes ergodiques pour des mesures diagonales. *Bull. Soc. Math. France* **112** (1984), no. 2, 143–175.
[64] 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.
[65] J-P. Conze, E. Lesigne. Sur un théorème ergodique pour des mesures diagonales. *C. R. Acad. Sci. Paris, Série I* **306** (1988), 491–493.
[66] I. Cornfeld, S. Fomin, Y. Sinai. Ergodic theory. Translated from the Russian by A. B. Sosinskii Grundlehren der Mathematischen Wissenschaften *Fundamental Principles of Mathematical Sciences* **245** Springer-Verlag, New York, 1982.
[67] L. Corwin, F. Greenleaf. Representations of nilpotent Lie groups and their applications. Part I. Basic theory and examples. *Cambridge Studies in Advanced Mathematics* **18**, Cambridge University Press, Cambridge, 1990.
[68] C. Demeter. Pointwise convergence of the ergodic bilinear Hilbert transform. *Illinois J. Math.* **51** (2007), no. 4, 1123-1158.
[69] C. Demeter, C. Thiele. On the two-dimensional bilinear Hilbert transform. *Amer. J. Math.* **132** (2010), no. 1, 201-256.
[70] J-M. Derrien, E. Lesigne. Un théorème ergodique polynomial ponctuel pour les endomorphismes exacts et les K-systèmes. *Ann. Inst. H. Poincaré Probab. Statist.* **32** (1996), no. 6, 765-778.
[71] M. Einsiedler, T. Ward, Ergodic theory with a view towards number theory. *Graduate Texts in Mathematics* **259**, Springer-Verlag London, Ltd., London, 2011.
[72] A. Fan, D. Schneider. Recurrence properties of sequences of integers. *Sci. China Math.* 53 (2010), no. 3, 641–656.
[73] A. Fish. Solvability of linear equations within weak mixing sets. To appear in *Isr. J. Math.*, `arXiv:0704.0600`.
[74] A. Fish. Polynomial largeness of sumsets and totally ergodic sets. To appear in *Online J. Anal. Comb.*, `arXiv:0711.3201`.
[75] A. Forrest. Two techniques in multiple recurrence. Ergodic theory and its connections with harmonic analysis (Alexandria, 1993), 273290, *London Math. Soc. Lecture Note Ser.* **205**, Cambridge Univ. Press, Cambridge, 1995.
[76] N. Frantzikinakis. The structure of strongly stationary systems. *J. Analyse Math.* **93** (2004), 359–388.
[77] N. Frantzikinakis. Multiple ergodic averages for three polynomials and applications. *Trans. Amer. Math. Soc.* **360** (2008), no. 10, 5435–5475.
[78] N. Frantzikinakis. Equidistribution of sparse sequences on nilmanifolds. *J. Analyse Math.* **109** (2009), 1–43.
[79] N. Frantzikinakis. Multiple recurrence and convergence for Hardy sequences of polynomial growth. *J. Analyse Math.* **112** (2010), 79–135.
[80] N. Frantzikinakis, B. Host, B. Kra. Multiple recurrence and convergence for sets related to the primes. *J. Reine Angew. Math.* **611** (2007), 131–144.
[81] N. Frantzikinakis, B. Host, B. Kra. The polynomial multidimensional Szemerédi Theorem along shifted primes. Preprint, `arXiv:1009.1484`.

[82] N. Frantzikinakis, M. Johnson, E. Lesigne, M. Wierdl. Powers of sequences and convergence. *Ergodic Theory Dynam. Systems* **30** (2010), no. 5, 1431–1456.
[83] N. Frantzikinakis, B. Kra. Polynomial averages converge to the product of integrals. *Isr. J. Math.* **148** (2005), 267–276.
[84] N. Frantzikinakis, B. Kra. Convergence of multiple ergodic averages for some commuting transformations. *Ergodic Theory Dynam. Systems* **25** (2005), no. 3, 799–809.
[85] N. Frantzikinakis, B. Kra. Ergodic averages for independent polynomials and applications. *J. London Math. Soc.* **74** (2006), no. 1, 131–142.
[86] 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.
[87] N. Frantzikinakis, E. Lesigne, M. Wierdl. Powers of sequences and recurrence. *Proc. Lond. Math. Soc.* (3) **98** (2009), no. 2, 504–530.
[88] N. Frantzikinakis, E. Lesigne, M. Wierdl. Random sequences and pointwise convergence of multiple ergodic averages. Preprint, `arXiv:1012.1130`.
[89] N. Frantzikinakis, R. McCutcheon. Ergodic Theory: Recurrence. Survey. *Encyclopedia of complexity and Systems Science.* Springer, (2009), Part 5, 3083–3095.
[90] N. Frantzikinakis, M. Wierdl. A Hardy field extension of Szemerédi's theorem. *Adv. Math.* **222**, (2009), 1–43.
[91] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Analyse Math.* **31** (1977), 204–256.
[92] H. Furstenberg. Recurrence in ergodic theory and combinatorial number theory. *Princeton University Press*, Princeton, 1981.
[93] H. Furstenberg. Poincaré recurrence and number theory. *Bull. Amer. Math. Soc. (N.S.)* **5** (1981), no. 3, 211-234.
[94] H. Furstenberg. Nonconventional ergodic averages. The legacy of John von Neumann (Hempstead, NY, 1988), *Proc. Sympos. Pure Math.*, **50**, Amer. Math. Soc., Providence, RI, (1990), 43–56.
[95] H. Furstenberg. Recurrent ergodic structures and Ramsey theory. *Proceedings of the International Congress of Mathematicians*, Vol. I, II (Kyoto, 1990), Math. Soc. Japan, Tokyo, (1991), 1057-1069,.
[96] H. Furstenberg. A polynomial Szemerédi theorem. Combinatorics, Paul Erdös is eighty, Vol. 2 (Keszthely, 1993) *Bolyai Soc. Math. Stud.*, **2**, Jànos Bolyai Math. Soc., Budapest, (1996), 253-269.
[97] H. Furstenberg, Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. *J. Analyse Math.* **34** (1979), 275–291.
[98] H. Furstenberg, Y. Katznelson, D. Ornstein. The ergodic theoretical proof of Szemerédi's theorem. *Bull. Amer. Math. Soc. (N.S.)* **7** (1982), no. 3, 527–552.
[99] H. Furstenberg, Y. Katznelson, B. Weiss. Ergodic theory and configurations in sets of positive density. Mathematics of Ramsey theory, *Algorithms Combin.*, **5**, Springer, Berlin, (1990), 184–198.
[100] H. Furstenberg, Y. Katznelson. A density version of the Hales-Jewett theorem. *J. Analyse Math.* **57** (1991), 64–119.
[101] 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.
[102] E. Glasner. Ergodic theory via joinings. *Mathematical Surveys and Monographs* **101**. American Mathematical Society, Providence, RI, 2003.
[103] W. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.* **11** (2001), 465–588.
[104] B. Green, T. Tao. The primes contain arbitrarily long arithmetic progressions. *Annals Math.* **167** (2008), 481–547.
[105] B. Green, T. Tao. Linear equations in primes. *Annals Math.* **171** (2010), 1753–1850. `arXiv:0606088`.
[106] B. Green, T. Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. To appear in *Annals Math.*, `arXiv:0709.3562`.
[107] B. Green, T. Tao. The Mobious function is strongly orthogonal to nilsequences. To appear in *Annals. Math.*, `arXiv:0807.1736`.
[108] J. Griesmer. Abundant configurations in sumsets with one dense summand. *Preprint*, `arXiv:1011.4657`.
[109] G. Hardy. Properties of logarithmico-exponential functions. *Proc. London Math. Soc.* (2) **10** (1912), 54–90.

[110] G. Hardy. Orders of Infinity. *Cambridge Tracts in Math. and Math. Phys.*, 12 (2nd edition), Cambridge, (1924).
[111] B. Host. Ergodic seminorms for commuting transformations and applications. *S*tudia Math. **195** (1) (2009), 31–49.
[112] B. Host, B. Kra. Convergence of Conze-Lesigne averages. *Ergodic Theory Dynam. Systems* **21** (2001), no. 2, 493-509.
[113] B. Host, B. Kra. An Odd Furstenberg-Szemerédi Theorem and quasi-affine systems. *J. Analyse Math.*, **86** (2002), 183–220.
[114] B. Host, B. Kra. Averaging along cubes. Modern dynamical systems and applications. *Cambridge Univ. Press*, Cambridge, (2004), 123–144.
[115] B. Host, B. Kra. Non-conventional ergodic averages and nilmanifolds. *Annals Math.* **161** (2005), 397–488.
[116] B. Host, B. Kra. Convergence of polynomial ergodic averages. *Isr. J. Math.* **149** (2005), 1–19.
[117] B. Host, B. Kra. Analysis of two step nilsequences. *Ann. Inst. Fourier (Grenoble)* **58** (2008), 1407–1453.
[118] B. Host, B. Kra. Parallelepipeds, nilpotent groups and Gowers norms. *Bull. Soc. Math. France* **136** (2008), no. 3, 405-437.
[119] B. Host, B. Kra. Uniformity seminorms on l^∞ and applications. *J. Analyse Math.* **108** (2009), 219–276.
[120] B. Host, B. Kra. Nil-Bohr sets of integers. To appear in *Ergodic Theory Dynam. Systems.*, arXiv:0903.1642.
[121] B. Host, B. Kra, A. Maass. Nilsequences and a topological structure theorem. *Advances in Math.* **224** (2010), 103-129.
[122] B. Host, A. Maass. Two step nilsystems and parallelepipeds. *Bull. Soc. Math. France* **135** (2007), no. 3, 367-405.
[123] E. Jenvey. Strong stationarity and De Finetti's theorem. *J. d'Analyse Math.* **73** (1997), 1–18.
[124] M. Johnson. Convergence of polynomial ergodic averages for some commuting transformations. *Illinois J. Math.* **53** (2009), no. 3, 865-882.
[125] R. Jones, M. Lacey, M. Wierdl. Integer sequences with big gaps and the pointwise ergodic theorem. *Ergodic Theory Dynam. Systems* **19** (1999), no.5, 1295–1308.
[126] S. Kalikow, R. McCutcheon. An outline of ergodic theory. *Cambridge Studies in Advanced Mathematics*, **122**. Cambridge University Press, Cambridge, 2010.
[127] T. Kamae, M. Mendès-France. Van der Corput's difference theorem. *Isr. J. Math.* **31** (1978), 335–342.
[128] A. Karatsuba. Estimates for trigonometric sums by Vinogradov's method, and some applications. *Trudy Mat. Inst. Steklov*, **112**, (1971), 241–255; English Transl., *Proc. Steklov Inst. Math.* **112** (1973), 251–265.
[129] Y. Katznelson. Chromatic numbers of Cayley graphs on $\mathbb{Z}$ and recurrence. *Combinatorica* **21** (2001) 211–219.
[130] B. Kra. The Green-Tao Theorem on arithmetic progressions in the primes: an ergodic point of view. *Bull. Amer. Math. Soc.* **43** (2006), 3–23.
[131] B. Kra. From combinatorics to ergodic theory and back again. *Proceedings of International Congress of Mathematicians*, Madrid 2006, Vol. III, 57–76.
[132] B. Kra. Ergodic methods in additive combinatorics. Additive combinatorics, 103-143, *CRM Proc. Lecture Notes*, **43**, Amer. Math. Soc., Providence, RI, 2007. *Proceedings of International Congress of Mathematicians*, Madrid 2006, Vol. III, 57–76.
[133] U. Krengel. Ergodic theorems. With a supplement by Antoine Brunel. *de Gruyter Studies in Mathematics*, **6**, Walter de Gruyter & Co., Berlin, (1985).
[134] I. Kriz. Large independent sets in shift-invariant graphs: solution of Bergelson's problem. *Graphs Combin.* **3** (1987), no. 2, 145-158.
[135] L. Kuipers, H. Niederreiter. Uniform distribution of sequences. Pure and Applied Mathematics. *Wiley-Interscience*, New York-London-Sydney, 1974.
[136] M. Lacey, K. Petersen, M. Wierdl, D. Rudolph. Random ergodic theorems with universally representative sequences. *Ann. Inst. H. Poincare Probab. Statist.* **30** (1994), no. 3, 353-395.
[137] A. Leibman. Multiple recurrence theorem for measure preserving actions of a nilpotent group. *Geom. Funct. Anal.* **8** (1998), 853–931.
[138] A. Leibman. Polynomial mappings of groups. *Israel J. Math.* **129** (2002), 29–60.
[139] A. Leibman. Lower bounds for ergodic averages. *Ergodic Theory Dynam. Systems* **22** (2002), no. 3, 863–872.

[140] 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.
[141] 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.
[142] A. Leibman. Convergence of multiple ergodic averages along polynomials of several variables. *Isr. J. Math.* **146** (2005), 303–316.
[143] A. Leibman. Rational sub-nilmanifolds of a compact nilmanifold. *Ergodic Theory Dynam. Systems* **26** (2006), no. 3, 787-798.
[144] A. Leibman. Orbits on a nilmanifold under the action of a polynomial sequence of translations. *Ergodic Theory Dynam. Systems* **27** (2007), 1239–1252.
[145] A. Leibman. Ergodic components of an extension by a nilmanifold. *Illinois J. Math.* **52** (2008), no. 3, 957-965.
[146] A. Leibman. Orbit of the diagonal in the power of a nilmanifold. *Trans. Amer. Math. Soc.* **362** (2010), 1619–1658.
[147] A. Leibman. Multiple polynomial sequences and nilsequences. To appear in *Ergodic Theory Dynam. Systems.*
[148] A. Leibman. A canonical form and the distribution of values of generalized polynomials, To appear in *Isr. J. Math.*
[149] D. Leitmann. On the uniform distribution of some sequences. *J. London Math. Soc.* (2) **14** (1976), no. 3, 430-432.
[150] E. Lesigne. Sur la convergence ponctuelle de certaines moyennes ergodiques. *C. R. Acad. Sci. Paris Sér. I Math.* **298** (1984), no. 17, 425-428.
[151] E. Lesigne. Théorèmes ergodiques ponctuels pour des mesures diagonales. Cas des systèmes distaux. [Pointwise ergodic theorems for diagonal measures. The case of distal systems] *Ann. Inst. H. Poincaré Probab. Statist.* **23** (1987), no. 4, 593-612.
[152] E. Lesigne. Théorèmes ergodiques pour une translation sur un nilvariété. *Ergodic Theory Dynam. Systems* **9** (1989), no. 1, 115-126.
[153] 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.
[154] E. Lesigne. Équations fonctionnelles, couplages de produits gauches et théorèmes ergodiques pour mesures diagonales. *Bull. Soc. Math. France* **121** (1993), no. 3, 315–351.
[155] E. Lesigne, B. Rittaud, T. de la Rue. Weak disjointness of measure preserving dynamical systems. *Ergodic Theory Dynam. Systems* **23** (2003), 1173–1198.
[156] N. Lyall, A. Magyar. Simultaneous polynomial recurrence. To appear in *Bull. Lond. Math. Soc.*
[157] A. Malcev. On a class of homogeneous spaces. *Amer. Math. Soc. Translation* **39** (1951).
[158] D. McClendon. On the maximal Weyl complexity of families of four polynomials. *Preprint*, Available at http://www.math.northwestern.edu/ dmm/papers/fourpolynomials.pdf
[159] R. McCutcheon. Elemental methods in ergodic Ramsey theory. *Lecture Notes in Mathematics* **1722**. Springer-Verlag, Berlin, 1999.
[160] R. McCutcheon. FVIP systems and multiple recurrence. *Israel J. Math.* **146** (2005), 157–188.
[161] R. McCutcheon, A. Quas. Generalized polynomials and mild mixing. *Canad. J. Math.* **61** (2009), no. 3, 656-673.
[162] D. Meiri. Generalized correlation sequences. M.Sc. thesis 1990. Tel Aviv University. Available at http://taalul.com/David/Math/ma.pdf
[163] W. Parry. Ergodic properties of affine transformations and flows on nilmanifolds. *Amer. J. Math.* **91** (1969), 757–771.
[164] W. Parry. Dynamical systems on nilmanifolds. *Bull. London Math. Soc.* **2** (1970), 37-40.
[165] W. Parry. Metric classification of ergodic nilflows and unipotent affines. *Amer. J. Math.* **93** (1971), 819-828.
[166] W. Parry. Dynamical representations in nilmanifolds. *Compositio Math.* **26** (1973), 159-174.
[167] K. Petersen. Ergodic theory. *Cambridge Studies in Advanced Mathematics*, **2**. Cambridge University Press, Cambridge, 1989.
[168] A. Potts. Multiple ergodic averages for flows and an application. *Preprint*, `arXiv:0910.3687`.

[169] Queffélec, Martine(F-LILL-LM) Substitution dynamical systemsspectral analysis. Second edition. *Lecture Notes in Mathematics* **1294**, Springer-Verlag, Berlin, 2010.

[170] J. Rosenblatt, M. Wierdl. Pointwise ergodic theorems via harmonic analysis. Ergodic theory and its connections with harmonic analysis (Alexandria, 1993). *London Math. Soc. Lecture Note Ser.*, **205**, Cambridge Univ. Press, Cambridge, (1995), 3–151.

[171] M. Rosenlicht. Hardy fields. *J. Math. Anal. Appl.* **91** (1983), no. 2, 297–311.

[172] D. Rudolph. Fundamentals of measurable dynamics. Ergodic theory on Lebesgue spaces. *Oxford Science Publications.* The Clarendon Press, Oxford University Press, New York, 1990

[173] 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.

[174] N. Shah. Limit distributions of polynomial trajectories on homogeneous spaces. *Duke Math. J.* **75** (1994), no. 3, 711-732.

[175] N. Shah. Invariant measures and orbit closures on homogeneous spaces for actions of subgroups generated by unipotent elements. Lie groups and ergodic theory (Mumbai, 1996), *Tata Inst. Fund. Res. Stud. Math.*, **14**, Tata Inst. Fund. Res., Bombay, (1998), 229-271.

[176] I. Stux. On the uniform distribution of prime powers. *Comm. Pure Appl. Math.* **27** (1974), 729-740.

[177] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.* **27** (1975), 299–345.

[178] T. Tao. A quantitative ergodic theory proof of Szemerédi's theorem. *Electron. J. Combin.* **1**3 (2006), 1–49.

[179] T. Tao. The ergodic and combinatorial approaches to Szemerédi's theorem. Additive combinatorics, *CRM Proc. Lecture Notes* **43**, Amer. Math. Soc., Providence, RI, (2007), 145-193.

[180] T. Tao. The dichotomy between structure and randomness, arithmetic progressions, and the primes. *International Congress of Mathematicians. Vol. I*, Eur. Math. Soc., Zrich, (2007), 581-608.

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

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

[183] T. Tao, T. Ziegler. The primes contain arbitrarily long polynomial progressions. *Acta Math.* **201** (2008), 213–305.

[184] J-P. Thouvenot. La demonstration de Furstenberg du théorème de Szemerédi sur les progressions arithmétiques. Seminaire Bourbaki, Vol. 1977/78, Expose No.518, *Lect. Notes Math.* **710** (1979), 221–232.

[185] J-P. Thouvenot. On the almost-sure convergence of ergodic means following certain subsequences of the integers (after Jean Bourgain). Seminaire Bourbaki, Vol. 1989/90, 42ème année, *Astérisque* **189-190**, Exp. No.719, (1990) 133–153.

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

[187] P. Walters. An introduction to ergodic theory. *Graduate Texts in Mathematics*, **79**, Springer-Verlag, New York-Berlin, (1982).

[188] B. Weiss. Single orbit dynamics. CBMS Regional Conference Series in Mathematics, **95**, *American Mathematical Society*, Providence, RI, (2000).

[189] M. Wierdl. Pointwise ergodic theorem along the prime numbers. *Israel J. Math.* **64** (1988), 315–336.

[190] D. Wolke. Zur Gleichverteilung einiger Zahlenfolgen. *Math. Z.* **142** (1975), 181-184.

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

[192] Q. Zhang. On convergence of the averages $\frac{1}{N}\sum_{n=1}^{N} f_1(R^n x) f_2(S^n x) f_3(T^n x)$, *Monatsh. Math.* **122** (1996), 275-300.

[193] T. Ziegler. An application of ergodic theory to a problem in geometric Ramsey theory. *Israel J. Math.* **114** (1999), 271–288.

[194] T. Ziegler. A nonconventional ergodic theorem for a nilsystem. *Ergodic Theory Dynam. Systems* **25** (2005), no. 4, 1357–1370.

[195] 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.

[196] 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