Van der Corput's difference theorem: some modern developments

Vitaly Bergelson, Joel Moreira[*]

Department of Mathematics, The Ohio State University, Columbus, OH 43210, USA

Abstract

We discuss various forms of the classical van der Corput's difference theorem and explore applications to and connections with the theory of uniform distribution, ergodic theory, topological dynamics and combinatorics.
© 2015 Royal Dutch Mathematical Society (KWG). Published by Elsevier B.V. All rights reserved.

Keywords: Uniform distribution; Van der Corput difference theorem

1. Introduction

Van der Corput's classical Difference Theorem [33] is traditionally stated as follows:

DT: Let $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers. If for any $h \in \mathbb{N} = \{1, 2, \ldots\}$, the sequence $(x_{n+h} - x_n)_{n=1}^{\infty}$ is uniformly distributed mod 1, then $(x_n)_{n=1}^{\infty}$ is uniformly distributed mod 1.

Let $f(x) \in \mathbb{R}[x]$ and let $x_n = f(n)$, $n \in \mathbb{N}$. Then for any $h \in \mathbb{N}$, the degree of the polynomial $f(n+h) - f(n) = x_{n+h} - x_n$ equals deg $f - 1$. This rather trivial observation, together with DT, allows one to obtain a streamlined inductive proof of Weyl's celebrated theorem, [63], which states that for any polynomial $f \in \mathbb{R}[x]$ which has at least one coefficient, other than the constant term, irrational, the sequence $x_n = f(n)$, $n \in \mathbb{N}$, is uniformly distributed mod 1. While the proof of DT can be condensed to just a few lines (see the proof of Theorem 2.3 in the next section), van der Corput's Difference Theorem contains in embryonic form a powerful

[*] Corresponding author.
E-mail addresses: vitaly@math.ohio-state.edu (V. Bergelson), moreira@math.ohio-state.edu (J. Moreira).

idea of Complexity Reduction. In the subsequent sections we will provide numerous examples of applications of various generalizations of DT. These applications include results in the theory of uniform distribution, as well as some new multiple recurrence theorems in Ergodic Theory and Topological Dynamics.

The structure of the paper is as follows. In Section 2 we provide a short proof of DT and also formulate and prove a general form of DT that will be utilized in later sections. Section 3 is devoted to various aspects of the phenomenon of well distribution. Section 4 deals with uniform distribution along Besicovitch almost periodic sequences. Among other things, we establish new uniform distribution results involving the sequence of squarefree numbers. In Section 5 we provide some additional applications of the (generalized) DT. In particular, we establish a version of Weyl's equidistribution theorem in the group $\mathbb{A}/\mathbb{Q}$ (where $\mathbb{A}$ is the additive group of adeles). We also prove a novel kind of ergodic theorem involving measure preserving actions of the additive and multiplicative groups of a countable field. Finally, in sub-Section 5.3 we prove a rather general 'non-linear' mean ergodic theorem involving a family of commuting measure preserving $\mathbb{R}$-actions. In the relatively short Section 6 we formulate a theorem due to Kátai which may be interpreted as a multiplicative version of DT and demonstrate its power by providing an application involving the classical Möbius function. In Section 7 we consider limits along idempotent ultrafilters and prove, with the help of an ultrafilter variant of DT, a polynomial ergodic theorem for mildly mixing transformations. Finally, in Section 8 we show that the idea of Complexity Reduction can be applied to multiple recurrence in topological dynamics, with new applications to combinatorics.

Regrettably, due to various constraints, numerous additional applications of, and connections with, DT were not included in this paper. See, for example, [11,19,20,22,30,37,40,45,48,50,56, 61].

2. Some variants of van der Corput's difference theorem

We start with the classical definition of uniform distribution, which goes back to the groundbreaking paper of Weyl [63].

Definition 2.1. A sequence $(x_n)_{n\in\mathbb{N}}$ taking values in the unit interval $[0, 1)$ is *uniformly distributed* if for any subinterval $[a, b) \subset [0, 1)$ the proportion of those $n \in \mathbb{N}$ for which $x_n \in [a, b)$ is $b - a$. More precisely, if

$$\lim_{N\to\infty} \frac{\left|\{n \in \{1, \ldots, N\} : x_n \in [a, b)\}\right|}{N} = b - a \tag{2.1}$$

for all $0 \le a < b \le 1$.

One can informally say that a sequence is uniformly distributed if each interval gets its fair share of points. Since there are uncountably many subintervals $[a, b) \subset [0, 1)$, it is a priori not clear from the definition that such sequences exist (although one should certainly believe that a sequence of i.i.d. random variables taking values in $[0, 1)$ and having probability density function 1 will have this property almost surely).

The following classical theorem, providing useful characterizations of uniform distribution, is originally due to Weyl [63]. For a proof see, for instance, [51, Theorems 1.1.1 and 1.2.1].

Theorem 2.2. *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence taking values in $[0, 1)$. The following are equivalent:*

(1) $(x_n)_{n\in\mathbb{N}}$ *is uniformly distributed.*
(2)

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}f(x_n) = \int_0^1 f(x)\,\mathrm{d}x \quad \forall f \in C[0, 1]$$

(3)

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e^{2\pi i h x_n} = 0 \quad \forall h \in \mathbb{N}.$$

Condition (3) of Theorem 2.2 is known as the *Weyl criterion* for uniform distribution. Using it, one can easily show, for instance, that for any irrational $\alpha \in \mathbb{R}$, the sequence $x_n = n\alpha$ mod 1[1] is uniformly distributed.

In this paper, as is customary in the theory of uniform distribution, we identify $[0, 1)$ with the torus $\mathbb{T} := \mathbb{R}/\mathbb{Z}$. The van der Corput difference theorem, DT, gives a sufficient condition for a sequence $(x_n)_{n\in\mathbb{N}}$ taking values in $\mathbb{T}$ to be uniformly distributed:

Theorem 2.3 (*Van der Corput's Difference Theorem, [33]*)**.** *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence taking values in the torus $\mathbb{T}$. Assume that for every $d \in \mathbb{N}$, the sequence $n \mapsto x_{n+d} - x_n$ is uniformly distributed. Then $(x_n)_{n\in\mathbb{N}}$ is uniformly distributed.*

In view of Weyl's criterion, the assumption of Theorem 2.3 is that, for any $d, h \in \mathbb{N}$,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e^{2\pi i h(x_{n+d}-x_n)} = 0$$

and to prove Theorem 2.3 we have to show that

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

for any $h \in \mathbb{N}$. Therefore, Theorem 2.3 is a corollary of the following version of DT.

Theorem 2.4. *Let $(u_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathbb{C}$. Assume that for every $d \in \mathbb{N}$ we have*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}u_{n+d}\overline{u_n} = 0. \tag{2.2}$$

Then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}u_n = 0.$$

---

[1] For a real number x, we denote by $\lfloor x \rfloor$ the largest integer not exceeding x, and denote by x mod 1 the number $x - \lfloor x \rfloor \in [0, 1)$.

Proof. Our goal will be achieved if we show that for any $\epsilon > 0$

$$\limsup_{N \to \infty} \left| \frac{1}{N} \sum_{n=1}^{N} u_n \right| < \epsilon.$$

Notice now that for any $\epsilon > 0$ and any $D \in \mathbb{N}$, if $N \in \mathbb{N}$ is large enough we have

$$\left| \frac{1}{N} \sum_{n=1}^{N} u_n - \frac{1}{N} \frac{1}{D} \sum_{n=1}^{N} \sum_{d=1}^{D} u_{n+d} \right| < \frac{\epsilon}{2}.$$

Hence it suffices to show that, if D is large enough,

$$\limsup_{N \to \infty} \left| \frac{1}{N} \frac{1}{D} \sum_{n=1}^{N} \sum_{d=1}^{D} u_{n+d} \right| < \frac{\epsilon}{2}.$$

Using the Cauchy–Schwarz inequality in $\mathbb{R}^N$ we have

$$\limsup_{N \to \infty} \left| \frac{1}{N} \frac{1}{D} \sum_{n=1}^{N} \sum_{d=1}^{D} u_{n+d} \right|^2 \leq \limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| \frac{1}{D} \sum_{d=1}^{D} u_{n+d} \right|^2$$

$$= \limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \frac{1}{D^2} \sum_{d_1, d_2 = 1}^{D} u_{n+d_1} \overline{u_{n+d_2}}$$

$$\leq \frac{1}{D^2} \sum_{d_1, d_2 = 1}^{D} \limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} u_{n+d_1} \overline{u_{n+d_2}}. \tag{2.3}$$

Note that, for $d_1 \neq d_2$, it follows from (2.2) that $\frac{1}{N} \sum_{n=1}^{N} u_{n+d_1} \overline{u_{n+d_2}} \to 0$ as $N \to \infty$. We conclude that the quantity in (2.3) is bounded by $\frac{D}{D^2} = \frac{1}{D}$ which is arbitrarily small for large enough D. $\quad\square$

As was alluded to in the introduction, DT allows one to easily show that for any polynomial $f \in \mathbb{R}[x]$ which has an irrational coefficient, other than the constant term, the sequence $(f(n) \bmod 1)_{n \in \mathbb{N}}$ is uniformly distributed. We will describe now another immediate application of DT.

A classical result in the theory of uniform distribution is Fejér's theorem, which we will presently formulate. Given a sequence $n \mapsto f(n)$ of real numbers, we define its discrete derivative $\Delta f : \mathbb{N} \to \mathbb{R}$ by the formula $\Delta f(n) = f(n+1) - f(n)$. We also denote by $\Delta^s f$ the iterated discrete derivative, i.e. $\Delta^s f = \Delta(\Delta^{s-1} f)$.

Theorem 2.5 (*Fejér; See, for Example, [51, Theorem 1.2.5]*). *Let $(f(n))_{n \in \mathbb{N}}$ be sequence in $\mathbb{R}$. Assume that $\Delta f(n)$ is eventually decreasing and satisfies*

$$\lim_{n \to \infty} \Delta f(n) = 0 \quad and \quad \lim_{n \to \infty} n \Delta f(n) = \infty.$$

Then the sequence $(f(n) \bmod 1)_{n \in \mathbb{N}}$ is uniformly distributed.

Definition 2.6. A function $f : \mathbb{R} \to \mathbb{R}$ is called *tempered* if for some $\ell \in \{0, 1, 2, \ldots\}$ the first $\ell + 1$ derivatives of f exist, are continuous on some interval (a, ∞) and satisfy the following conditions:

(1) $f^{(\ell+1)}(s)$ (eventually) decreases to 0.
(2) $\lim_{s\to\infty} f^{(\ell)}(s) = \lim_{s\to\infty} s f^{(\ell+1)}(s) = \infty$.

The sequence $\big(f(n)\big)_{n\in\mathbb{N}}$ is called a *tempered sequence* and ℓ is called the *degree* of f.

Here are some examples of tempered sequences.

- $x_n = \sum_{i=1}^{k} c_i n^{\alpha_i}$, where $c_i, \alpha_i \in \mathbb{R}$, $c_k, \alpha_k > 0$ and $\alpha_1 < \cdots < \alpha_k$.
- $x_n = cn^{\alpha}\big(\cos(\log^{\beta} n) + d\big)$, where $\alpha, \beta, c, d \in \mathbb{R}$ satisfy $\alpha, c > 0$, $\beta < 1$, and $d > 1$.

Fejér's theorem together with van der Corput's DT readily implies:

Corollary 2.7. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a tempered function. Then the sequence* $\big(f(n) \bmod 1\big)_{n\in\mathbb{N}}$ *is uniformly distributed.*

The notion of uniform distribution in the torus $\mathbb{T}$ can naturally be extended to a much more general setup. For instance, one can consider sequences (functions) defined on more general domains and taking values in various metric spaces. In order to make a generalization of Definition 2.1, one needs an appropriate replacement for the discrete interval $\{1, \ldots, N\}$ in (2.1). Such replacement is provided by the notion of Følner sequence.

Definition 2.8. Let G be a locally compact Hausdorff group. A (left) *Følner sequence* is a sequence $(F_N)_{N\in\mathbb{N}}$ of compact positive measure subsets of G asymptotically invariant under left translations. More precisely:

$$\lim_{N\to\infty} \frac{\lambda\big(g F_N \cap F_N\big)}{\lambda(F_N)} = 1 \quad \forall g \in G \tag{2.4}$$

where $g F_N := \{gx : x \in F_N\}$ and λ is the (left) Haar measure on G.

Not every locally compact group has a Følner sequence. Groups admitting Følner sequences are called amenable. It is well known that any abelian group G is amenable. We can now define uniform distribution of a function $u : G \to K$ from an amenable group G to a compact group K.

Definition 2.9. Let G be a σ-compact locally compact amenable group with Haar measure λ, let $(F_N)_{N\in\mathbb{N}}$ be a Følner sequence in G, let K be a compact group with normalized Haar measure μ and let $u : G \to K$ be a continuous function. We say that u is *uniformly distributed* (with respect to (F_N)) if, for every open set $U \subset K$ with boundary of measure 0,

$$\lim_{N\to\infty} \frac{\lambda\big(\{t \in F_N : u(t) \in U\}\big)}{\lambda(F_N)} = \mu(U). \tag{2.5}$$

In later chapters we will use the following general form of Theorem 2.2.

Theorem 2.10. *Let* G *be a* σ-*compact locally compact amenable group with Haar measure* λ, *let* $(F_N)_{N\in\mathbb{N}}$ *be a Følner sequence in* G, *let* K *be a compact abelian group with Haar measure* μ *and let* $u : G \to K$ *be a measurable function. The following are equivalent:*

(1) *u is uniformly distributed with respect to* (F_N).
(2)

$$\lim_{N\to\infty} \frac{1}{\lambda(F_N)} \int_{F_N} f\big(u(t)\big)\, d\lambda(t) = \int_K f(x)\, d\mu(x) \quad \forall f \in C(K)$$

(3)

$$\lim_{N \to \infty} \frac{1}{\lambda(F_N)} \int_{F_N} \chi\big(u(t)\big) \, d\lambda(t) = 0 \quad \forall \chi \in \hat{K} \setminus \{0_{\hat{K}}\}$$

where $\hat{K}$ is the Pontryagin dual of K.

Sketch of the proof (*cf. [60, Exercise 1.1.1]*). The implication (2) $\Rightarrow$ (3) is trivial and the implication (3) $\Rightarrow$ (2) follows directly from the Stone–Weierstrass theorem (and the fact that characters separate points). To prove the implication (1) $\Rightarrow$ (2), let us assume that u is uniformly distributed. Condition (2.5) states that

$$\lim_{N \to \infty} \frac{1}{\lambda(F_N)} \int_{F_N} f\big(u(t)\big) \, d\lambda(t) = \int_K f(x) \, d\mu(x) \tag{2.6}$$

whenever f is the indicator function of an open set U with boundary of 0 measure. Since (2.6) is linear in f, it follows that (2.6) holds whenever f is a finite linear combination of indicator functions of open sets with boundary of 0 measure. Now let $f \in C(K)$. For any $\epsilon > 0$, one can find $f_1, f_2 : K \to \mathbb{R}$ which are finite linear combinations of indicator functions of open sets with boundary of 0 measure and satisfy $f_1 \leq f \leq f_2$ and $\int_K f_2 - f_1 \, d\mu < \epsilon$. It follows that (2.6) holds for continuous functions.

To prove the converse implication (2) $\Rightarrow$ (1), one can approximate the indicator function of any open set with boundary of 0 measure from above and from below by continuous functions, using Urysohn's lemma. One can then proceed as in the proof of (1) $\Rightarrow$ (2) above. $\quad\square$

Van der Corput's difference theorem can also be adapted to the following generality:

Theorem 2.11 (*General Form of DT*). *Let G be a σ-compact locally compact amenable group with Haar measure λ, let $(F_N)_{N \in \mathbb{N}}$ be a Følner sequence in G, let K be a compact abelian group with Haar measure μ and let $u : G \to K$ be a measurable function. Assume that for every $d \in G \setminus \{1_G\}$, the function $t \mapsto u(td) - u(t)$ is uniformly distributed in K. Then u is uniformly distributed.*

We will derive Theorem 2.11 from the following result involving functions taking values in a Hilbert space which can be interpreted as yet another form of DT (cf. the remark after Theorem 2.2 in [7]).

Theorem 2.12 (*DT for Hilbert Spaces*). *Let G be a σ-compact locally compact amenable group with Haar measure λ and let $u : G \to H$ be a continuous bounded map into a Hilbert space H. Let $(F_N)_{N \in \mathbb{N}}$ be a Følner sequence in G. Assume that*

$$\lim_{D \to \infty} \frac{1}{\lambda(F_D)} \int_{F_D} \limsup_{N \to \infty} \left| \frac{1}{\lambda(F_N)} \int_{F_N} \langle u(sh), u(s) \rangle \, d\lambda(s) \right| \, d\lambda(h) = 0.$$

Then we have

$$\lim_{N \to \infty} \left\| \frac{1}{\lambda(F_N)} \int_{F_N} u(s) \, d\lambda(s) \right\| = 0.$$

While it is possible to adapt the proof of Theorem 2.3 to this general setting we instead choose to employ ideas, going back to Bass [2] that involve properties of positive definite functions.

Proof. Define a sequence $(a_N)_{N \in \mathbb{N}}$ in H via the Bochner integral

$$a_N := \frac{1}{\lambda(F_N)} \int_{F_N} u(s) \, d\lambda(s).$$

Assume for the sake of a contradiction that the theorem is false. After passing, if needed, to a subsequence, we can assume that $a := \lim_{N \to \infty} \|a_N\|$ exists and is not 0. Define, for each $h \in G$

$$\gamma(h) := \lim_{N \to \infty} \frac{1}{\lambda(F_N)} \int_{F_N} \langle u(sh) - a_N, u(s) - a_N \rangle \, d\lambda(s). \tag{2.7}$$

Passing to a further subsequence we can assume that the limit in (2.7) exists for a countable dense subset of h. Since γ is continuous, it follows that γ can be defined for every $h \in G$. Using the Følner property (2.4) we can rewrite $\gamma(h)$ as

$$\gamma(h) = -a + \lim_{N \to \infty} \frac{1}{\lambda(F_N)} \int_{F_N} \langle u(sh), u(s) \rangle \, d\lambda(s). \tag{2.8}$$

Therefore

$$\lim_{D \to \infty} \frac{1}{\lambda(F_D)} \int_{h \in F_D} \gamma(h) \, d\lambda(h) = -a. \tag{2.9}$$

On the other hand, it follows directly from (2.7) that γ is a positive definite function. Indeed, let $f : G \to \mathbb{C}$ be a function with finite support. Then

$$\sum_{g,h \in G} f(g) \overline{f(h)} \gamma(h^{-1} g)$$

$$= \sum_{g,h \in G} f(g) \overline{f(h)} \lim_{N \to \infty} \frac{1}{\lambda(F_N)} \int_{F_N} \langle u(nh^{-1}g) - a_N, u(n) - a_N \rangle \, d\lambda(n)$$

$$= \sum_{g,h \in G} f(g) \overline{f(h)} \lim_{N \to \infty} \frac{1}{\lambda(F_N)} \int_{F_N} \langle u(mg) - a_N, u(mh) - a_N \rangle \, d\lambda(m)$$

$$= \lim_{N \to \infty} \frac{1}{\lambda(F_N)} \int_{F_N} \left\| \sum_{g \in G} f(g) \big(u(mg) - a_N \big) \right\|^2 \, d\mu(m)$$

$$\geq 0.$$

By the Naimark dilation theorem (see, for example, [54, Theorem 4.8]), there exists a unitary representation $(U_g)_{g \in G}$ of G in some Hilbert space V and some vector $v \in V$ such that $\gamma(h) = \langle U_h v, v \rangle$. Invoking the mean ergodic theorem (see, for instance, [44, Theorem 3.3]) we conclude that

$$-a = \lim_{D \to \infty} \frac{1}{\lambda(F_D)} \int_{h \in F_D} \gamma(h) \, d\lambda(h)$$

$$= \lim_{D \to \infty} \frac{1}{\lambda(F_D)} \int_{h \in F_D} \langle U_h v, v \rangle \, d\lambda(h) = \langle Pv, v \rangle = \|Pv\|^2$$

where $P : V \to V$ is the orthogonal projection onto the subspace of invariant functions. The last equation contradicts the hypothesis that $a > 0$. $\quad \square$

We can now derive Theorem 2.11 from Theorem 2.12:

Proof of Theorem 2.11. We will use the Weyl criterion, statement (3) in Theorem 2.10. Let $\chi \in \hat{K}$ be nontrivial. Let $v(s) = \chi\big(u(s)\big) \in \mathbb{C}$. Viewing $\mathbb{C}$ as a one dimensional Hilbert space we have, for every $s, h \in G$, $h \neq 1_G$

$$\langle v(sh), v(s) \rangle = \chi\big(u(sh)\big)\overline{\chi\big(u(s)\big)} = \chi\big(u(sh) - u(s)\big).$$

Since the map $s \mapsto u(sh) - u(s)$ is uniformly distributed, it follows from the Weyl criterion that

$$\lim_{N \to \infty} \frac{1}{\lambda(F_N)} \int_{F_N} \langle v(sh), v(s) \rangle \, d\lambda(s) = \lim_{N \to \infty} \frac{1}{\lambda(F_N)} \int_{F_N} \chi\big(u(sh) - u(s)\big) \, d\lambda(s) = 0$$

for any $h \in G \setminus \{1_G\}$. If we average the previous equation over all h we will also trivially get 0. Therefore the conditions of Theorem 2.12 are met and hence

$$0 = \lim_{N \to \infty} \left\| \frac{1}{\lambda(F_N)} \int_{F_N} v(s) \, d\lambda(s) \right\| = \lim_{N \to \infty} \left| \frac{1}{\lambda(F_N)} \int_{F_N} \chi\big(u(s)\big) \, d\lambda(s) \right|.$$

It follows now from Theorem 2.10 that u is uniformly distributed. □

3. Well distribution

A sequence $(u_n)_{n \in \mathbb{N}}$ taking values in the torus $\mathbb{T}$ is *well distributed* if for every $0 \leq a < b \leq 1$

$$\lim_{N \to \infty} \frac{\left| \{n \in \{M+1, \ldots, M+N\} : x_n \in [a, b)\} \right|}{N} = b - a \tag{3.1}$$

uniformly in M. An equivalent definition is that $(u_n)_{n \in \mathbb{N}}$ is uniformly distributed along any Følner sequence in $\mathbb{N}$. It follows from the Weyl criterion that (u_n) is well distributed if and only if for every $h \in \mathbb{N}$

$$\lim_{N \to \infty} \sup_{M \in \mathbb{N}} \left| \frac{1}{N} \sum_{n=M}^{M+N} e^{2\pi i h u_n} \right| = 0.$$

It is not hard to see that a trivial modification of the proof of Theorem 2.3 allows one to establish the following version of DT for well distribution (one can also derive it from Theorem 2.11).

Theorem 3.1. *Let $(x_n)_{n \in \mathbb{N}}$ be a sequence taking values in the torus $\mathbb{T}$. Assume that for every $d \in \mathbb{N}$, the sequence $n \mapsto x_{n+d} - x_n$ is well distributed. Then $(x_n)_{n \in \mathbb{N}}$ is well distributed.*

3.1. Well distribution and Hardy fields

It follows from Theorem 2.5 that the sequence $x_n = \sqrt{n} \mod 1$ is uniformly distributed. Since the differences between consecutive terms $\sqrt{n+1} - \sqrt{n}$ converge to 0, it is clear that $(x_n)_{n \in \mathbb{N}}$ is not well distributed. On the other hand, it is easy to see that the sequence $n\alpha \mod 1$ is well distributed for any irrational α. What about $(n\alpha + \sqrt{n}) \mod 1$? As it turns out, this sequence is well distributed, and a quick way to prove this is to utilize (a special case of) Theorem 2.12. More generally, we have the following:

Theorem 3.2. *Let $(v_n)_{n \in \mathbb{N}}$ be a non-decreasing sequence in $\mathbb{R}$ such that*

$$\lim_{N \to \infty} \sup_{M \in \mathbb{N}} \frac{v_{M+N} - v_M}{N} = 0. \tag{3.2}$$

Then for every $\alpha \notin \mathbb{Q}$ the sequence $u_n = n\alpha + v_n \mod 1$ is well distributed in $\mathbb{T}$.

Remark 3.3. (1) Theorem 3.2 extends Theorem 1.3.3 in [51], where, under somewhat stronger assumptions, it is concluded that u_n is uniformly distributed (by the way, Theorem 1.3.3 in [51] also appears in van der Corput's seminal paper [33]).

(2) Under the slightly stronger condition that $v_{N+1} - v_N \to 0$ as $N \to \infty$, it is actually true that $u_n = a_n + v_n$ is well distributed for any well distributed sequence a_n. This follows from Theorem 4.2.3 in [51].

Examples of sequences (v_n) satisfying (3.2) are $v_n = \sqrt{n}$, $v_n = \log n$ and $v_n = cn^a$ $\left(\cos(\log^b n) + d\right)$, where $a, b, c, d \in \mathbb{R}$ satisfy $a, c > 0$, $a, b < 1$, $d > 1$. More generally, if $f : \mathbb{R} \to \mathbb{R}$ is a non-decreasing smooth function such that $f'(x) \to 0$ as $x \to \infty$, then $v_n = f(n)$ also satisfies (3.2). Notice that $\log n$ is not even uniformly distributed in $\mathbb{T}$ but, in view of Theorem 3.2, $n\alpha + \log n$ is well distributed.

We need the following elementary lemma:

Lemma 3.4. *Let $(v_n)_{n \in \mathbb{N}}$ be a non-decreasing sequence in $\mathbb{R}$ satisfying (3.2). Then for every pair of sequences $(M_m), (N_m)$ of natural numbers such that $N_m - M_m \to \infty$ and every $d \neq 0$ we have*

$$\lim_{m \to \infty} \frac{1}{N_m - M_m} \sum_{n=M_m+1}^{N_m} \exp\big(2\pi i k(v_{n+d} - v_d)\big) = 1.$$

Proof. Observe that $\big|\exp(i\theta) - 1\big| \leq |\theta|$ for any $\theta \in \mathbb{R}$ (geometrically this says that a chord is shorter than the corresponding arc). Then

$$\left| \sum_{n=M_m+1}^{N_m} \exp\big(2\pi i k(v_{n+d} - v_d)\big) - 1 \right| \leq \sum_{n=M_m+1}^{N_m} \big|\exp\big(2\pi i k(v_{n+d} - v_d)\big) - 1\big|$$

$$\leq 2\pi k \sum_{n=M_m+1}^{N_m} \big|v_{n+d} - v_n\big|$$

$$= 2\pi k \sum_{i=1}^{d} \big(v_{N_m+i} - v_{M_m+i}\big).$$

Dividing by $N_m - M_m$ and using (3.2) d times we conclude that

$$\lim_{m \to \infty} \left| \frac{1}{N_m - M_m} \sum_{n=M_m+1}^{N_m} \exp\big(2\pi i k(v_{n+d} - v_d)\big) - 1 \right| = 0$$

as desired. $\square$

Proof of Theorem 3.2. We need to show that for all sequences $(M_m), (N_m)$ such that $N_m - M_m \to \infty$ and every $k \neq 0$ we have

$$\lim_{m \to \infty} \frac{1}{N_m - M_m} \sum_{n=M_m+1}^{N_m} \exp(2\pi i k u_n) = 0. \tag{3.3}$$

Let $w(n) = \exp(2\pi i k u_n)$. In order to prove (3.3), we will use Theorem 2.12 with the Hilbert space being $\mathbb{C}$, the group $G = \mathbb{Z}$, the Haar measure λ being simply the counting measure and the

Følner sequence being defined by $F_m = \{M_m + 1, \ldots, N_m\}$. We have

$$w(n + d)\overline{w(n)} = \exp\big(2\pi i k(u_{n+d} - u_n)\big) = \exp\big(2\pi i k(d\alpha + v_{n+d} - v_n)\big).$$

Taking Cesàro limits and appealing to Lemma 3.4 we obtain

$$\lim_{m\to\infty} \frac{1}{|F_m|} \sum_{n\in F_m} w(n + d)\overline{w(n)} = \exp(2\pi i k d\alpha) \lim_{m\to\infty} \frac{1}{|F_m|} \sum_{n\in F_m} \exp\big(2\pi i k(v_{n+d} - v_n)\big)$$

$$= \exp(2\pi i k d\alpha).$$

Taking the Cesàro limit in d we deduce that

$$\lim_{D\to\infty} \frac{1}{D} \sum_{d=1}^{D} \lim_{m\to\infty} \frac{1}{N_m - M_m} \sum_{n=M_m+1}^{N_m} w(n + d)\overline{w(n)} = 0.$$

Now (3.3) follows from Theorem 2.12 and this finishes the proof. □

Observe that if $(au_n + bv_n)_{n\in\mathbb{N}}$ is well distributed in $\mathbb{T}$ for every $(a, b) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$, then it follows from Weyl's criterion (Theorem 2.10) that the sequence $(u_n, v_n)_{n\in\mathbb{N}}$ is well distributed in $\mathbb{T}^2$. This observation should be juxtaposed with the following corollary of Theorem 3.2.

Corollary 3.5. *There exists a sequence $(u_n, v_n) \in \mathbb{T}^2$ which is not uniformly distributed but such that for any $a, b \in \mathbb{Z}$ with $a \neq 0$ we have that $au_n + bv_n$ is well distributed in $\mathbb{T}$.*

Proof. Let $u_n = n\alpha$ for some irrational $\alpha \in \mathbb{R}$, let $v_n = \log n$ and apply Theorem 3.2. □

Recall that $\Delta^s f$ denotes the sth iterated discrete derivative of f. Bootstrapping Theorem 3.2 with the help of DT we obtain the following more general result.

Corollary 3.6. *Let $s \in \mathbb{N}$ and let $p \in \mathbb{R}[x]$ be a polynomial with $\deg p = s$ whose leading coefficient is irrational. Let $f : \mathbb{N} \to \mathbb{R}$ be such that*

$$\lim_{N\to\infty} \sup_{M\in\mathbb{N}} \left| \frac{1}{N} \sum_{n=M}^{M+N} \Delta^s f(n) \right| = 0. \tag{3.4}$$

Then the sequence $\big(p(n) + f(n) \bmod 1\big)_{n\in\mathbb{N}}$ is well distributed.

Proof. We proceed by induction on s, the case $s = 1$ being Theorem 3.2. Next assume that $s > 1$ and that the result holds for $s - 1$. Let $u_n = p(n) + f(n) \bmod 1$ and let $d \in \mathbb{N}$ be arbitrary. The difference $u_{n+d} - u_n$ can be written as $\tilde{p}(n) + \tilde{f}(n)$, where $\tilde{p}(n) = p(n + d) - p(n)$ is a polynomial of degree $s - 1$ with leading coefficient irrational and $\tilde{f}(n) = \Delta f(n) + \Delta f(n + 1) + \cdots + \Delta f(n + d - 1)$ is a sequence satisfying (3.4) with $s - 1$ instead of s. By the induction hypothesis, $u_{n+d} - u_n$ is well distributed for every $d \in \mathbb{N}$. It follows from DT (in the form of Theorem 3.1) that u_n is well distributed as well. □

Observe that if $f : \mathbb{R} \to \mathbb{R}$ is a smooth function such that $f^{(s)}(x) \to 0$ as $x \to \infty$, then the sequence $f(n)$ satisfies (3.4).

We will now demonstrate the usefulness of Corollary 3.6 by providing a relatively short proof of a result of Boshernitzan pertaining to Hardy fields. First we need to introduce some relevant definitions. Two real valued functions f, g defined for all large enough $x \in \mathbb{R}$ are *equivalent* if there exists $M \in \mathbb{R}$ such that $f(x) = g(x)$ for all $x > M$. A *germ* is an equivalence class of functions under this equivalence relation.

Definition 3.7. A Hardy field is a collection of germs of smooth functions $f : \mathbb{R} \to \mathbb{R}$ that forms a field (with respect to the operations of pointwise addition and multiplication) and is closed under differentiation.

Observe that any function representing a non-zero germ in a Hardy field must be eventually non-zero (since it is invertible). It follows that any non-zero element of a Hardy field is eventually either positive or negative. Since for any Hardy field function f, the derivative f' belongs to the same Hardy field, we deduce that f is eventually monotone, and hence the limit $\lim_{x \to \infty} f(x)$ always exists in $[-\infty, \infty]$. Many familiar functions belong to some Hardy field, including any function obtained from e^x, $\log x$ and polynomial functions by composition and standard arithmetic operations. On the other hand, (non-constant) periodic functions do not belong to any Hardy field.

A function $f : \mathbb{R} \to \mathbb{R}$ is called *subpolynomial* if there exists $n \in \mathbb{N}$ such that $f(x)/x^n \to 0$ as $x \to \infty$.

Theorem 3.8 (*Boshernitzan, [28]*). *Assume* $f : \mathbb{R} \to \mathbb{R}$ *belongs to a Hardy field and is subpolynomial. Then the sequence* $(f(n) \bmod 1)_{n \in \mathbb{N}}$ *is uniformly distributed in* $\mathbb{T}$ *if and only if for every* $p(x) \in \mathbb{Q}[x]$ *we have*

$$\lim_{x \to \infty} \frac{\log x}{|f(x) - p(x)|} = 0. \tag{3.5}$$

Proof. First assume that

$$\lim_{x \to \infty} \frac{|f(x) - p(x)|}{\log x} = a < \infty$$

for some $p \in \mathbb{Q}[x]$. Then $f(x) = (a + g(x)) \log x + p(x)$ for some function $g(x)$ which approaches 0 as $x \to \infty$. Let M be a common multiple of the denominators of the coefficients of p. Then $Mf(n) = (Ma + g(n)) \log n \bmod 1$ for every $n \in \mathbb{N}$ grows too slowly to be uniformly distributed (see, for example, Theorem 1.2.6 in [51]).

We move now to the proof of the converse direction. Let f be a function satisfying the conditions of the theorem. Let $d = d(f)$ be the smallest integer such that $f(x)/x^d$ is bounded and let $L = \lim f(x)/x^d$. (The limit exists because f belongs to a Hardy field.)

We will prove the result by induction on d. If L is irrational, then the uniform distribution (and indeed well distribution) of $(f(n) \bmod 1)_{n \in \mathbb{N}}$ follows directly from either Theorem 3.2 or Corollary 3.6.

Observe that for any $p \in \mathbb{Q}[x]$, the sequence $(p(n) \bmod 1)_{n \in \mathbb{N}}$ is periodic. Hence $(f(n) + p(n) \bmod 1)_{n \in \mathbb{N}}$ is uniformly distributed if and only if $(f(n) \bmod 1)_{n \in \mathbb{N}}$ is. In particular, if $L \in \mathbb{Q} \setminus \{0\}$, then the uniform distribution of $(f(n) \bmod 1)_{n \in \mathbb{N}}$ is equivalent to that of $(g(n) \bmod 1)_{n \in \mathbb{N}}$, with $g(x) = f(x) - Lx^d$ satisfying the conditions of the theorem and either $d(g) < d(f)$ or $L(g) = 0$. Therefore it suffices to assume that $L = 0$.

Assume first that $d = 1$. Then L'Hôpital's rule implies that $f'(x) \to 0$ and, taking into account that $f(x)/\log(x) \to \infty$, we obtain that $xf'(x) \to \infty$ as $x \to \infty$. It now follows from Theorem 2.5 that $(f(n) \bmod 1)_{n \in \mathbb{N}}$ is uniformly distributed in $\mathbb{T}$.

Next we deal with the case $d = 2$ (a typical example is $f(x) = x \log x$). The reason why we cannot use induction here is that the derivative $f'(x)$ may not be uniformly distributed mod 1. By

L'Hôpital's rule we have that $f'(x) \to \infty$ and $f''(x) \to 0$. Therefore also $f'(x)/x^2 f''(x) \to 0$ and now the uniform distribution of $(f(n) \bmod 1)_{n \in \mathbb{N}}$ follows from Exercise 1.2.26 in [51].[2]

Finally, to deal with $d > 2$, we apply Theorem 2.3. Take an arbitrary $h \in \mathbb{N}$ and let $f_h(x) = f(x+h) - f(x)$. Since f is a subpolynomial function in a Hardy field, one can show that f_h is also a subpolynomial function in a Hardy field and that $f_h(x)/x^{d-1} \to 0$. Therefore $d(f_h) \le d - 1$. On the other hand, $f_h(x)/x \to \infty$, and in fact $|f_h(x) - p(x)|/x \to \infty$ for any $p \in \mathbb{Q}[x]$. This implies that $\log(x)/|f_h(x) - p(x)| \to 0$ and now the result follows by induction. $\quad\square$

Recently, Theorem 3.8 was extended to sequences of the form $f(p_n)$ where f is a subpolynomial Hardy field function and p_n is the nth prime:

Theorem 3.9 (*See [15, Theorem 3.1]*)**.** *Assume that $f : \mathbb{R} \to \mathbb{R}$ belongs to a Hardy field and is subpolynomial and let $(p_n)_{n \in \mathbb{N}}$ be the increasing sequence listing all the prime numbers. Then the sequence $(f(p_n) \bmod 1)_{n \in \mathbb{N}}$ is uniformly distributed in $\mathbb{T}$ if and only if for every $p(x) \in \mathbb{Q}[x]$ we have*

$$\lim_{x \to \infty} \frac{\log x}{|f(x) - p(x)|} = 0.$$

See also Corollary 4.10 for another result related to Theorem 3.8.

3.2. Cigler's characterization of well distribution and some of its corollaries

For completeness of the picture, we present in this subsection a result due to Cigler [31] that provides an explanation for the dynamical underpinnings of the notion of well distribution. We also present another result of Cigler [32] which states that for any tempered function f, the sequence $f(n)$ cannot be well distributed, complementing Corollary 2.7.

Theorem 3.10. *Let $x = (x_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{T}$. Consider x as an element of $\mathbb{T}^{\mathbb{N}}$ and let $S : \mathbb{T}^{\mathbb{N}} \to \mathbb{T}^{\mathbb{N}}$ be the shift map. Then $(x_n)_{n \in \mathbb{N}}$ is well distributed in $\mathbb{T}$ if and only if for any S-invariant measure μ on the orbit closure $\overline{\{S^n x : n \in \mathbb{N}\}}$ of x, the pushforward $\pi_* \mu$ by the projection $\pi : \mathbb{T}^{\mathbb{N}} \to \mathbb{T}$ onto the first coordinate is the Lebesgue measure λ on $\mathbb{T}$.*

Proof. Assume first that $(x_n)_{n \in \mathbb{N}}$ is well distributed and let μ be an invariant measure on $X := \overline{\{S^n x : n \in \mathbb{N}\}}$. Let $f \in C(\mathbb{T})$ be arbitrary. We need to show that $\int_X f \circ \pi \, d\mu = \int_{\mathbb{T}} f \, d\lambda$. For each $y \in X$, let $k_i(y)$ be a sequence in $\mathbb{N}$ such that $S^{k_i(y)} x \to y$ as $i \to \infty$. We have

$$\int_X f \circ \pi \, d\mu = \int_X f\big(\pi(y)\big) \, d\mu(y)$$

$$= \lim_{N \to \infty} \int_X \frac{1}{N} \sum_{n=1}^{N} f\big(\pi(S^n y)\big) \, d\mu(y)$$

$$= \lim_{N \to \infty} \int_X \lim_{i \to \infty} \frac{1}{N} \sum_{n=1}^{N} f\big(\pi(S^{n+k_i(y)} x)\big) \, d\mu(y)$$

---

[2] Curiously enough, this exercise is based on yet another result of van der Corput, Theorem 1.2.7 in [51].

$$= \lim_{N \to \infty} \int_X \lim_{i \to \infty} \frac{1}{N} \sum_{n=1+k_i(y)}^{N+k_i(y)} f\big(\pi(S^n x)\big) \, \mathrm{d}\mu(y)$$

$$= \lim_{N \to \infty} \int_X \lim_{i \to \infty} \frac{1}{N} \sum_{n=1+k_i(y)}^{N+k_i(y)} f(x_n) \, \mathrm{d}\mu(y).$$

Since $(x_n)_{n \in \mathbb{N}}$ is well distributed, for any $\epsilon > 0$, if N is large enough we have

$$\left| \frac{1}{N} \sum_{n=1+k_i(y)}^{N+k_i(y)} f\big(\pi(S^n x)\big) - \int_{\mathbb{T}} f(x) \, \mathrm{d}\lambda(x) \right| < \epsilon$$

regardless of $i \in \mathbb{N}$ and $y \in X$. Therefore, we deduce that $\int_X f \circ \pi \, \mathrm{d}\mu = \int_{\mathbb{T}} f(x) \, \mathrm{d}\lambda(x)$.

Now assume that for every invariant measure μ on X we have $\pi_* \mu = \lambda$. Let $f \in C(X)$ and assume, for the sake of a contradiction, that

$$A_N := \frac{1}{N} \sum_{n=a_N+1}^{a_N+N} f(x_n)$$

does not converge to $\int_{\mathbb{T}} f \, \mathrm{d}\lambda$ for some sequence $(a_N)_{N \in \mathbb{N}}$. Passing, if necessary, to a subsequence we can assume that $\lim_{k \to \infty} A_{N_k} = A$ exists and does not equal $\int_{\mathbb{T}} f \, \mathrm{d}\lambda$. Let

$$\mu_k = \frac{1}{N_k} \sum_{n=a_{N_k}+1}^{a_{N_k}+N_k} \delta_{S^n x}$$

be the corresponding average of point masses at the orbit of x. Passing to a further subsequence we can assume that μ_k converges as $k \to \infty$ to some S-invariant measure μ in the weak* topology. By the assumption, $\pi_* \mu = \lambda$ and therefore

$$\int_{\mathbb{T}} f \, \mathrm{d}\lambda = \int_X f\big(\pi(y)\big) \, \mathrm{d}\mu(y)$$

$$= \lim_{k \to \infty} \frac{1}{N_k} \sum_{n=a_{N_k}+1}^{a_{N_k}+N_k} f\big(\pi(S^n x)\big)$$

$$= \lim_{k \to \infty} \frac{1}{N_k} \sum_{n=a_{N_k}+1}^{a_{N_k}+N_k} f(x_n) = A.$$

This contradiction finishes the proof. $\quad \square$

Corollary 3.11. *For every tempered function $f : \mathbb{R} \to \mathbb{R}$, the sequence $(f(n) \bmod 1)_{n \in \mathbb{N}}$ in $\mathbb{T}$ is not well distributed.*

Proof. Let $x_n = f(n) \bmod 1$ and $x = (x_n)_{n \in \mathbb{N}}$ be the corresponding sequence in $\mathbb{T}^{\mathbb{N}}$. Let ℓ be the degree of f, so that $f^{(\ell)}$ goes to infinity but $f^{(\ell+1)}$ goes to 0. Then each of the functions $f, f', \ldots, f^{(\ell)}$ is tempered and, moreover, each non-zero linear combination of these functions (with integer coefficients) is tempered. In view Corollary 2.7 and Theorem 2.10, the tuple $\big(f(n), f'(n), \ldots, f^{(\ell)}(n)\big)$ is uniformly distributed in $\mathbb{T}^{\ell+1}$. In particular, that sequence is

dense in $\mathbb{T}^{\ell+1}$ so there exists some sequence $n_k \to \infty$ such that $f^{(i)}(n_k) \to 0$ as $k \to \infty$ for all $i = 0, \dots, \ell$.

For each $h \in \mathbb{N}$ we now get

$$f(n_k + h) = \sum_{i=0}^{\ell} \frac{f^{(i)}(n_k)}{i!} h^i + \frac{f^{(\ell+1)}(t)}{(\ell+1)!} h^{(\ell+1)}$$

for some $t \in [n_k, n_k + h]$. After reducing modulo 1 we deduce that $x_{n_k+h} \to 0$ as $k \to \infty$ for every $h \in \mathbb{N}$. Translating to a dynamical viewpoint, this means that the point $(0, 0, \dots)$ belongs to the orbit closure of x. Since this point is fixed under the shift map S, it supports a point mass which is an invariant measure. The projection under $\pi : \mathbb{T}^{\mathbb{N}} \to \mathbb{T}$ of this measure is then the point mass at 0 which is not the Lebesgue measure. It now follows from Theorem 3.10 that x is not well distributed. $\quad\square$

Corollary 3.12. *Almost every sequence in $\mathbb{T}^{\mathbb{N}}$ (with the product Lebesgue measure) is not well distributed.*

Proof. Notice that the shift S on $\mathbb{T}^{\mathbb{N}}$ is ergodic with respect to the product Lebesgue measure, and hence almost every $\omega \in \mathbb{T}^{\mathbb{N}}$ has a dense orbit (under S) in $\mathbb{T}^{\mathbb{N}}$. In particular, the orbit closure of ω contains the fixed point $(0, 0, \dots)$ which supports an invariant (atomic) measure δ whose projection under π is not the Lebesgue measure. $\quad\square$

4. Uniform distribution along Besicovitch almost periodic subsequences

The results obtained in this section heavily depend on ideas introduced in papers by Mendès France [52] and Daboussi and Mendès France [36]. In particular, we establish a refined version of DT involving Besicovitch almost periodic functions and derive some interesting new applications.

Definition 4.1. • For a subset $A \subset \mathbb{N}$, we define its upper density and Banach lower density, respectively, by

$$\bar{d}(A) = \limsup_{N \to \infty} \frac{|A \cap \{1, \dots, N\}|}{N}$$

$$d_*(A) = \lim_{N \to \infty} \inf_{M \in \mathbb{N}} \frac{|A \cap \{M+1, \dots, M+N\}|}{N}.$$

• A function $f : \mathbb{N} \to \mathbb{C}$ is *Besicovitch almost periodic* if for every $\epsilon > 0$ there exist $k \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_k \in \mathbb{T}$, $c_1, \dots, c_k \in \mathbb{C}$ such that

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| f(n) - \sum_{j=1}^{k} c_j e^{2\pi i \alpha_j n} \right| < \epsilon. \tag{4.1}$$

• A function $f : \mathbb{N} \to \mathbb{C}$ is *uniform Besicovitch almost periodic* if for every $\epsilon > 0$ there exist $k \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_k \in \mathbb{T}$, $c_1, \dots, c_k \in \mathbb{C}$ such that

$$\limsup_{N \to \infty} \sup_{M \in \mathbb{Z}} \frac{1}{N} \sum_{n=M+1}^{M+N} \left| f(n) - \sum_{j=1}^{k} c_j e^{2\pi i \alpha_j n} \right| < \epsilon.$$

- An increasing sequence $(n_k)_{k \in \mathbb{N}}$ in $\mathbb{N}$ is a *(uniform) Besicovitch sequence* if the indicator function 1_A of the set $A = \{n_k : k \in \mathbb{N}\}$ is (uniform) Besicovitch almost periodic and A has positive upper density (positive lower Banach density).

The following lemma provides a way to check that certain functions are Besicovitch almost periodic.

Lemma 4.2 (*[26], cf. also [3, Corollary 3.21]*). *A function $f : \mathbb{N} \to \mathbb{C}$ is (uniform) Besicovitch almost periodic if there exists a compact abelian group K, a point $a \in K$ and a Riemann integrable function $\phi : K \to \mathbb{C}$ such that $f(n) = \phi(na)$.*

In the next proposition we collect some examples of Besicovitch sequences.

Proposition 4.3. (1) *For any $\alpha > 0$, the sequence $k \mapsto \lfloor k\alpha \rfloor$ is a uniform Besicovitch sequence.*
(2) *The increasing sequence $(n_k)_{k \in \mathbb{N}}$ of squarefree numbers is a Besicovitch sequence.*
(3) *More generally, for any set $Q \subset \mathbb{N}$ such that $\sum_{m \in Q} 1/m < \infty$, the increasing sequence $(n_k)_{k \in \mathbb{N}}$ listing the numbers n without a divisor in Q is a Besicovitch sequence.*
(4) *For any irrational $\alpha \in \mathbb{R}$, the set $A := \{n \in \mathbb{N} : \gcd(n, \lfloor n\alpha \rfloor) = 1\}$ has a Besicovitch almost periodic indicator function, therefore the increasing enumeration of the elements of A is a Besicovitch sequence.*

Proof. (1) Let $\beta = 1/\alpha$ and let $A = \{\lfloor k\alpha \rfloor : k \in \mathbb{N}\}$. Observe that $\lfloor k\alpha \rfloor = n \iff n\beta \leq k < n\beta + \beta$ and therefore $n \in A \iff n\beta \mod 1 \in (1 - \beta, 1]$. Let $\phi : \mathbb{T} \to \mathbb{C}$ be the indicator function of the set $(1 - \beta, 1]$. It follows from Lemma 4.2 that the sequence $k \mapsto \lfloor k\alpha \rfloor$ is uniform Besicovitch.
(3) For each $M \in \mathbb{N}$, let Q_M be the set of natural numbers not divisible by an element of $Q \cap [1, M]$. The indicator function 1_{Q_M} is periodic and hence can be written as a finite sum $\sum_{j=1}^{k} c_j e^{2\pi i \alpha_j n}$ (with all the $\alpha_j \in \mathbb{Q}$).
Let $\epsilon > 0$ and choose M so that $\sum_{m > M} 1_Q(m)/m < \epsilon$. It is clear that $Q \subset Q_M$. Moreover any number in $Q_M \setminus Q$ must be divisible by some element of $Q \cap [M + 1, \infty)$. Therefore, for each $N \in \mathbb{N}$, the cardinality of $\{1, \ldots, N\} \cap (Q_M \setminus Q)$ is at most $\sum_{m=M+1}^{\infty} 1_Q(m)N/m < N\epsilon$. We conclude that

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| 1_Q(n) - \sum_{j=1}^{k} c_j e^{2\pi i \alpha_j n} \right| < \epsilon.$$

(4) This is the content of the paper [59]. $\square$

We have the following result (cf. [52,36,55]) which can be viewed as a refinement of Theorem 2.3.

Theorem 4.4. *Let $(u_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{T}$. If for every $h \in \mathbb{N}$ the difference $n \mapsto u_{n+h} - u_n$ is uniformly distributed in $\mathbb{T}$ then for any Besicovitch sequence $(n_k)_{k \in \mathbb{N}}$ the sequence $(u_{n_k})_{k \in \mathbb{N}}$ is uniformly distributed. Moreover, if for every $h \in \mathbb{N}$ the sequence $(u_{n+h} - u_n)_{n \in \mathbb{N}}$ is well distributed and $(n_k)_{k \in \mathbb{N}}$ is uniform Besicovitch, then $(u_{n_k})_{k \in \mathbb{N}}$ is well distributed.*

We will actually prove a multidimensional version of Theorem 4.4. First, we will need the following multidimensional version of Definition 4.1.

Definition 4.5. Let $d \in \mathbb{N}$.

- For a set $A \subset \mathbb{N}^d$ we define the *upper density* by the formula

$$\bar{d}(A) = \limsup_{N \to \infty} \frac{\left| A \cap \{1, \ldots, N\}^d \right|}{N^d}$$

and the *lower Banach density* by the formula

$$d_*(A) = \liminf_{N \to \infty} \inf_{(M_1, \ldots, M_d) \in \mathbb{N}^d} \frac{\left| A \cap \prod_{i=1}^{d} \{M_i + 1, \ldots, M_i + N\} \right|}{N^d}.$$

- A function $f : \mathbb{N}^d \to \mathbb{C}$ is *Besicovitch almost periodic* if for any $\epsilon > 0$ there exists $\ell \in \mathbb{N}$ and $\alpha_1, \ldots, \alpha_\ell \in \mathbb{T}^d$, $c_1, \ldots, c_\ell \in \mathbb{C}$ such that

$$\limsup_{N \to \infty} \frac{1}{N^d} \sum_{n \in [1,N]^d} \left| f(n) - \sum_{j=1}^{\ell} c_j \exp 2\pi i \langle \alpha_j, n \rangle \right| < \epsilon.$$

- We say that $f : \mathbb{N}^d \to \mathbb{C}$ is *uniform Besicovitch almost periodic* if one can replace the above limit by its uniform version:

$$\limsup_{N \to \infty} \sup_{m \in \mathbb{N}^d} \frac{1}{N^d} \sum_{n \in m+[1,N]^d} \left| f(n) - \sum_{j=1}^{\ell} c_j \exp 2\pi i \langle \alpha_j, n \rangle \right| < \epsilon$$

where $[1, N] = \{1, \ldots, N\}$.
- A sequence $n : \mathbb{N}^d \to \mathbb{N}^d$ is *increasing* if whenever $a, b \in \mathbb{N}^d$ and for some coordinate $j \in \{1, \ldots, d\}$, one has $a_j \leq b_j$, then also $n(a)_j \leq n(b)_j$.
- An increasing sequence $n : \mathbb{N}^d \to \mathbb{N}^d$ is a *(uniform) Besicovitch sequence* if the indicator function 1_A of the set $A := \{n(k) : k \in \mathbb{N}^d\}$ is (uniform) Besicovitch almost periodic and A has positive upper density (positive lower Banach density).

Theorem 4.6. *Let $d \in \mathbb{N}$, let K be a compact abelian group and let $u : \mathbb{N}^d \to K$ be such that for any $h \in \mathbb{N}^d$ the map $n \mapsto u(n + h) - u(n)$ is uniformly distributed in K. Let $n : \mathbb{N}^d \to \mathbb{N}^d$ be an increasing Besicovitch sequence. Then the sequence $k \mapsto u\big(n(k)\big)$ is uniformly distributed in K.*

Moreover, if the maps $n \mapsto u(n+h) - u(n)$ are well distributed and n is uniform Besicovitch, then $k \mapsto u\big(n(k)\big)$ is well distributed in K.

Proof. Let $\chi \in \hat{K} \setminus \{0_{\hat{K}}\}$ be an arbitrary non-constant character on the Pontryagin dual $\hat{K}$ of the compact abelian group K. We need to show that

$$\lim_{N \to \infty} \frac{1}{N^d} \sum_{k \in m+[1,N]^d} \chi\Big(u\big(n(k)\big)\Big) = 0$$

uniformly in $m \in \mathbb{N}^d$. Let $A = \{n(k) : k \in \mathbb{N}^d\}$ and let 1_A be the indicator function of A. Also, for each $m \in \mathbb{N}^d$ and $N \in \mathbb{N}$, let $I(m, N) = m + [1, N]^d = \prod_i [m_i, m_i + N]$ and let $\tilde{I}(m, N) = \prod_i [n(m_i), n(m_i + N)]$. Observe that $n^{-1}\big(\tilde{I}(m, N)\big) = I(m, N)$ because n is increasing, and therefore $n\big(I(m, N)\big) = A \cap \tilde{I}(m, N)$. Hence

$$\liminf_{N \to \infty} \frac{N^d}{|\tilde{I}(m, N)|} = \liminf_{N \to \infty} \frac{|A \cap \tilde{I}(m, N)|}{|\tilde{I}(m, N)|} \geq d_*(A) > 0.$$

Moreover

$$\frac{1}{N^d} \sum_{k \in I(m,N)} \chi\big(u\big(n(k)\big)\big) = \frac{1}{N^d} \sum_{n \in \tilde{I}(m,N)} 1_A(n) \chi\big(u(n)\big).$$

Fix $\epsilon > 0$ and let $\alpha_1, \dots, \alpha_s \in \mathbb{T}^d$ and $c_1, \dots, c_s \in \mathbb{C}$ such that

$$\limsup_{N \to \infty} \frac{1}{|\tilde{I}(m, N)|} \sum_{n \in \tilde{I}(m,N)} \left| 1_A(n) - \sum_{j=1}^{s} c_j \exp(2\pi i \langle \alpha_j, n \rangle) \right| < \epsilon.$$

We deduce that

$$\limsup_{N \to \infty} \frac{1}{N^d} \left| \sum_{n \in \tilde{I}(m,N)} 1_A(n) \chi\big(u(n)\big) \right|$$

$$= \limsup_{N \to \infty} \frac{1}{N^d} \left| \sum_{n \in \tilde{I}(m,N)} \left(1_A(n) - \sum_{j=1}^{s} c_j \exp(2\pi i \langle \alpha_j, n \rangle) \right. \right.$$

$$\left. \left. + \sum_{j=1}^{s} c_j \exp(2\pi i \langle \alpha_j, n \rangle) \right) \chi\big(u(n)\big) \right|$$

$$< \epsilon \limsup_{N \to \infty} \frac{|\tilde{I}(m, N)|}{N^d} + \limsup_{N \to \infty} \frac{1}{N^d} \left| \sum_{n \in \tilde{I}(m,N)} \sum_{j=1}^{s} c_j \exp(2\pi i \langle \alpha_j, n \rangle) \cdot \chi\big(u(n)\big) \right|$$

$$\leq \frac{\epsilon}{d_*(A)} + \sum_{j=1}^{s} c_j \frac{1}{d_*(A)} \limsup_{N \to \infty} \left| \frac{1}{|\tilde{I}(m, N)|} \sum_{n \in \tilde{I}(m,N)} \exp(2\pi i \langle \alpha_j, n \rangle) \cdot \chi\big(u(n)\big) \right|.$$

It suffices now to show that the uniform Cesàro limit of the sequence $v : n \mapsto \exp\big(2\pi i \langle \alpha_j, n \rangle\big) \cdot \chi\big(u(n)\big)$ is 0. To show this we will use Theorem 3.1. We have

$$v(n + h)\overline{v(n)} = \exp\big(2\pi i \langle \alpha_j, h \rangle\big) \cdot \chi\big(u(n + h) - u(n)\big).$$

Therefore, averaging over n, with h fixed we obtain

$$\frac{1}{|\tilde{I}(m, N)|} \sum_{n \in \tilde{I}(m,N)} v(n + h)\overline{v(n)}$$

$$= \exp\big(2\pi i \langle \alpha_j, h \rangle\big) \frac{1}{|\tilde{I}(m, N)|} \sum_{n \in \tilde{I}(m,N)} \chi\big(u(n + h) - u(n)\big). \tag{4.2}$$

It follows from Weyl's criterion, together with the hypothesis that $u(n + h) - u(n)$ is well distributed, that the limit as $N \to \infty$ of the quantity in (4.2) is 0. It now follows from DT (in the form of Theorem 3.1) that the uniform Cesàro limit of $v(n)$ is 0 and we are done. $\quad \square$

The argument used in part (1) of Proposition 4.3 can be utilized to show the following:

Proposition 4.7. *Let* $\alpha, \beta \in [1, \infty)$. *Then the sequence* $n : (k_1, k_2) \mapsto (\lfloor k_1 \alpha \rfloor, \lfloor k_2 \beta \rfloor)$ *is uniform Besicovitch.*

Putting Proposition 4.7 together with Theorem 4.6, we conclude that, for instance, the sequence $(n, m) \mapsto (\alpha \lfloor n\beta \rfloor^3, \sqrt{\lfloor m\gamma \rfloor})$ is uniformly distributed in $\mathbb{T}^2$ for any irrational $\alpha \in \mathbb{R}$ and $\beta, \gamma \geq 1$.

Observe that Theorem 4.4 does not remain true if one only requires that u_n is itself uniformly distributed. For instance, for any irrational $\alpha \in (0, 1)$, the sequence $n \mapsto n\alpha \mod 1$ is uniformly distributed, and, in view of Proposition 4.3, the increasing sequence $k \mapsto \lfloor k\alpha \rfloor$ is a Besicovitch sequence. However, it is not hard to show that $\lfloor k/\alpha \rfloor \alpha \mod 1$ is not uniformly distributed (indeed, this sequence always lands in the interval $(1 - \alpha, 1)$). On the other hand, it can be proved that $\lfloor k\alpha \rfloor \beta \mod 1$ *is* uniformly distributed whenever $\alpha\beta$ and β are irrational (cf. [51, Theorem 5.1.8]).

By looking more closely at the proof of Theorem 4.4, one notices that the condition that $u_{n+h} - u_n$ is uniformly distributed for every $h \in \mathbb{N}$ can be replaced with the (weaker) condition that $u_n + n\alpha$ is uniformly distributed for every $\alpha \in \mathbb{T}$. More precisely, we only need $u_n + n\alpha$ to be uniformly distributed for the frequencies $\alpha = \alpha_j$ that appear in the periodic approximation (4.1) of the (indicator function of the) Besicovitch sequence (n_k). This improvement of Theorem 4.4 was obtained by Daboussi and Mendès France in [36] (the authors called the set of such frequencies the *spectrum* of (u_n)).

Of particular interest is the case when all the frequencies needed to approximate the (indicator function of the) Besicovitch sequence (n_k) are rational:

Definition 4.8. • A function $f : \mathbb{N} \to \mathbb{C}$ is *rational Besicovitch almost periodic* if for every $\epsilon > 0$ there exist $k \in \mathbb{N}$ and $\alpha_1, \ldots, \alpha_k \in \mathbb{Q}$, $c_1, \ldots, c_k \in \mathbb{C}$ such that

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| f(n) - \sum_{j=1}^{k} c_j e^{2\pi i \alpha_j n} \right| < \epsilon.$$

• An increasing sequence $(n_k)_{k \in \mathbb{N}}$ in $\mathbb{N}$ is a *rational Besicovitch sequence* if the indicator function 1_A of the set $A = \{n_k : k \in \mathbb{N}\}$ is rational Besicovitch almost periodic and A has positive upper density.

It follows from (the proof of) Proposition 4.3 that the sequence $(n_k)_{k \in \mathbb{N}}$ enumerating the squarefree numbers is rational Besicovitch.

Theorem 4.9. *Let* $u : \mathbb{N} \mapsto \mathbb{T}$ *and assume that for every* $\alpha \in \mathbb{Q}$, *the sequence* $n \mapsto u(n) + n\alpha$ *is uniformly distributed. Let* $(n_k)_{k \in \mathbb{N}}$ *be a rational Besicovitch sequence. Then the sequence* $k \mapsto u(n_k)$ *is uniformly distributed in* $\mathbb{T}$.

Theorem 4.9 immediately implies that if (n_k) is the increasing enumeration of the set of squarefree numbers then, for any irrational α, the sequence $k \mapsto n_k \alpha \mod 1$ is uniformly distributed. A more sophisticated application involves Boshernitzan's theorem (Theorem 3.8). It is easy to see that a function $f : \mathbb{R} \to \mathbb{R}$ satisfies (3.5) for every $p \in \mathbb{Q}[x]$ if and only if $f(x) + x\alpha$ does, for a rational α. Therefore, we obtain the following corollary of Theorem 3.8:

Corollary 4.10. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be subpolynomial and a member of a Hardy field. Let* (n_k) *be a rational Besicovitch sequence. Assume that condition* (3.5) *holds for every* $p \in \mathbb{Q}[x]$. *Then the sequence* $\left(f(n_k) \mod 1 \right)_{k \in \mathbb{N}}$ *is uniformly distributed.*

5. Some additional applications of DT

In this section we demonstrate the versatility of DT by providing three more applications of Theorem 2.12.

5.1. A Weyl type theorem in $\mathbb{A}/\mathbb{Q}$

In this subsection, with the help of (a special case of) Theorem 2.11, we establish an "adelic" version of Weyl's theorem on equidistribution of polynomials.

Fix a prime p. Given $x, y \in \mathbb{Q}$, write $x - y = p^n \cdot \frac{a}{b}$ where $a, b, n \in \mathbb{Z}$, $b > 0$ and both a and b are coprime with p. Let $d_p(x, y) = p^{-n}$. It is not hard to check that d_p is a metric on $\mathbb{Q}$. We denote by $\mathbb{Q}_p$ the completion of $\mathbb{Q}$ with respect to this metric and extend d_p to $\mathbb{Q}_p$. A more concrete description of $\mathbb{Q}_p$ is

$$\mathbb{Q}_p = \left\{ \sum_{i=N}^{\infty} x_i p^i : x_i \in \{0, \ldots, p-1\}; \ N \in \mathbb{Z} \right\}.$$

We let $\mathbb{Z}_p \subset \mathbb{Q}_p$ be the set $\{n \in \mathbb{Q}_p : d_p(n, 0) \leq 1\}$. Equivalently, $\mathbb{Z}_p$ consists of sums $\sum_{i=N}^{\infty} x_i p^i$ where $N \geq 0$. Observe that $\mathbb{Z} \subset \mathbb{Z}_p$.

Next, let $\mathbb{A}$ denote the additive group of the adeles:

$$\mathbb{A} = \left\{ (a_\infty, a_2, a_3, a_5, \ldots) \in \left(\mathbb{R} \times \prod_{p \ \text{prime}} \mathbb{Q}_p \right) : a_p \in \mathbb{Z}_p \text{ for all but finitely many } p \right\}.$$

If $x \in \mathbb{Q}$ and p is a prime which does not divide the denominator of x then $x \in \mathbb{Z}_p$; in particular, $x \in \mathbb{Z}_p$ for all but finitely many p. Therefore, one can embed $\mathbb{Q}$ in $\mathbb{A}$ by identifying $x \in \mathbb{Q}$ with $(x, x, x, \ldots) \in \mathbb{A}$. Given $x = \sum_{i=N}^{\infty} x_i p^i \in \mathbb{Q}_p$ with $x_i \in \{0, \ldots, p-1\}$ we denote by

$$f_p(x) = \sum_{i=N}^{-1} x_i p^i \tag{5.1}$$

the p-fractional part of x (with the convention that $f_p(x) = 0$ when $N > -1$). Therefore $x - f_p(x) \in \mathbb{Z}_p$ for any $x \in \mathbb{Q}_p$. Observe that $f_p(x)$ is in $\mathbb{Q}$ for any $x \in \mathbb{Q}_p$ since (5.1) is a finite sum. Moreover, the denominator of $f_p(x)$ is a power of p; it follows that $f_p(x) \in \mathbb{Z}_q$ for any $p \neq q$ and $x \in \mathbb{Q}_p$. Given $x \in \mathbb{A}$, the sum $\tilde{f}(x) = \sum_p f_p(x_p) \in \mathbb{Q}$ has only finitely many non-zero terms and hence is well defined. Note that each p-adic coordinate of $x - \tilde{f}(x)$ is in $\mathbb{Z}_p$ and therefore $x - \tilde{f}(x) \in \mathbb{R} \times \prod_p \mathbb{Z}_p$. Finally, let

$$f(x) = \left(\tilde{f}(x) + \lfloor x_\infty - \tilde{f}(x) \rfloor \right) \in \mathbb{Q} \tag{5.2}$$

and observe that $x - f(x) \in [0, 1) \times \prod_p \mathbb{Z}_p$.

The topology on $\mathbb{A}$ is generated by a basis of open sets of the form

$$U_\infty \times \prod_{p \in F} U_p \times \prod_{p \notin F} \mathbb{Z}_p$$

where $U_\infty \subset \mathbb{R}$ is open, F is a finite set of primes and $U_p \subset \mathbb{Q}_p$ is an open set.

The following lemma is well known; we give a proof for the sake of completeness.

Lemma 5.1. *The subgroup $\mathbb{Q} \subset \mathbb{A}$ is discrete and co-compact.*

Proof. Let $x \in \mathbb{Q}$ be an arbitrary rational number. On the one hand, observe that the set

$$U_x = (x - 1, x + 1) \times \prod_p (x + \mathbb{Z}_p) \subset \mathbb{A}$$

is open, because $x \in \mathbb{Z}_p$ for all but finitely many primes p (those which divide the denominator of x). On the other hand, for any rational $y \in U_x$, the difference $y - x$ belongs to every $\mathbb{Z}_p$ and hence to $\mathbb{Z}$. However, that difference is also in $(-1, 1)$, and so it must be 0. We just showed that x is the only rational in U_x, and because $x \in \mathbb{Q}$ was arbitrary it follows that $\mathbb{Q}$ is a discrete subset of $\mathbb{A}$.

Next we focus on the quotient group $K := \mathbb{A}/\mathbb{Q}$. Let $\tilde{K} := [0, 1) \times \prod_p \mathbb{Z}_p \subset \mathbb{A}$. We define a group structure in $\tilde{K}$ as follows: given $x, y \in \tilde{K}$, let $x + y := (x + y) - f(x + y)$, where the sums and the subtraction on the right are the (usual) pointwise operations in $\mathbb{A}$, and $f : \mathbb{A} \to \mathbb{Q}$ is defined by (5.2). Note that this group operation is different from the pointwise addition.

We claim that K is isomorphic to $\tilde{K}$. To show this, let $\phi : \mathbb{A} \to \tilde{K}$ be the map $\phi(x) = x - f(x) \in \tilde{K} \subset \mathbb{A}$. First we show that ϕ is a group homomorphism. Indeed, for any $x, y \in \mathbb{A}$, the difference (in $\mathbb{A}$) between $\phi(x + y)$ and $\phi(x) + \phi(y)$ is a rational number which belongs to the set $\tilde{K} - \tilde{K} = (-1, 1) \times \prod_p \mathbb{Z}_p$, and 0 is the only such number.

Next, observe that ϕ is surjective. Indeed, for $x \in \tilde{K} \subset \mathbb{A}$ we have $f(x) = 0$ and hence $\phi(x) = x$. To see that the kernel of ϕ is $\mathbb{Q}$, notice that $f(x) \in \mathbb{Q}$ for any $x \in \mathbb{A}$, and so, on the one hand, $x = f(x) \Rightarrow x \in \mathbb{Q}$, and on the other hand, $x \in \mathbb{Q} \Rightarrow \phi(x) \in \mathbb{Q} \cap \tilde{K} = \{0\}$.

Finally, observe that ϕ is not quite a homeomorphism, because the open set $[0, 1/2) \times \prod_p \mathbb{Z}_p \subset \tilde{K}$ cannot be written as $\phi(U)$ where $U \subset \mathbb{A}$ is an open set invariant under $\mathbb{Q}$. In fact, if we endow $K \cong \tilde{K}$ with the topology consisting of the open sets $\phi(U)$ with $U \subset \mathbb{A}$ open and invariant under $\mathbb{Q}$, we see that the topology on K is the same as $\mathbb{T} \times \prod_p \mathbb{Z}_p$ (but keep in mind that the group structure in K is *not* the product (or pointwise) structure). $\quad\square$

A function $g : \mathbb{Q} \to K$ is a *polynomial* if it takes the form $g(x) = \sum_{j=0}^m \alpha_j x^j$ for some $\alpha_j \in \mathbb{A}$, $j = 0, \ldots, m$. We say that the polynomial g *has an irrational coefficient* if for some $j > 0$ one has $\alpha_j \in \mathbb{A} \setminus \mathbb{Q}$.

Theorem 5.2. *Let $g_1, \ldots, g_d : \mathbb{Q} \to K$ be polynomials. If for any $(n_1, \ldots, n_d) \in \mathbb{Q}^d \setminus \{0\}$ the sum $\sum n_j g_j : \mathbb{Q} \to K$ is not a constant map, then the sequence $(g_1, \ldots, g_d) : \mathbb{Q} \to K^d$ is well distributed.*

In particular, this sequence has a dense image; a special case of this fact was established in [34] using different methods. A related result, concerning a kind of uniform distribution of polynomial sequences in $\mathbb{A}/\mathbb{Q}$ and more general settings, was obtained by D. Cantor in [29]. Cantor used significantly more sophisticated machinery from number theory.

Our proof of Theorem 5.2 is very similar to the classical proof of (multidimensional version of) Weyl's theorem using DT and rests on the Weyl criterion, Theorem 2.10. We state in the following proposition the precise version of Weyl's criterion we will need.

Proposition 5.3. *A sequence $f : \mathbb{Q} \to K^d$ is well distributed if and only if for any Følner sequence $(F_N)_{N \in \mathbb{N}}$ in $\mathbb{Q}$ and any non-constant character $\chi : K^d \to \mathbb{C}$ we have:*

$$\lim_{N \to \infty} \frac{1}{|F_N|} \sum_{u \in F_N} \chi\big(f(x)\big) = 0.$$

Proof. Apply Theorem 2.10 with $G = \mathbb{Q}$ with the discrete topology, λ being the counting measure on $\mathbb{Q}$ and the compact group (K in the statement of Theorem 5.2) being K^d. $\quad\square$

The next step is to give a description of the characters of K^d. It is a classical fact that the Pontryagin dual of K is $\mathbb{Q}$ (with the discrete topology). More precisely, with any point $r \in \mathbb{Q}$ we associate the character $\chi_r : K \to \mathbb{C}$ described by

$$\chi_r : u \mapsto \exp\left(2\pi i\left(\tilde{f}(ru) - nu_\infty\right)\right).$$

More generally, the Pontryagin dual of K^d is $\mathbb{Q}^d$ and, for $r = (r_1, \ldots, r_d) \in \mathbb{Q}^d$ we define the character $\chi_r : K^d \to \mathbb{C}$ by the formula

$$\chi_r : u = (u_1, \ldots, u_d) \mapsto \prod_{j=1}^{d} \chi_{r_j}(u_j).$$

Observe that, if we denote χ_1 by simply χ, then $\chi_r(u) = \chi(ru)$ for any $r \in \mathbb{Q}$ and $u \in \mathbb{A}$.

Proof of Theorem 5.2. Let $r \in \mathbb{Q}^d \setminus \{0\}$ be arbitrary and let $(F_N)_{N \in \mathbb{N}}$ be an arbitrary Følner sequence for the additive group $(\mathbb{Q}, +)$. In view of Proposition 5.3, we need to show that

$$\lim_{N \to \infty} \frac{1}{|F_N|} \sum_{x \in F_N} \chi_r\big(g_1(x), \ldots, g_d(x)\big) = 0. \tag{5.3}$$

Expanding the definition of χ_r we get

$$\chi_r\big(g_1(x), \ldots, g_d(x)\big) = \prod_{j=1}^{d} \chi_{r_j}\big(g_j(x)\big) = \chi\left(\sum_{j=1}^{d} r_j g_j(x)\right). \tag{5.4}$$

Without loss of generality, we may (and will) assume that $g_i(0) = 0$. We proceed by induction on the highest degree of the polynomials $g_1, \ldots, g_d$. Assume first that the highest degree is 1, so that $g_j(x) = \alpha_j x$ where all $\alpha_j \in \mathbb{A}$ and, for every $r = (r_1, \ldots, r_d) \in \mathbb{Q}^d \setminus \{0\}$, we have $\alpha := \sum_{j=1}^{d} r_j \alpha_j \notin \mathbb{Q}$.

In view of (5.3) and (5.4) it suffices to show that

$$\lim_{N \to \infty} \frac{1}{|F_N|} \sum_{x \in F_N} \chi(\alpha x) = 0$$

for any $\alpha \in \mathbb{A} \setminus \mathbb{Q}$ and any additive Følner sequence $(F_N)_{N \in \mathbb{N}}$ in $\mathbb{Q}$. Since $\alpha \in \mathbb{A} \setminus \mathbb{Q}$, the projection $\alpha + \mathbb{Q} \in K$ is nonzero. Therefore there exists some character of K which does not vanish at α. Since K is the Pontryagin dual of $\mathbb{Q}$, this means that $\chi(\alpha y) \neq 1$ for some $y \in \mathbb{Q}$.

Next take $\epsilon > 0$ and let $N \in \mathbb{N}$ be such that

$$\frac{\left|(F_N + y) \bigtriangleup F_N\right|}{|F_N|} \leq \left|1 - \chi(\alpha y)\right| \epsilon.$$

Then, on the one hand, we have

$$\left|\frac{1}{|F_N|} \sum_{x \in F_N} \chi(\alpha x) - \frac{1}{|F_N|} \sum_{x \in F_N + y} \chi(\alpha x)\right| \leq \frac{\left|(F_N + y) \bigtriangleup F_N\right|}{|F_N|} < \left|1 - \chi(\alpha y)\right| \epsilon,$$

and, on the other hand, we have

$$\frac{1}{|F_N|} \sum_{x \in F_N + y} \chi(\alpha x) = \frac{1}{|F_N|} \sum_{x \in F_N} \chi\big(\alpha(x + y)\big) = \chi(\alpha y) \frac{1}{|F_N|} \sum_{x \in F_N} \chi(\alpha x).$$

Putting things together we conclude that

$$\left| \left(1 - \chi(\alpha y)\right) \frac{1}{|F_N|} \sum_{x \in F_N} \chi(\alpha x) \right| \leq \left| 1 - \chi(\alpha y) \right| \epsilon$$

which finally implies that

$$\left| \frac{1}{|F_N|} \sum_{x \in F_N} \chi(\alpha x) \right| < \epsilon.$$

This finishes the proof of the theorem in the case when all the polynomials $g_1, \ldots, g_d$ have degree 1.

Next assume that $s := \max\{\deg g_1, \ldots, \deg g_d\} > 1$ and that the result holds for $s - 1$. We will apply DT, Theorem 2.11, with the role of G being played by $\mathbb{Q}$ endowed with the discrete topology and the counting Haar measure. For any $h \in \mathbb{Q}$ and $j \in \{1, \ldots, d\}$, the difference $x \mapsto g_j(x + h) - g_j(x)$ is a polynomial of degree at most $s - 1$. It follows from the induction hypothesis that

$$g_h(x) := g(x + h) - g(x) = \left(g_1(x + h) - g_1(x), \ldots, g_d(x + h) - g_d(x)\right)$$

is uniformly distributed in K^d for any $h \in \mathbb{Q} \setminus \{0\}$. The result now follows from Theorem 2.11. $\quad\square$

The technique employed in this section works in more general settings. For instance, one can obtain a similar theorem for polynomials $p : K \to \mathbb{A}_K / K$, where K is a number field and $\mathbb{A}_K$ is the group of adeles over K. These results also extend to the case of polynomials with several variables. In order to keep the exposition short, we do not pursue these possibilities any further.

5.2. An application to the $\{x + y, xy\}$ problem

A long standing open question in Ramsey theory asks whether for any finite partition $\mathbb{N} = C_1 \cup \cdots \cup C_r$ of the natural numbers, there exists a cell C_i of the partition and a pair $(x, y) \in \mathbb{N}^2$ such that $\{x, y, x + y, xy\} \subset C_i$. The answer is not known even if one only asks for $\{x + y, xy\} \subset C_i$ and $(x, y) \neq (2, 2)$. In [23] a similar question pertaining to partitions of fields was answered positively. In particular, the following result was proved:

Theorem 5.4. *For every infinite field K and every finite partition $K = C_1 \cup \cdots \cup C_r$ there exists a cell C_i and infinitely many $x, y \in K$ such that $\{x, x + y, xy\} \subset C_i$.*

The proof of Theorem 5.4 in [23] is based on ergodic-theoretic techniques. At its core lies a new ergodic theorem, whose proof hinges on a new application of the following form of DT.

Proposition 5.5. *Let H be a Hilbert space, let K be a countable field, let $(F_N)_{N \in \mathbb{N}}$ be a Følner sequence in the multiplicative group K^* and let $u : K^* \to H$ be a bounded function. Assume that*

$$\lim_{D \to \infty} \frac{1}{|F_D|} \sum_{d \in F_D} \limsup_{N \to \infty} \left| \frac{1}{|F_N|} \sum_{x \in F_N} \langle u(xd), u(x) \rangle \right| = 0.$$

Then

$$\lim_{N \to \infty} \frac{1}{|F_N|} \sum_{x \in F_N} u(x) = 0.$$

Proof. Apply Theorem 2.12 with G being the multiplicative group K^*, endowed with the discrete topology, and λ being the counting measure. $\square$

Theorem 5.6. *Let K be a countable infinite field, let $(A_x)_{x \in K}$ and $(M_y)_{y \in K^*}$ be unitary representations, respectively, of the additive and multiplicative groups $(K, +)$ and $(K^*, \times)$ on the same Hilbert space H and satisfying the distributivity law*

$$M_y A_x = A_{xy} M_y \quad \forall x, y \in K, \ y \neq 0. \tag{5.5}$$

Then for every sequence $(F_N)_{N \in \mathbb{N}}$ of finite subsets of K^* which is simultaneously a Følner sequence for both[3] $(K, +)$ and $(K^*, \times)$, and for every $f \in H$ we have

$$\lim_{N \to \infty} \frac{1}{|F_N|} \sum_{x \in F_N} M_x A_x f = P f$$

where P is the orthogonal projection onto the invariant subspace $\{h \in H : A_x h = M_y h = h \ \forall x, y \in K, \ y \neq 0\}$.

Proof. Each $f \in H$ has a unique representation of the form $f = Pf + g$, where $g = f - Pf$ is orthogonal to the invariant subspace. It is clear that $M_x A_x Pf = Pf$, so it suffices to show that the Cesàro limit of $M_x A_x g$ is 0. The idea is to apply Proposition 5.5 to the function $u : (K^*, \times) \to H$ defined by $u(x) = M_x A_x g$.

Let $d \in K^*$. Using (5.5) and the fact that the representations are unitary we have

$$\langle u(xd), u(x) \rangle = \langle M_{xd} A_{xd} g, M_x A_x g \rangle = \langle A_{-x} M_d A_{xd} g, g \rangle = \langle A_{x(d-1/d)} g, M_{1/d} g \rangle.$$

Observe that $x \mapsto x(d - 1/d)$ is an automorphism of the group $(K, +)$ for every $d \neq 1$. Applying twice the mean ergodic theorem (see, for instance, [8, Theorem 5.5]) we have

$$\lim_{D \to \infty} \frac{1}{|F_D|} \sum_{d \in F_D} \lim_{N \to \infty} \frac{1}{|F_N|} \sum_{x \in F_N} \langle u(dx), u(x) \rangle = \langle P_A g, P_M g \rangle$$

where P_A is the orthogonal projection onto the additively invariant subspace $\{h \in H : A_x h = h \ \forall x \in K\}$ and P_M is the orthogonal projection onto the multiplicatively invariant subspace $\{h \in H : M_x h = h \ \forall x \in K^*\}$. One can show (see [23, Lemma 3.1]) that P_M and P_A commute. Therefore $\langle P_A g, P_M g \rangle = \|P_A P_M g\|^2$, but $P_A P_M g = P_M P_A g$ is invariant under both the additive and the multiplicative representation. This implies that $P_A P_M g = Pg = 0$. We can now invoke Proposition 5.5 and conclude that

$$\lim_{N \to \infty} \frac{1}{|F_N|} \sum_{x \in F_N} M_x A_x g = 0$$

which finishes the proof. $\square$

5.3. A non-linear multiple convergence theorem for ergodic $\mathbb{R}$-actions

In this subsection we demonstrate the power of DT (in the form of Theorem 2.12 with $G = \mathbb{R}$) in yet another situation. Namely, we obtain a rather general kind of non-linear ergodic theorem involving several commuting measure preserving $\mathbb{R}$-actions.

---

[3] Such sequences exist in any countable field K (see Proposition 2.4 in [23]).

Definition 5.7. An increasing function $p \in C^2(b, \infty)$ defined in some interval (b, ∞) is called *admissible* if it satisfies the following:

(1) The derivative p' is monotone.
(2) For every $c \in (0, 1)$ there exist $\epsilon < 1$ such that for large enough $\tau \in \mathbb{R}$ we have
$\epsilon < p(c\tau)/p(\tau) < 1 - \epsilon$.

Observe that the inverse function p^{-1} of an admissible function p exists in some interval $(\tilde{b}, \infty)$ and is also an admissible function. Moreover, the composition $p_1 \circ p_2$ of two admissible functions is again an admissible function. In other words, the class of (germs of) admissible functions forms a group under composition of functions. Observe that if $p \in C^2(b, \infty)$ has monotone derivative and if for every $\epsilon > 0$ there exists a positive α such that

$$\lim_{s \to \infty} \frac{s^\alpha}{f(s)} = \lim_{s \to \infty} \frac{f(s)}{s^{\alpha+\epsilon}} = 0,$$

then f is an admissible function. This observation gives the following examples.

Example 5.8. The following are admissible function:

(1) $p(s) = \sum_{i=1}^{k} c_i s^{\alpha_i}$, where $c_i, \alpha_i \in \mathbb{R}, c_k, \alpha_k > 0, \alpha_1 < \cdots < \alpha_k$.
(2) $p(s) = s^\alpha \log^\beta s$, where $\alpha > 0$ and $\beta \in \mathbb{R}$.
(3) $p(s) = s^\alpha \left(\cos(\log^\beta s) + 2 \right)$, where $\alpha > 0$ and $\beta < 1$.
(4) $p(s) = s^\alpha \left(1 + \frac{\sin s}{\log s} \right)$ for $\alpha > 0$.

Observe that (3) and (4) in Example 5.8 do not belong to any Hardy field.

Here now is the formulation of the first theorem of this subsection.

Theorem 5.9. *Let k, ℓ be positive integers and let $(X, \mathcal{B}, \mu)$ be a probability space. For each $i \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, \ell\}$ let $(T_{i,j}^s)_{s \in \mathbb{R}} : X \to X$ be a continuous ergodic[4] measure preserving $\mathbb{R}$-action and let $p_{i,j} : \mathbb{R} \to \mathbb{R}$ be admissible functions. Assume that all the $T_{i,j}$ commute (in the sense that $T_{i,j}^s T_{\tilde{i},\tilde{j}}^{\tilde{s}} = T_{\tilde{i},\tilde{j}}^{\tilde{s}} T_{i,j}^s$ for any $i, \tilde{i} \in \{1, \ldots, k\}, j, \tilde{j} \in \{1, \ldots, \ell\}$ and $s, \tilde{s} \in \mathbb{R}$) and that $\lim p_{i,j}(s)/p_{\tilde{i},\tilde{j}}(s) = 0$ whenever $i > \tilde{i}$ or ($i = \tilde{i}$ and $j > \tilde{j}$).*
Then for any $f_1, \ldots, f_k \in L^\infty(X)$ we have

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \prod_{i=1}^{k} \left[\left(\prod_{j=1}^{\ell} T_{i,j}^{p_{i,j}(s)} \right) f_i \right] ds = \prod_{i=1}^{k} \int_X f_i \, d\mu \quad in \ L^2(X). \tag{5.6}$$

The strategy of the proof of Theorem 5.9 is to apply a change of variables and then use (an appropriate version of) DT. We will need the following technical lemma, asserting essentially that a change of variables involving an admissible function yields an equivalent method of summation. Similar "change of variable" lemmas appeared for instance in [17, Lemma 7.7] and [1, Lemma 2.2], but we were unable to find the precise formulation we need in the literature.

---

[4] An action $(T_g)_{g \in G}$ of a group G on a probability space $(X, \mathcal{B}, \mu)$ by measurable maps is ergodic if any set $B \in \mathcal{B}$ satisfying $T_g^{-1}(B) = B$ for every $g \in G$ is trivial, in the sense that $\mu(B) \in \{0, 1\}$.

Lemma 5.10. *Let H be a Hilbert space, let $a : \mathbb{R} \to H$ be a continuous and bounded function and let $\sigma \in C^2(0, \infty)$ be an admissible function. Then*

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau a(s) - a\big(\sigma(s)\big) \, ds = 0 \quad in \ norm, \tag{5.7}$$

where the integral is understood in the sense of Bochner.

Proof. Assume, without loss of generality, that $\|a(s)\| \le 1$ for every s. We will use the notation $A_{[x,y]} := \frac{1}{y-x} \int_x^y a(s) \, ds$. Assume that $\lim_{\tau \to \infty} A_{[0,\tau]}$ exists and call it L. Let f be the inverse of σ. Since the contribution of the initial segment $[0, \tau_0]$ is asymptotically negligible when one considers long range Cesàro averages, we may and will assume, without loss of generality, that f is defined on $(0, \infty)$.

We claim that for any $c \in (0, 1)$ and any function $u : \mathbb{R} \to \mathbb{R}$ satisfying $0 \le u(s) \le \sigma\big(cf(s)\big)$ we have

$$\lim_{s \to \infty} A_{[u(s),s]} = L. \tag{5.8}$$

Indeed,

$$A_{[u(s),s]} = \frac{s}{s - u(s)} A_{[0,s]} - \frac{u(s)}{s - u(s)} A_{[0,u(s)]}. \tag{5.9}$$

Since $\sigma(c\tau) \le \delta\sigma(\tau)$ for some δ depending on c (but not on τ), taking $\tau = f(s)$ we get that $u(s) \le \delta s$. It follows that the coefficients of the linear combination in (5.9) add up to 1 and stay bounded as $s \to \infty$. Since both $A_{[0,s]}$ and $A_{[0,u(s)]}$ approach L as $s \to \infty$, this proves the claim (5.8).

We now carry out some computations. Making the change of variables $s = f(t)$ we have

$$\int_0^\tau a\big(\sigma(s)\big) \, ds = \int_0^{\sigma(\tau)} a(t) f'(t) \, dt = \int_0^{\sigma(\tau)} a(t) \int_0^t f''(u) \, du \, dt.$$

Since both functions a and f'' are bounded in the relevant intervals, we may switch the order of integrals and obtain

$$\int_0^\tau a\big(\sigma(s)\big) \, ds = \int_0^{\sigma(\tau)} f''(u) \int_u^{\sigma(\tau)} a(t) \, dt \, du$$

$$= \int_0^{\sigma(\tau)} f''(u)\big(\sigma(\tau) - u\big) A_{[u,\sigma(\tau)]} \, du. \tag{5.10}$$

For future use, we now compute the following integral

$$\int_x^y f''(u)\big(\sigma(\tau) - u\big) \, du = f'(u)\big(\sigma(\tau) - u\big)\Big|_x^y + \int_x^y f'(u) \, du$$

$$= f'(y)\big(\sigma(\tau) - y\big) - f'(x)\big(\sigma(\tau) - x\big) + f(y) - f(x). \tag{5.11}$$

Let $\epsilon > 0$, let $c \in (1 - \epsilon, 1)$ and let $M = \sup_\tau \sigma'(\tau)/\sigma'(c\tau)$. Observe that M is finite, as a consequence of L'Hôpital's rule (and of the fact that σ is an admissible function). We will show that we can truncate the integral in (5.10) at $u = \sigma(c\tau)$, at the cost of an error of size at most $M\epsilon\tau$.

Since $\|a(s)\| \leq 1$, we have $\|A_{[x,y]}\| \leq 1$ for every $x < y$. Also, since f is an admissible function, the sign of f'' does not change. Therefore, using (5.11) we obtain

$$
\left\| \int_{\sigma(c\tau)}^{\sigma(\tau)} f''(u)\big(\sigma(\tau) - u\big) A_{[u,\sigma(\tau)]} \, du \right\| \leq \left| \int_{\sigma(c\tau)}^{\sigma(\tau)} f''(u)\big(\sigma(\tau) - u\big) \, du \right|
$$

$$
= \left| -f'\big(\sigma(c\tau)\big)\big(\sigma(\tau) - \sigma(c\tau)\big) + f\big(\sigma(\tau)\big) - f\big(\sigma(c\tau)\big) \right|
$$

$$
= \left| \frac{\sigma(\tau) - \sigma(c\tau)}{-\sigma'(c\tau)} + \tau(1 - c) \right|
$$

$$
= \tau(1 - c) \left| 1 - \frac{\sigma'(b)}{\sigma'(c\tau)} \right|
$$

for some $b \in (c\tau, \tau)$. If σ' is decreasing, the last expression is bounded by $\tau\epsilon$. If σ' is increasing, it is bounded by $\tau\epsilon(M - 1)$. In either case we obtain

$$
\left\| \frac{1}{\tau} \int_{\sigma(c\tau)}^{\sigma(\tau)} f''(u)\big(\sigma(\tau) - u\big) A_{[u,\sigma(\tau)]} \, du \right\| \leq M\epsilon. \tag{5.12}
$$

Putting (5.10) and (5.12) together we obtain

$$
\frac{1}{\tau} \int_0^\tau a\big(\sigma(s)\big) \, ds = \frac{1}{\tau} \int_0^{\sigma(c\tau)} f''(u)\big(\sigma(\tau) - u\big) A_{[u,\sigma(\tau)]} \, du + R_1
$$

where $R_1 = R_1(\tau) \in H$ satisfies $\|R_1\| \leq M\epsilon$.

On the other hand, for $u \in \big(0, \sigma(c\tau)\big)$, using (5.8) with $s = \phi(\tau)$ (and hence $\tau = f(s)$) and τ large enough, we obtain that $\|A_{[u,\sigma(\tau)]} - L\| < \epsilon$. Therefore

$$
\frac{1}{\tau} \int_0^\tau a\big(\sigma(s)\big) \, ds = \frac{1}{\tau} \int_0^{\sigma(c\tau)} f''(u)\big(\sigma(\tau) - u\big) L \, du + R_2 + R_1
$$

for some $R_2 = R_2(\tau) \in H$ satisfying $\|R_2\| \leq \epsilon$. Calculating the last integral using (5.11), taking the limit as $\tau \to \infty$ and letting $\epsilon \to 0$ we conclude that

$$
\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau a\big(\sigma(s)\big) \, ds = L. \quad \square
$$

We will need yet another lemma:

Lemma 5.11. *Let $(X, \mathcal{B}, \mu, (T^s)_{s\in\mathbb{R}})$ be an ergodic continuous measure preserving $\mathbb{R}$-action and let $b : \mathbb{R} \to \mathbb{R}$ be a continuous function such that $\lim_{s\to\infty} b(s) = 0$. Let $f \in L^1(X)$ and let $(g_s)_{s\in\mathbb{R}}$ be a family of functions in $L^\infty(X)$ such that $\sup_s \|g_s\|_{L^\infty} < \infty$. Then*

$$
\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \int_X f \cdot g_s \, d\mu \, ds = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \int_X T^{b(s)} f \cdot g_s \, d\mu \, ds
$$

in the sense that the existence of one limit implies the existence of the other and the equality.

Proof. It suffices to show that

$$
\lim_{s \to \infty} \int_X \big(f - T^{b(s)} f\big) \cdot g_s \, d\mu = 0.
$$

Let $M = \sup_s \|g_s\|_{L^\infty}$. Then we have

$$\left| \int_X \left(f - T^{b(s)} f \right) \cdot g_s \, \mathrm{d}\mu \right| \leq M \left\| f - T^{b(s)} f \right\|_{L^1}.$$

Since $b(s) \to 0$ and T is a continuous action, we deduce that $T^{b(s)} f \to f$ in $L^1(X)$ and hence

$$\limsup_{s \to \infty} \left| \int_X \left(f - T^{b(s)} f \right) \cdot g_s \, \mathrm{d}\mu \right| \leq M \lim_{s \to \infty} \left\| f - T^{b(s)} f \right\|_{L^1} = 0. \quad \square$$

The next proposition uses DT. The exact form we need follows from Theorem 2.12 by letting $G = \mathbb{R}$ with its usual topology and the Lebesgue Haar measure λ, and letting $(F_N)_{N \in \mathbb{N}}$ be any sequence of intervals $[0, b_N]$ with $b_N \to \infty$.

Proposition 5.12. *Let k, ℓ be positive integers and let $(X, \mathcal{B}, \mu)$ be a probability space. For each $i \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, \ell\}$ let $(T_{i,j}^s)_{s \in \mathbb{R}} : X \to X$ be a continuous ergodic measure preserving $\mathbb{R}$-action and let $q_{i,j} : \mathbb{R} \to \mathbb{R}$ be an admissible function. Assume that all the $T_{i,j}$ commute, that $\lim q_{1,1}'(s) = 1$ as $s \to \infty$ and that $\lim q_{i,j}'(s) = 0$ for each $(i, j) \neq (1, 1)$.*

If either $k = 1$, or $k > 1$ and Theorem 5.9 holds for $k - 1$, then for any $f_1, \ldots, f_k \in L^\infty(X)$ with $\int_X f_1 d\mu = 0$ we have

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \prod_{i=1}^k \left[\left(\prod_{j=1}^\ell T_{i,j}^{q_{i,j}(s)} \right) f_i \right] \mathrm{d}s = 0.$$

Proof. With Theorem 2.12 in mind (with $G = \mathbb{R}$ and $(F_N)_{N \in \mathbb{N}}$ being any sequence of intervals $[0, \tau_N]$ with $\tau_N \to \infty$), let $a(s) = \prod_{i=1}^k \left[\left(\prod_{j=1}^\ell T_{i,j}^{q_{i,j}(s)} \right) f_i \right]$ and $h \in \mathbb{R}$. We have

$$\langle a(s + h), a(s) \rangle = \int_X \prod_{i=1}^k \left[\left(\prod_{j=1}^\ell T_{i,j}^{q_{i,j}(s+h)} \right) f_i \cdot \left(\prod_{j=1}^\ell T_{i,j}^{q_{i,j}(s)} \right) f_i \right] \mathrm{d}\mu$$

$$= \int_X \prod_{i=1}^k \left[\left(\prod_{j=1}^\ell T_{i,j}^{q_{i,j}(s)} \right) \left[f_i \cdot \left(\prod_{j=1}^\ell T_{i,j}^{q_{i,j}(s+h) - q_{i,j}(s)} \right) f_i \right] \right] \mathrm{d}\mu.$$

The mean value theorem implies that $\lim_s q_{i,j}(s + h) - q_{i,j}(s) = 0$ for each fixed h and every $(i, j) \neq (1, 1)$, and also that $\lim_s q_{1,1}(s + h) - q_{1,1}(s) - h = 0$ for any fixed h. In view of Lemma 5.11 we have

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \langle a(s + h), a(s) \rangle \, \mathrm{d}s$$

$$= \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \int_X \left(\prod_{j=1}^\ell T_{1,j}^{q_{1,j}(s)} \right) (f_1 \cdot T_{1,1}^h f_1) \cdot \prod_{i=2}^k \left(\prod_{j=1}^\ell T_{i,j}^{q_{i,j}(s)} \right) (f_i^2) \, \mathrm{d}\mu \, \mathrm{d}s. \quad (5.13)$$

If $k = 1$ the expression (5.13) reduces to

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \int_X \left(\prod_{j=1}^\ell T_{1,j}^{q_{1,j}(s)} \right) (f_1 \cdot T_{1,1}^h f_1) \, \mathrm{d}\mu \, \mathrm{d}s.$$

Using the fact that each $T_{1,j}^{q_{1,j}(s)}$ preserves μ, we have

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \langle a(s+h), a(s) \rangle \, \mathrm{d}s = \int_X \left(f_1 \cdot T_{1,1}^h f_1 \right) \mathrm{d}\mu. \tag{5.14}$$

It follows from the mean ergodic theorem (see [53]) that the Cesàro limit in h of the right hand side of (5.14) is $\left(\int_X f_1 d\mu \right)^2 = 0$ and hence, by Theorem 2.12 the proof is complete.

Assume now that $k > 1$ and Theorem 5.9 is true for $k - 1$. Since each $T_{k,j}$ preserves μ we can apply $\left(\prod_{j=1}^\ell T_{k,j}^{-q_{k,j}(s)} \right)$ to the right hand side of (5.13). Letting $\tilde{f}_1 = f_1 \cdot T_{1,1}^h f_1$ and $\tilde{f}_i = f_i^2$ for $i = 2, \ldots, k - 1$ we have

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \langle a(s+h), a(s) \rangle \, \mathrm{d}s$$

$$= \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \int_X \prod_{i=1}^{k-1} \left(\prod_{j=1}^\ell T_{i,j}^{q_{i,j}(s)} T_{k,j}^{-q_{k,j}(s)} \right) \tilde{f}_i \cdot f_k^2 \, \mathrm{d}\mu \, \mathrm{d}s. \tag{5.15}$$

It is not hard to check that the right hand side of (5.15) satisfies the conditions of Theorem 5.9 with $k - 1$. Therefore

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \langle a(s+h), a(s) \rangle \, \mathrm{d}s = \int_X f_1 \cdot T_{1,1}^h f_1 \, \mathrm{d}\mu \cdot \prod_{i=2}^k \int_X f_i^2 \, \mathrm{d}\mu.$$

By the mean ergodic theorem we get

$$\lim_{H \to \infty} \frac{1}{H} \int_0^H \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \langle a(s+h), a(s) \rangle \, \mathrm{d}s \, \mathrm{d}h = \left(\int_X f_1 \, \mathrm{d}\mu \right)^2 \cdot \prod_{i=2}^k \int_X f_i^2 \, \mathrm{d}\mu = 0.$$

The result in question follows now from Theorem 2.12.   □

We can now prove Theorem 5.9:

Proof of Theorem 5.9. We proceed by induction. Assume that either $k = 1$, or $k > 1$ and the theorem has been proved for $k - 1$. After decomposing $f_1 = \int_X f_1 \, \mathrm{d}\mu + g_1$, where $\int_X g_1 \, \mathrm{d}\mu = 0$, the left hand side of (5.6) becomes the sum of two terms. The first term is

$$\int_X f_1 \, \mathrm{d}\mu \cdot \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \prod_{i=2}^k \left[\left(\prod_{j=1}^\ell T_{i,j}^{p_{i,j}(s)} \right) f_i \right] \mathrm{d}s. \tag{5.16}$$

Since we assumed that either $k = 1$, or the theorem holds for $k - 1$, it follows that (5.16) equals the left hand side of (5.6) as desired. The second term in the decomposition is

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \left(\prod_{j=1}^\ell T_{1,j}^{p_{1,j}(s)} \right) g_1 \cdot \prod_{i=2}^k \left[\left(\prod_{j=1}^\ell T_{i,j}^{p_{i,j}(s)} \right) f_i \right] \mathrm{d}s$$

and all that remains to prove is that it equals 0. Therefore we have reduced the Theorem to the case when $\int_X f_1 \, \mathrm{d}\mu = 0$.

Let σ be the inverse of $p_{1,1}$. Since $p_{1,1}$ is an admissible function, so is σ. Let $q_{i,j} = p_{i,j} \circ \sigma$ for each $i = 1, \ldots, k$ and $j = 1, \ldots \ell$; observe that these are all admissible functions. Since for

every $(i, j) \neq (1, 1)$, $p_{i,j}(s)/p_{1,1}(s) \to 0$ as $s \to \infty$, also $q_{i,j}(s)/s \to 0$ as $s \to \infty$. In view of Lemma 5.10, the left hand side of (5.6) can be replaced with

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \prod_{i=1}^k \left[\left(\prod_{j=1}^\ell T_{i,j}^{q_{i,j}(s)} \right) f_i \right] \mathrm{d}s.$$

The result now follows from Proposition 5.12. □

6. A multiplicative variant of DT

Definition 6.1. A function $F : \mathbb{N} \to \mathbb{C}$ is *multiplicative* if for any coprime $a, b \in \mathbb{N}$ we have $F(ab) = F(a)F(b)$.

Classical examples of multiplicative functions include the Möbius function $\mu(n)$ and the Liouville function $\lambda(n)$. The Liouville function is defined by $\lambda(p) = -1$ for any prime p and $\lambda(ab) = \lambda(a)\lambda(b)$ for *every* $a, b \in \mathbb{N}$. The Möbius function is equal to the Liouville function for every squarefree number, and $\mu(n) = 0$ whenever n is divisible by a perfect square k^2 with $k \geq 2$. The following theorem, due to Kátai, has a clear similarity to DT, in the form of Theorem 2.4.

Theorem 6.2 ([49]). *Let* $a : \mathbb{N} \to \mathbb{C}$ *be bounded and let* $F : \mathbb{N} \to \mathbb{N}$ *be a multiplicative function. If for any distinct primes* p, q,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^N a(pn)\overline{a(qn)} = 0 \tag{6.1}$$

then

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^N a(n)F(n) = 0. \tag{6.2}$$

Limiting statements of the form (6.2) are central to analytic number theory. For instance, if $a(n) \equiv 1$ and $F(n) = \mu(n)$ (the Möbius function), then (6.2) is equivalent to the Prime Number theorem. If $a(n)$ is replaced with a periodic function, then (6.2) is equivalent to the prime number theorem in arithmetic progressions. (Unfortunately, neither of these theorems can be easily derived from Theorem 6.2.) For $a(n) = e^{2\pi i \alpha n}$, where α is irrational, an application of Theorem 6.2 gives a Theorem of Daboussi (cf. [35]), stating that, for this choice of $a(n)$, (6.2) holds for any multiplicative function F. More generally, one can use Theorem 6.2 to deduce the following result similar in spirit to (but much simpler than) the main result in [43]:

Theorem 6.3. *Let* $f \in \mathbb{R}[x]$ *be a polynomial with an irrational coefficient (other than the constant term) and let* $F : \mathbb{N} \to \mathbb{C}$ *be a bounded multiplicative function. Then*

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

Proof. In view of Theorem 6.2 it suffices to show that (6.1) holds for $a(n) = e^{2\pi i f(n)}$. For any primes $p \neq q$, the map $n \mapsto f(pn) - f(qn)$ is a polynomial with an irrational (non-constant)

coefficient. Therefore

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} a(pn)\overline{a(qn)} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i \left(f(pn) - f(qn) \right)} = 0$$

and this finishes the proof. □

7. Limits along ultrafilters

In this section we will use a variant of DT for limits along idempotent ultrafilters. Since Følner sequences are only available in amenable groups, the form of DT in Theorem 2.12 does not apply to non-amenable groups, such as a (non-commutative) free group or $SL(n, \mathbb{Z})$. Nevertheless, it is possible to study multiple recurrence in the setting of non-amenable groups through the use of ultrafilters. The following definitions and facts about ultrafilters can be found, for instance, in [9,47].

Definition 7.1. An *ultrafilter* on a countable group G is a family p of subsets of G which is closed under intersections, supersets and which contains exactly one member of any finite partition of G. The set of all ultrafilters is denoted by βG.

The set βG can be identified with the Stone–Čech compactification of G. In particular, the operation on G can be lifted to βG. Of special interest are ultrafilters $p \in \beta G$ which satisfy $p \cdot p = p$; such ultrafilters are called *idempotent*.

Given a map $f : G \to K$ from G to a compact Hausdorff space K and an ultrafilter $p \in \beta G$, we denote by $p\text{-}\lim_g f(g)$ the (unique) point $x \in K$ with the property that $\{g \in G : f(g) \in U\} \in p$ for any neighborhood U of x. It follows easily from the definitions that for any $f : G \to K$ from G to a compact Hausdorff space and for any $p \in \beta G$, $p\text{-}\lim_g f(g)$ exists and is unique. One property which will be important for the exposition in this section is the following:

$$p\text{-}\lim_g p\text{-}\lim_h f(gh) = p\text{-}\lim_g f(g) \quad \text{whenever } p \in \beta G \text{ is idempotent.} \tag{7.1}$$

The following proposition is the analog of Theorem 3.2 in the setting of limits along ultrafilters. It appeared originally as Theorem 2.3 in [21].

Proposition 7.2. *Let G be a countable group, let H be a Hilbert space, let $u : G \to H$ be a bounded sequence and let $p \in \beta G$ be an idempotent ultrafilter. Then*

$$\text{If } p\text{-}\lim_h p\text{-}\lim_g \langle u(hg), u(g) \rangle = 0 \quad \text{then } p\text{-}\lim_g u(g) = 0$$

where the last limit is in the weak topology.

The proof presented here is due to Schnell [58].

Proof. For each $N \in \mathbb{N}$ we have that

$$p\text{-}\lim_g u(g) = p\text{-}\lim_{g_1} p\text{-}\lim_{g_2} \cdots p\text{-}\lim_{g_N} \frac{1}{N} \sum_{k=1}^{N} u(g_k \cdots g_N).$$

Taking norms and using the Cauchy–Schwarz inequality, we get:

$$\left\| p\text{-}\lim_g u(g) \right\|^2 \leq p\text{-}\lim_{g_1} p\text{-}\lim_{g_2} \cdots p\text{-}\lim_{g_N} \frac{1}{N^2} \left\| \sum_{k=1}^{N} u(g_k \cdots g_N) \right\|^2$$

$$= p\text{-}\lim_{g_1} p\text{-}\lim_{g_2} \cdots p\text{-}\lim_{g_N} \frac{1}{N^2} \left\langle \sum_{k=1}^{N} u(g_k \cdots g_N), \sum_{l=1}^{N} u(g_l \cdots g_N) \right\rangle$$

$$= \frac{1}{N^2} \sum_{k,l=1}^{N} p\text{-}\lim_{g_1} p\text{-}\lim_{g_2} \cdots p\text{-}\lim_{g_N} \langle u(g_k \cdots g_N), u(g_l \cdots g_N) \rangle.$$

Now we use the fact that p is an idempotent ultrafilter to get:

$$= \frac{1}{N^2} \sum_{k=1}^{N} p\text{-}\lim_{g_k \to \infty} \|u(n_k)\|^2 + \frac{2}{N^2} \sum_{k<l} p\text{-}\lim_{g_k \to \infty} p\text{-}\lim_{g_l \to \infty} \langle u(g_k g_l), u(g_l) \rangle$$

$$= \frac{1}{N} p\text{-}\lim_{g \to \infty} \|u(g)\|^2.$$

Since N can be chosen arbitrarily we conclude that $p\text{-}\lim_{n \to \infty} x_n = 0$. $\quad\square$

Proposition 7.2 can be used to give a quick proof of the following polynomial extension of Khintchine's recurrence theorem (see, for instance, [7, Theorem 5.1]).

Theorem 7.3 (*Cf. [7, Theorem 3.12]*)**.** *Let (X, μ, T) be an invertible measure preserving system, let $q \in \mathbb{Z}[x]$ be a polynomial satisfying $q(0) = 0$ and let $p \in \beta\mathbb{Z}$ be an idempotent ultrafilter. Then there exists an orthogonal projection $P : L^2(X) \to L^2(X)$ such that for any $f \in L^2(X)$*

$$p\text{-}\lim_n f \circ T^{q(n)} = Pf \quad \text{in the weak topology.}$$

In particular, for any $\epsilon > 0$, the set $R := \left\{ n \in \mathbb{Z} : \mu(A \cap T^{q(n)}A) > \mu(A)^2 - \epsilon \right\}$ is in p and hence non-empty.

In fact, Theorem 7.3 implies that R is an IP*-set (see Definition 8.1). This result was first obtained in [10] (see also Section 3 in [7]). One can deduce combinatorial corollaries from recurrence results using Furstenberg's correspondence principle (see [38,39,5]). In particular, Theorem 7.3 implies that for any set $A \subset \mathbb{N}$ with positive upper density $\bar{d}(A) := \limsup |A \cap [1, N]|/N > 0$ and for any $q \in \mathbb{Z}[x]$ with $q(0) = 0$, there exist many $x, y \in \mathbb{N}$ such that $\{x, x + q(y)\} \subset A$.

Next we define the notion of a p-mixing measure preserving system.

Definition 7.4. Let G be a countable group and let $p \in \beta G$. A measure preserving system $(X, \mu, (T_g)_{g \in G})$ is called *p-mixing* if for every $f \in L^2(X)$ we have $p\text{-}\lim_g T_g f = \int_X f \, d\mu$ in the weak topology.

When $G = \mathbb{Z}$ and $p \in \beta\mathbb{Z}$, a measure preserving system (X, μ, T) is called *totally p-mixing* if for every $\ell \in \mathbb{N}$ the system (X, μ, T^ℓ) is p-mixing. Equivalently, (X, μ, T) is totally p-mixing if for every $f \in L^2(X)$ and every $\ell \in \mathbb{N}$ we have $p\text{-}\lim_n T^{\ell n} f = \int_X f \, d\mu$ in the weak topology.

7.1. A polynomial ergodic theorem for totally p-mixing systems

In this subsection we will use Proposition 7.2 to deduce the following result.

Theorem 7.5. *Let $p \in \beta\mathbb{N}$ be an idempotent ultrafilter and let (X, μ, T) be a totally p-mixing measure preserving system. Let $p_1, \ldots, p_k \in \mathbb{Z}[x]$ be polynomials such that for every $i \neq j$, the*

polynomial $p_i - p_j$ is not constant. For every $f_1, \ldots, f_k \in L^\infty(X)$ we have:

$$p\text{-}\lim_n \prod_{i=1}^k T^{p_i(n)} f_i = \prod_{i=1}^k \int_X f_i \, d\mu \quad \text{in the weak topology.} \tag{7.2}$$

If the system is p-mixing but not totally p-mixing, the conclusion of Theorem 7.5 may not hold. In fact, there are p-mixing systems satisfying $p\text{-}\lim_n T^{2n} f = f$ for every $f \in L^2$, see [14].

The proof of Theorem 7.5 utilizes the PET-induction scheme developed in [6]. Two polynomials $p_1, p_2 \in \mathbb{Z}[x]$ are *equivalent* if they have the same degree and leading coefficient, in other words, if $p_2 - p_1$ has degree strictly smaller than p_2. Given a finite family $P = \{p_1, \ldots, p_k\}$ of polynomials, let d be the maximal degree of the p_i and, for each $j = 1, \ldots, d$ let s_j be the number of equivalence classes in P of polynomials of degree j. The vector $(s_1, \ldots, s_d)$ is called the *characteristic vector* of P. Since $s_d > 0$, the characteristic vector is unique. For example, the family $P = \{x, 2x - 1, 3x, x^3 + 2x^2, x^3 + 1\}$ has characteristic vector $(3, 0, 1)$.

We order the characteristic vectors by letting $(s_1, \ldots, s_d) < (\tilde{s}_1, \ldots, \tilde{s}_{\tilde{d}})$ if and only if either $d < \tilde{d}$ or both $d = \tilde{d}$ and the maximum j for which $s_j \neq \tilde{s}_j$ satisfies $s_j < \tilde{s}_j$. For example $(1, 2, 3) < (0, 0, 0, 1)$ and $(9, 3, 5, 2, 4) < (1, 7, 6, 2, 4)$.

We can now prove Theorem 7.5.

Proof of Theorem 7.5. Theorem 7.5 We will prove this theorem by induction on the characteristic vector of $P = \{p_1, \ldots, p_k\}$. If P has characteristic vector (1), then P must consist of a single linear polynomial and (7.2) holds trivially. Next assume that the characteristic vector $(s_1, \ldots, s_d)$ of P is larger than the vector (1) and that Theorem 7.5 has been proved for any family $\tilde{P}$ with a smaller characteristic vector. Without loss of generality we can assume that $\int_X f_i = 0$ for every $i = 1, \ldots, k$. We will also assume that the polynomials in P have 0 constant term.

Let $q \in P$ have the smallest degree. If we denote by e the degree of q, then $s_e \geq 1$ and $s_j = 0$ for each $j < e$.

Let $u_n = \prod_{p \in P} T^{p(n)} f_p$. We need to show that $p\text{-}\lim_n u_n = 0$. We consider separately two cases.

Case 1 ($e > 1$). Let $Q \subset P$ be the equivalence class of q. With Proposition 7.2 in mind, we look at the inner product $\langle u_{n+h}, u_n \rangle$:

$$\langle u_{n+h}, u_n \rangle = \int_X \prod_{p \in P} T^{p(n+h)} f_p \cdot T^{p(n)} f_p \, d\mu$$

$$= \int_X \prod_{p \in P} T^{p(n+h)-q(n)} f_p \cdot T^{p(n)-q(n)} f_p \, d\mu$$

$$= \int_X f_q \cdot \prod_{\tilde{p} \in P_h} T^{\tilde{p}(n)} \tilde{f}_{\tilde{p}} \, d\mu$$

where P_h is the family of polynomials consisting of the maps $\tilde{p} : n \mapsto p(n) - q(n)$ for $p \in P \setminus \{q\}$, associated with the function $\tilde{f}_{\tilde{p}} = f_p$, and the maps $\tilde{p}_h : n \mapsto p(n + h) - q(n)$ for any $p \in P$, associated with the function $\tilde{f}_{\tilde{p}} = f_p$.

Notice that, for every $p \in P \setminus Q$ and every $h \in \mathbb{N}$, the polynomial $n \mapsto p(n) - q(n)$ is equivalent to $n \mapsto p(n + h) - q(n)$ and both have the same degree as p. Next, observe that whenever $p_1, p_2 \in P \setminus Q$, the polynomial $n \mapsto p_1(n) - q(n)$ is equivalent to

$n \mapsto p_2(n) - q(n)$ if and only if p_1 is equivalent to p_2. Finally note that if $p \in Q$, then the degree of $p(n) - q(n)$ and $p(n + h) - q(n)$ is strictly smaller than e. Since all polynomials in P have degree at least 2, the family P_h contains no two polynomials p_1, p_2 whose difference $p_1 - p_2$ is a constant (for all but finitely many h). It follows from the observations above that the characteristic vector of P_h is strictly smaller than that of P. From the induction hypothesis we obtain

$$p\text{-}\lim_n \langle u_{n+h}, u_n \rangle = \int_X f_q \left(p\text{-}\lim_n \prod_{\tilde{p} \in P_h} T^{\tilde{p}(n)} \tilde{f}_{\tilde{p}} \right) \mathrm{d}\mu = 0$$

and hence, the desired result $p\text{-}\lim_n u_n = 0$ follows from the ultrafilter version of DT (Proposition 7.2).

Case 2 ($e = 1$). In this case, for any $p \in P$ with degree 1, the polynomials $p(n) - q(n)$ and $p(n+h) - q(n)$ have a constant difference. To overcome this difficulty, let $n \mapsto a_i n$, $i = 1, \ldots, s$ be all the polynomials of degree 1 in P, and assume that $q(n) = a_1 n$. Observe that all the $a_i \in \mathbb{Z}$ are distinct. Let Q be the set of polynomials in P with degree at least 2. We now have

$$\langle u_{n+h}, u_n \rangle = \int_X \prod_{i=1}^s T^{a_i(n+h)} f_i \cdot T^{a_i n} f_i \cdot \prod_{p \in Q} T^{p(n+h)} f_p \cdot T^{p(n)} f_p \, \mathrm{d}\mu$$

$$= \int_X \prod_{i=1}^s T^{(a_i - a_1)n} \left(T^{a_i h} f_i \cdot f_i \right) \cdot \prod_{p \in Q} T^{p(n+h)-q(n)} f_p \cdot T^{p(n)-q(n)} f_p \, \mathrm{d}\mu$$

$$= \int_X T^{a_1 h} f_1 \cdot f_1 \cdot \prod_{\tilde{p} \in P_h} T^{\tilde{p}(n)} \tilde{f}_{\tilde{p}} \, \mathrm{d}\mu$$

where the family P_h consists of the polynomials $\tilde{p} : n \mapsto (a_i - a_1)n$ for $i = 2, \ldots, s$, associated with the function $\tilde{f}_{\tilde{p}} = T^{a_i} f_i \cdot f_i$ together with the polynomials $\tilde{p} : n \mapsto p(n) - q(n)$ and $\tilde{p}_h : n \mapsto p(n + h) - q(n)$, both associated with the function $\tilde{f}_{\tilde{p}} = \tilde{f}_{\tilde{p}_h} = f_p$.

Applying the same reasoning as in the first case we deduce that the family P_h has a smaller characteristic vector than P. It follows from the induction hypothesis that

$$p\text{-}\lim_n \langle u_{n+h}, u_n \rangle = \int_X T^{a_1 h} f_1 \cdot f_1 \, \mathrm{d}\mu \cdot \prod_{p \in P_h} \int_X \tilde{f}_p \, \mathrm{d}\mu.$$

Since each $\tilde{f}_p \in L^\infty(X)$, the absolute value of the second term in this expression is bounded, say by M. Hence we deduce that

$$p\text{-}\lim_h \left| p\text{-}\lim_n \langle u_{n+h}, u_n \rangle \right| \leq p\text{-}\lim_h \left| \int_X T^{a_1 h} f_1 \cdot f_1 \, \mathrm{d}\mu \right| \cdot M = 0.$$

Finally, applying Proposition 7.2 we obtain the desired conclusion. □

Recall that a measure preserving system is called *mildly mixing* if for any $f \in L^2(X)$ and any $p = p + p \in \beta\mathbb{N}$ we have $p\text{-}\lim_n T^n f = \int_X f \, \mathrm{d}\mu$. We note in passing that every mixing system is mildly mixing and every mildly mixing system is weakly mixing, but both inclusions are strict (see, for example [42]).

Since for every idempotent $p \in \beta\mathbb{N}$ and $\ell \in \mathbb{N}$, the ultrafilter $\ell p \in \beta\mathbb{N}$[5] is also idempotent, it follows that a mildly mixing system is in fact totally p-mixing for every idempotent ultrafilter p. This allows us to deduce from Theorem 7.5 the following corollary:

Corollary 7.6. *Let (X, μ, T) be a mildly mixing measure preserving system and let $p_1, \ldots, p_k \in \mathbb{Z}[x]$ be polynomials. If for every $i \neq j$, the polynomial $p_i - p_j$ is not constant, then for every $f_1, \ldots, f_k \in L^\infty(X)$ we have (7.2). In particular, for every $\epsilon > 0$ and $f_0 \in L^2(X)$, the set*

$$\left\{ n : \left| \int_X f_0 \cdot \prod_{i=1}^k T^{p_i(n)} f_i \, d\mu - \prod_{i=0}^k \int_X f_i \, d\mu \right| < \epsilon \right\}$$

is IP*.

7.2. An ultrafilter analog of joint ergodicity

The mean ergodic theorem states that a measure preserving system $(X, \mathcal{B}, \mu, T)$ is ergodic if and only if for every $f \in L^2(X)$, the averages $\frac{1}{N} \sum_{n=1}^N T^n f$ converge in norm to $\int f \, d\mu$ as $N \to \infty$. The following definition is an extension of the notion of ergodicity to several transformations:

Definition 7.7. Let $k \in \mathbb{N}$ and let $T_1, \ldots, T_k$ be invertible measure preserving transformations on the same probability space $(X, \mathcal{B}, \mu)$. We say that the maps $T_1, \ldots, T_k$ are *jointly ergodic* if for any $f_1, \ldots, f_k \in L^\infty(X)$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^N T_1^n f_1 \cdots T_k^n f_k = \int_X f_1 \, d\mu \cdots \int_X f_k \, d\mu \quad \text{in } L^2.$$

The notion of joint ergodicity was introduced in [4], where the authors gave the following necessary and sufficient conditions for a tuple of invertible commuting transformations to be jointly ergodic.

Theorem 7.8 ([4]). *Let $k \in \mathbb{N}$ and let $T_1, \ldots, T_k$ be invertible measure preserving transformations on the same probability space $(X, \mathcal{B}, \mu)$. Then $T_1, \ldots, T_k$ are jointly ergodic if and only if all the transformations $T_i T_j^{-1}$ and $T_1 \times \cdots \times T_k$ are ergodic.*

Later, in [24], the definition of joint ergodicity was (somewhat modified and) extended to actions of more general groups. Let G be a countable group[6] and let $(T_1^g)_{g \in G}, \ldots, (T_k^g)_{g \in G}$ be measure preserving actions of G on a probability space $(X, \mathcal{B}, \mu)$. We say that these actions *commute* if for any $g, h \in G$ and $i \neq j$ we have $T_i^g T_j^h = T_j^h T_i^g$. Note that when G is non-commutative, this equality may fail for $i = j$.

In the above setting, whenever $1 \leq i \leq j \leq k$, we denote by $T_{[i,j]}^g$ the measure preserving map $T_{[i,j]}^g = T_i^g T_{i+1}^g \cdots T_j^g$. Observe that the commutativity assumption implies that $\left(T_{[i,j]}^g\right)_{g \in G}$ is an action of G.

---

[5] The ultrafilter ℓp is defined by $A \in \ell p \iff A/\ell \in p$ (where, in turn, A/ℓ is defined by $x \in A/\ell \iff \ell x \in A$).

[6] In [24] the setup is that of general locally compact amenable groups.

Definition 7.9. Let G be a countable amenable group, let $(F_N)_{N\in\mathbb{N}}$ be a Følner sequence in G, let $k \in \mathbb{N}$ and let $(T_1^g)_{g\in G}, \ldots, (T_k^g)_{g\in G}$ be commuting measure preserving actions of G on a probability space $(X, \mathcal{B}, \mu)$. The actions $T_1, \ldots, T_k$ are called *jointly ergodic* if for every $f_1, \ldots, f_k \in L^\infty(X)$

$$\lim_{N\to\infty} \frac{1}{|F_N|} \sum_{g\in F_N} \prod_{i=1}^k T_{[1,i]}^g f_i = \prod_{i=1}^k \int_X f_i \, d\mu \quad \text{in } L^2(X).$$

One has the following theorem.

Theorem 7.10 (*Cf. [24, Theorems 2.4 and 2.6]*)**.** *Let G be a countable amenable group, let $(F_N)_{N\in\mathbb{N}}$ be a Følner sequence in G, let $k \in \mathbb{N}$ and let $(T_1^g)_{g\in G}, \ldots, (T_k^g)_{g\in G}$ be commuting measure preserving actions of G on a probability space $(X, \mathcal{B}, \mu)$. Then the actions $T_1, \ldots, T_k$ are jointly ergodic if and only if*

$$\left(T_{[i,j]}^g\right)_{g\in G} \quad \text{are ergodic for every } i \leq j$$

and $T_1 \times T_{[1,2]} \times \cdots \times T_{[1,k]}$ is ergodic on $(X^k, \mathcal{B}^{\otimes k}, \mu^{\otimes k})$.

One can show that when $G = \mathbb{Z}$, Theorem 7.10 implies Theorem 7.8.

In this subsection we establish necessary and sufficient conditions, analogous to those obtained in [4,24], for joint p-mixing.

Definition 7.11. Let G be a countable group, let $p \in \beta G$ be an idempotent ultrafilter, let $k \in \mathbb{N}$ and let $(T_1^g)_{g\in G}, \ldots, (T_k^g)_{g\in G}$ be commuting measure preserving actions of G on a probability space $(X, \mathcal{B}, \mu)$. The actions $T_1, \ldots, T_k$ are called *jointly p-mixing* if for every $f_1, \ldots, f_k \in L^\infty(X)$

$$p\text{-}\lim_g \prod_{i=1}^k T_{[1,i]}^g f_i = \prod_{i=1}^k \int_X f_i \, d\mu \quad \text{weakly in } L^2(X). \tag{7.3}$$

Theorem 7.12. *Let G be a countable group, let $p \in \beta G$ be an idempotent ultrafilter, let $k \in \mathbb{N}$ and let $(T_1^g)_{g\in G}, \ldots, (T_k^g)_{g\in G}$ be commuting measure preserving actions of G on a probability space $(X, \mathcal{B}, \mu)$. Then the actions $T_1, \ldots, T_k$ are jointly p-mixing if and only if*

$$\left(T_{[i,j]}^g\right)_{g\in G} \quad \text{are p-mixing for every } i \leq j. \tag{7.4}$$

Proof. We first prove that (7.4) implies joint p-mixing, i.e., (7.3). We proceed by induction on $k \in \mathbb{N}$. When $k = 1$, (7.3) reduces to the very definition of p-mixing.

Assume now that the result has been proved for $k - 1$ transformations. Since both sides of (7.3) are linear on each f_i, and for constant f_k the result follows by induction, we can reduce the statement to the case when $\int_X f_k \, d\mu = 0$. Now the right hand side of (7.3) vanishes, so all we need to show is that $p\text{-}\lim_g u_g = 0$, where $u_g = \prod_{i=1}^k T_{[1,i]}^g f_i$. In order to obtain this we will resort to DT, in the form of Proposition 7.2. Let $h \in G$ be an arbitrary element different from the identity.

$$\langle u_{gh}, u_g \rangle = \int_X \prod_{i=1}^k T_{[1,i]}^{gh} f_i \cdot T_{[1,i]}^g f_i \, d\mu$$

$$= \int_X \prod_{i=1}^{k} T^g_{[1,i]}\left(T^h_{[1,i]} f_i \cdot f_i\right) d\mu$$

$$= \int_X T^g_1\left(T^h_1 f_1 \cdot f_1 \cdot \prod_{i=2}^{k} T^g_{[2,i]}\left(T^h_{[1,i]} f_i \cdot f_i\right)\right) d\mu$$

$$= \int_X T^h_1 f_1 \cdot f_1 \cdot \prod_{i=2}^{k} T^g_{[2,i]}\left(T^h_{[1,i]} f_i \cdot f_i\right) d\mu.$$

Putting $\tilde{T}_i = T_{i+1}$ and $\tilde{f}_i = T^h_{[1,i+1]} f_{i+1} \cdot f_{i+1}$ for each $i = 1, \ldots, k-1$, observe that $\tilde{T}_{[i,j]} = T_{[i+1,j+1]}$. In particular, by induction we obtain that

$$p\text{-}\lim_g \prod_{i=2}^{k} T^g_{[2,i]}\left(T^h_{[1,i]} f_i \cdot f_i\right) = p\text{-}\lim_g \prod_{i=1}^{k-1} \tilde{T}^g_{[1,i]} \tilde{f}_i = \prod_{i=1}^{k-1} \int_X \tilde{f}_i \, d\mu$$

and hence we deduce that

$$p\text{-}\lim_g \langle u_{gh}, u_g \rangle = \int_X T^h_1 f_1 \cdot f_1 \cdot \prod_{i=1}^{k-1} \int_X \tilde{f}_i \, d\mu = \prod_{i=1}^{k} \int_X T^h_{[1,i]} f_i \cdot f_i \, d\mu.$$

Finally, taking the p-lim in h we conclude that

$$p\text{-}\lim_h p\text{-}\lim_g \langle u_{gh}, u_g \rangle = \prod_{i=1}^{k}\left(\int_X f_i \, d\mu\right)^2 = 0$$

and the result follows from Proposition 7.2.

Now we prove that joint p-mixing implies (7.4). It follows directly from (7.3), putting all functions but one equal to the constant 1, that for any i one has $p\text{-}\lim_g T^g_{[1,i]} f = \int_X f \, d\mu$. In particular $T_{[1,i]}$ is p-mixing for every i. Next, let $1 < i \leq j \leq k$ and let $f \in L^\infty(X)$. We can assume, without loss of generality, that $\int_X f \, d\mu = 0$, and we need to show that $\tilde{f} := p\text{-}\lim_g T^g_{[i,j]} f$ is equal to 0 (where the limit is in the weak topology of $L^2(X)$).

Take $f_j = f$, $f_{i-1} = \tilde{f}$ and $f_r = 1$ for all the other r. Then from (7.3) we obtain

$$0 = \int_X f \, d\mu \int_X \tilde{f} \, d\mu = p\text{-}\lim_g \prod_{r=1}^{k} T^g_{[1,r]} f_r = p\text{-}\lim_g T^g_{[1,i-1]}\left(\tilde{f} \cdot T^g_{[i,j]} f\right) = \int_X (\tilde{f})^2 \, d\mu.$$

We conclude that indeed $\tilde{f} = 0$. Observe that in order to invoke (7.3) we used implicitly the (easily checkable) fact that $\tilde{f} \in L^\infty$. $\square$

8. A topological variant of the difference theorem

As we have seen in previous sections, several versions of DT are useful in the study of multiple recurrence in ergodic theory. The phenomenon of (multiple) recurrence can also be studied from a topological standpoint. In this setting there is no good analog of ergodic averages, and hence the Hilbertian DT (Theorem 2.12) does not apply. Nevertheless, one can still adapt the fundamental idea of complexity reduction present in DT to this setup. This method was developed in [16], with a somewhat related ideas appearing earlier in [27].

Definition 8.1. Let $\mathcal{F}$ denote the family of all finite non-empty subsets of $\mathbb{N}$. Given an increasing sequence $(n_k)_{k \in \mathbb{N}}$ we define the *IP-set generated by* (n_k) to be the (image of the) map $n : \mathcal{F} \to \mathbb{N}$ defined by $n_\alpha = \sum_{k \in \alpha} n_k$. A set $A \subset \mathbb{N}$ is called an IP* *set* if it has nonempty intersection with every IP-set.

A famous theorem of Hindman [46] states that for any finite partition of an IP-set, one of the cells of the partition contains an IP-set. Equivalently, it states that a finite intersection of IP* sets is still an IP* set.

In this section we will deal with topological dynamical systems, namely with pairs (X, T) where X is a compact Hausdorff space and $T : X \to X$ is a homeomorphism. The system (X, T) (or the map T) is called *minimal* if there is no non-empty compact subset $Y \subset X$ such that $T^{-1}Y \subset Y$.

Definition 8.2. A family of sequences $A = \{a_1(n), \ldots, a_k(n)\}$ of integers is a *family of multiple IP topological recurrence* if for any minimal system (X, T), any non-empty open set $U \subset X$ and any IP set $(n_\alpha)_{\alpha \in \mathcal{F}}$, there exists $\alpha \in \mathcal{F}$ such that

$$U \cap T^{-a_1(n_\alpha)}U \cap \cdots \cap T^{-a_k(n_\alpha)}U \neq \varnothing.$$

Equivalently, A is a family of multiple IP topological recurrence if the set

$$\{n \in \mathbb{Z} : U \cap T^{-a_1(n)}U \cap \cdots \cap T^{-a_k(n)}U \neq \varnothing\}$$

is an IP* set.

Remark 8.3. It is an easy consequence of Hindman's theorem that whenever $\{a_1(n), \ldots, a_k(n)\}$ and $\{b_1(n), \ldots, b_\ell(n)\}$ are families of multiple IP topological recurrence, (X, T), (Y, S) are minimal systems and $U \subset X$, $V \subset Y$ are open sets, there exists $n \in \mathbb{N}$ such that

$$U \cap T^{-a_1(n)}U \cap \cdots \cap T^{-a_k(n)}U \neq \varnothing \neq V \cap S^{b_1(n)}V \cap \cdots \cap S^{-b_\ell(n)}V.$$

In fact, the set of such n is IP*.

Multiple IP topological recurrence has strong connections with combinatorics. This connection between topological dynamics and Ramsey theory was established by Furstenberg and Weiss in [41]. Using the methods developed in [41] one can show that if $\{a_1(n), \ldots, a_k(n)\}$ is a family of multiple IP topological recurrence, then for any finite partition of the natural numbers $\mathbb{N} = C_1 \cup \cdots \cup C_r$ there exist $i \in \{1, \ldots, r\}$, $x \in C_i$ and $n \in \mathbb{N}$ such that $x + a_j(n) \in C_i$ for every $j = 1, \ldots, k$. Moreover, n can be chosen from any prescribed IP-set.

Definition 8.4. A subset $S \subset \mathbb{N}$ is *syndetic* if it has bounded gaps[7]; in other words if for some $n \in \mathbb{N}$, the union $S \cup (S - 1) \cup \cdots \cup (S - n) = \mathbb{N}$. A set $T \subset \mathbb{N}$ is *thick* if it contains arbitrarily long intervals. A set $A \subset \mathbb{N}$ is called *piecewise syndetic* if it is the intersection $A = S \cap T$ of a syndetic set S and a thick set T.

For any finite partition of a piecewise syndetic set, one of the cells is piecewise syndetic. More generally, for any semigroup G, a set $S \subset G$ is called (left) *syndetic* if there exists a finite set $F \subset G$ such that $F^{-1}S := \{g \in G : \exists x \in F \text{ s.t. } xg \in S\} = G$. A set $T \subset G$ is called (right) *thick* if it has non-empty intersection with every syndetic set and a set $A \subset G$ is called piecewise syndetic if it equals the intersection of a syndetic set and a thick set.

---

[7] Note that a set $S \subset \mathbb{N}$ is syndetic if and only if it has positive lower Banach density (cf. Definition 4.1).

One can show that if $\{a_1(n), \ldots, a_k(n)\}$ is a family of multiple IP topological recurrence and $A \subset \mathbb{N}$ is a piecewise syndetic set, then there exist $x \in A$ and $n \in \mathbb{N}$ such that $\{x, x + a_1(n), \ldots, x + a_k(n)\} \subset A$. Moreover, n can be chosen from any prescribed IP-set.

The family $\{a(n)\}$ consisting the single degenerated sequence $a(n) \equiv 0$ is trivially a family of multiple IP topological recurrence. The following key lemma, which is the analog of DT in this setting, will allow us to quickly obtain some interesting applications:

Lemma 8.5 (*Reduction of Complexity (IP Version)*). *Let* $\{a_1(n), \ldots, a_s(n)\}$, $n \in \mathbb{N}$ *be a finite family of sequences of positive integers and put* $a_\ell(0) = 0$ *for all* $\ell = 1, \ldots, s$. *If for any finite set* $F \subset \mathbb{N} \cup \{0\}$ *the family of sequences*

$$\{n \mapsto a_\ell(n + h) - a_1(n) - a_\ell(h) \mid h \in F, \ \ell \in \{1, \ldots, s\}\}$$

is a family of multiple IP topological recurrence, then so is the family $\{a_1(n), \ldots, a_k(n)\}$, $n \in \mathbb{N}$.

Proof. Let X be a compact Hausdorff space, let $T : X \to X$ be a minimal homeomorphism and let $U \subset X$ be a non-empty open set. Since the system is minimal, there exists some $r \in \mathbb{N}$ such that $U \cup T^{-1}U \cup T^{-2}U \cup \cdots \cup T^{-r}U = X$. Let $(n_\alpha)_{\alpha \in \mathcal{F}}$ be an arbitrary IP-set in $\mathbb{N}$. Let $\alpha_1 \in \mathcal{F}$ be arbitrary, let $U_1 = U$ and $t_1 = 0$. For each $k \in \mathbb{N}$ we will construct inductively $\alpha_k \in \mathcal{F}$, $t_k \in \{0, \ldots, r\}$ and $U_k \subset T^{-t_k}U$ a non-empty open set such that

$$T^{-a_\ell(n_{\alpha_{j+1} \cup \alpha_{j+2} \cup \cdots \cup \alpha_k})}U_j \supset U_k \quad \text{for each } j < k \text{ and } \ell \in \{1, \ldots, s\}. \tag{8.1}$$

For simplicity, whenever $j \leq k$, denote $n_{\alpha_{j+1} \cup \alpha_{j+2} \cup \cdots \cup \alpha_k}$ by $m_{j,k}$, with the understanding that $m_{k,k} = 0$. Assume we have already chosen α_i, t_i, U_i for $i < k$. Consider the family of sequences of the form $n \mapsto a_\ell(n + m_{j,k-1}) - a_1(n) - a_\ell(m_{j,k-1})$ for each $j \in \{1, \ldots, k-2\}$ and $\ell \in \{1, \ldots, s\}$. By hypothesis this is a family of multiple IP topological recurrence, so there exists some $\alpha_k \in \mathcal{F}$, disjoint from $\bigcup_{i<k} \alpha_i$ and such that

$$V_k := \bigcap_{\ell=1}^{s} \bigcap_{j=1}^{k-1} T^{-\left(a_\ell(n_{\alpha_k} + m_{j,k-1}) - a_1(n_{\alpha_k}) - a_\ell(m_{j,k-1})\right)}U_{k-1} \neq \varnothing. \tag{8.2}$$

Next choose $t_k \in \{0, \ldots, r\}$ such that $U_k := T^{-a_0(n_{\alpha_k})}V_k \cap T^{-t_k}U \neq \varnothing$. To see how (8.1) follows from (8.2), let $x \in U_k$, let $\ell \in \{1, \ldots, s\}$ and let $j < k$. If $j = k - 1$ then the first inclusion below degenerates to a trivial equality. In any case we have:

$$T^{-a_\ell(m_{j,k})}U_j = T^{-a_1(n_{\alpha_k})}T^{-\left(a_\ell(m_{j,k}) - a_1(n_{\alpha_k}) - a_\ell(m_{j,k-1})\right)}T^{-a_\ell(m_{j,k-1})}U_j$$

$$\supset T^{-a_1(n_{\alpha_k})}T^{-\left(a_\ell(m_{j,k}) - a_1(n_{\alpha_k}) - a_\ell(m_{j,k-1})\right)}U_{k-1}$$

$$\supset T^{-a_1(n_{\alpha_k})}V_k$$

$$\supset U_k.$$

Finally, take $k > j$ such that $t_k = t_j =: t$ (such a pair must exist since each t_i belongs to the finite set $\{0, \ldots, r\}$). From (8.1) it follows that $T^{-a_\ell(m_{j,k})}U_j \supset U_k$ for each $\ell \in \{1, \ldots, s\}$, which implies that $U_k \subset (T^{-t}U) \cap T^{-a_\ell(m_{j,k})}(T^{-t}U) = T^{-t}(U \cap T^{-a_\ell(m_{j,k})}U)$ and therefore

$$\varnothing \neq T^t(U_k) \subset U \cap \bigcap_{\ell=1}^{s} T^{-a_\ell(m_{j,k})}U. \quad \square$$

Remark 8.6. One can state and prove the previous lemma (using the same method) for the case of regular (i.e., not IP) recurrence, weakening both the assumptions and the conclusion.

We now present some applications of Lemma 8.5.

Theorem 8.7 (*Topological IP Van der Waerden Theorem (Cf. [41, Theorem 3.2])*). *Let (X, T) be a minimal system and let $U \subset X$ be a non-empty open set. Then for any $k \in \mathbb{N}$ and any IP-set $(n_\alpha)_{\alpha \in \mathcal{F}}$ there exists $\alpha \in \mathcal{F}$ such that*

$$U \cap T^{-n_\alpha} U \cap T^{-2n_\alpha} \cap \cdots \cap T^{-kn_\alpha} U \neq \varnothing.$$

Proof. We prove the claim by induction on k. The case $k = 0$ is vacuously true. Next assume the result is true for $k - 1$. We want to show that the family of sequences $\{n, 2n, \ldots, kn\}$ is a family of multiple IP topological recurrence. Letting $a_\ell(n) = \ell n$ we see that $a_\ell(n+h) - a_1(n) - a_\ell(h) = (\ell - 1)n$ for any $h \in \mathbb{N} \cup \{0\}$. Therefore for any finite subset $F \subset \mathbb{N} \cup \{0\}$ we have

$$\left\{ n \mapsto a_\ell(n+h) - a_1(n) - a_\ell(h) \mid h \in F, \ \ell \in \{1, \ldots, s\} \right\} \subset \{0, n, 2n, \ldots, (k-1)\ell\}.$$

By the induction hypothesis, this is a family of multiple IP topological recurrence. It follows from Lemma 8.5 that $\{n, 2n, \ldots, kn\}$ is also a family of multiple IP topological recurrence, which finishes the proof. $\square$

As a corollary we obtain van der Waerden's celebrated theorem [62] stating that for any finite partition $\mathbb{N} = C_1 \cup \cdots \cup C_r$, one of the cells C_i contains arbitrarily long arithmetic progressions (Theorem 8.7 actually implies that any piecewise syndetic set contains arbitrarily long arithmetic progressions). Moreover, the common difference can be chosen from any prescribed IP-set.

Theorem 8.8 (*Topological IP Sárközy Theorem (Cf. Section 1.3 in [16])*). *Let (X, T) be a minimal system and let $U \subset X$ be a non-empty open set. Then for any IP-set $(n_\alpha)_{\alpha \in \mathcal{F}}$ there exists $\alpha \in \mathcal{F}$ such that*

$$U \cap T^{-n_\alpha^2} U \neq \varnothing.$$

Proof. We need to show that the family $\{n^2\}$ consisting of a single sequence $a_1(n) = n^2$, $n \in \mathbb{N}$ is a family of IP topological recurrence; we will use Lemma 8.5 to that end. Let $F \subset \mathbb{N} \cup \{0\}$ be an arbitrary finite set and let $k = 2 \max F$. One can easily check that

$$\left\{ n \mapsto a_1(n+h) - a_1(n) - a_1(h) \mid h \in F \right\} \subset \{0, n, 2n, \ldots, kn\}.$$

It follows from Theorem 8.7 that this is a family of multiple IP topological recurrence. Applying Lemma 8.5 we conclude that $\{n^2\}$ is a family of IP topological recurrence as desired. $\square$

A combinatorial corollary of Theorem 8.8 states that any piecewise syndetic set $A \subset \mathbb{N}$ contains a pair of the form $\{x, x + n^2\}$ (this result also follows from a stronger statement first obtained independently by Furstenberg [38] and Sárközy [57], which involves sets of positive upper density in $\mathbb{N}$). In addition, n can be chosen from any prescribed IP-set.

A finitistic analog of IP-sets is IP_s-sets, defined as follows: let $s \in \mathbb{N}$ and denote by $\mathcal{F}_s$ the family of non-empty subsets of $\{1, \ldots, s\}$. Given an increasing sequence $(n_k)_{k=1}^s$ of natural numbers, define the map $n : \mathcal{F}_s \to \mathbb{N}$ by the formula $n_\alpha = \sum_{k \in \alpha} n_k$. The IP_s-set generated by (n_k) is the set $\{n_\alpha : \alpha \in \mathcal{F}_s\}$. A set $A \subset \mathbb{N}$ that intersects non-trivially any IP_s set is called an IP_s^* set. There exists an analog of Hindman's theorem, stating that the

intersection of two IP$_s^*$-sets contains an IP$_r^*$-subset, for some r which depends only on s (and in particular the intersection is non-empty; for a proof of this fact see, for instance, [25]).

Using the PET induction scheme and applying Lemma 8.5 repeatedly, one can prove the general IP polynomial van der Waerden theorem, i.e., the fact that any family $\{a_1(n), \ldots, a_k(n)\}$ where each $a_i \in \mathbb{Z}[x]$ and $a_i(0) = 0$ is a family of multiple IP topological recurrence (see [16]). One actually has the following ostensibly stronger theorem (see [16] and [18, Theorem 7]).

Theorem 8.9. *Let X be a compact Hausdorff space, let $k \in \mathbb{N}$, let $T_1, \ldots, T_k : X \to X$ be commuting minimal homeomorphisms, and let $U \subset X$ be a non-empty open set. Then for any $p_1, \ldots, p_k \in \mathbb{Z}[x]$ with $p_i(0) = 0$ there exists $s \in \mathbb{N}$ such that the set*

$$\left\{ n \in \mathbb{N} : U \cap T_1^{-p_1(n)} U \cap \cdots \cap T_k^{-p_k(n)} U \neq \varnothing \right\}$$

is IP$_s^*$

On the other hand, even the family consisting of the single function $a(n) = \lfloor n^\alpha \rfloor$ where $\alpha \in (0, 1)$ is *not* a family of multiple IP topological recurrence. In order to study recurrence along such sequences we will make use of the notion of central* sets in a semigroup. We refer the reader to [12, Definition 3.1 (b)] for a definition of central sets. A subset of a semigroup is called a central* set if it has non-empty intersection with every central set. The only properties of central* sets that we will need are listed in the following proposition.

Proposition 8.10. (1) *If A, B are central*, then so is the intersection $A \cap B$.*
(2) *If $\phi : S \to R$ is a semigroup isomorphism and $C \subset S$ is central*, then so is $\phi(C) \subset R$.*
(3) *[13, Theorem 3.5] If $A \subset \mathbb{N}$ is an IP$_s^*$ set, then A is central* in the semigroup $(\mathbb{N}, \times)$.*

We prove below the following theorem:

Theorem 8.11. *Let $k, m \in \mathbb{N}$, let X be a compact Hausdorff space, let $k \in \mathbb{N}$, let $T_1, \ldots, T_k : X \to X$ be commuting minimal homeomorphisms, let $U \subset X$ be a nonempty open set and, for each $i \in \{1, \ldots, k\}$ let $a_i \in \mathbb{Z}[x]$ with $a_i(0) = 0$ and let $b_i(n) = \left\lfloor a_i\left(n^{1/m}\right) \right\rfloor$. Then the set*

$$\left\{ n : U \cap T_1^{-b_1(n)} U \cap \cdots \cap T_k^{-b_k(n)} U \neq \varnothing \right\}$$

is central* in the semigroup of mth powers, $\left(\{n^m : n \in \mathbb{N}\}, \times\right)$.

Corollary 8.12. *Let $k, t, m \in \mathbb{N}$. For each $j \in \{1, \ldots, t\}$ let X_j be a compact Hausdorff space, let $T_{j,1}, \ldots, T_{j,k} : X_j \to X_j$ be commuting minimal homeomorphisms, let $U_j \subset X_j$ be open and, for each $i \in \{1, \ldots, k\}$ let $a_i^{(j)} \in \mathbb{Z}[x]$ with $a_i^{(j)}(0) = 0$ and let $b_i^{(j)}(n) = \left\lfloor a_i^{(j)}\left(n^{1/m}\right) \right\rfloor$. Then there exists $n \in \mathbb{N}$ such that for each $j \in \{1, \ldots, t\}$ we have*

$$U_j \cap T_{j,1}^{-b_1^{(j)}(n)} U_j \cap \cdots \cap T_{j,k}^{-b_k^{(j)}(n)} U_j \neq \varnothing.$$

Here is a combinatorial application of Corollary 8.12.

Corollary 8.13. *Let $k, t, m \in \mathbb{N}$. For each $j \in \{1, \ldots, t\}$ let $A_j \subset \mathbb{N}^k$ be a piecewise syndetic set and, for each $i \in \{1, \ldots, k\}$ let $a_i^{(j)} \in \mathbb{Z}[x]$ with $a_i^{(j)}(0) = 0$ and let $b_i^{(j)}(n) = \left\lfloor a_i^{(j)}\left(n^{1/m}\right) \right\rfloor$.*

Then there exists $n \in \mathbb{N}$ such that for each $j \in \{1, \ldots, t\}$ there exists $x_j \in \mathbb{N}^k$ such that

$$\left\{ x_j + b_i^{(j)}(n) \cdot e_i : i \in \{1, \ldots, k\} \right\} \subset A_j$$

where $\{e_1, \ldots, e_k\}$ form the canonical basis for $\mathbb{N}^d$.

Proof of Theorem 8.11. Let

$$R_a := \left\{ n : \bigcap_{i=0}^{k} T^{-a_i(n)} U \neq \varnothing \right\} \quad \text{and} \quad R_b := \left\{ n : \bigcap_{i=0}^{k} T^{-b_i(n)} U \neq \varnothing \right\}.$$

It follows from Theorem 8.9 that R_a is an IP_s^* set for some $s \in \mathbb{N}$. Hence, by [13, Theorem 3.5] we deduce that R_a is central* in the semigroup $(\mathbb{N}, \times)$. Let $R = \{n^m : n \in R_a\}$.

It is clear that $R \subset R_b$. It thus suffices to show that R is a central* subset of the semigroup $\left(\{n^m : n \in \mathbb{N}\}, \times \right)$. Indeed, the map $n \mapsto n^m$ is a semigroup isomorphism between $(\mathbb{N}, \times)$ and $\left(\{n^m : n \in \mathbb{N}\}, \times \right)$, and hence the image of any central* set is again a central* set in the image semigroup. □

We conclude this section (and the paper) with a concrete example illustrating Corollary 8.13.

Example 8.14. For any piecewise syndetic sets $A, B \subset \mathbb{N}$ there exist $x, y, n \in \mathbb{N}$ such that

$$\left\{ x, x + \lfloor \sqrt{n} \rfloor, x + \lfloor n^{7/2} \rfloor \right\} \subset A \qquad \left\{ y, y + \lfloor n^{5/2} \rfloor, y + \lfloor n^{7/2} \rfloor \right\} \subset B.$$

References

[1] T. Austin, Norm convergence of continuous-time polynomial multiple ergodic averages, Ergodic Theory Dynam. Systems 32 (2) (2012) 361–382.

[2] J. Bass, Suites uniformément denses, moyennes trigonométriques, fonctions pseudo-aléatoires, Bol. Soc. Math. France 87 (1959) 1–64.

[3] A. Bellow, V. Losert, The weighted pointwise ergodic theorem and the individual ergodic theorem along subsequences, Trans. Amer. Math. Soc. 288 (1) (1985) 307–345.

[4] D. Berend, V. Bergelson, Jointly ergodic measure-preserving transformations, Israel J. Math. 49 (4) (1984) 307–314.

[5] Vitaly Bergelson, Ergodic Ramsey theory, in: Logic and Combinatorics (Arcata, Calif., 1985), in: Contemp. Math., vol. 65, Amer. Math. Soc., Providence, RI, 1987, pp. 63–87.

[6] V. Bergelson, Weakly mixing PET, Ergodic Theory Dynam. Systems 7 (3) (1987) 337–349.

[7] V. Bergelson, Ergodic Ramsey theory—an update, in: Ergodic Theory of $\mathbf{Z}^d$ Actions, in: London Math. Soc. Lecture Note Ser., vol. 228, Cambridge Univ. Press, Cambridge, 1996, pp. 1–61.

[8] V. Bergelson, Combinatorial and diophantine applications of ergodic theory, in: B. Hasselblatt, A. Katok (Eds.), Handbook of Dynamical Systems, Vol. 1B, Elsevier, 2006, pp. 745–841.

[9] V. Bergelson, Ultrafilters, IP sets, dynamics, and combinatorial number theory, in: Ultrafilters Across Mathematics, in: Contemp. Math., vol. 530, Amer. Math. Soc., Providence, RI, 2010, pp. 23–47.

[10] V. Bergelson, H. Furstenberg, R. McCutcheon, IP-sets and polynomial recurrence, Ergodic Theory Dynam. Systems 16 (5) (1996) 963–974.

[11] V. Bergelson, I.J. Håland Knutson, Weak mixing implies weak mixing of higher orders along tempered functions, Ergodic Theory Dynam. Systems 29 (5) (2009) 1375–1416.

[12] V. Bergelson, N. Hindman, Nonmetrizable topological dynamics and Ramsey theory, Trans. Amer. Math. Soc. 320 (1) (1990) 293–320.

[13] V. Bergelson, N. Hindman, On IP* sets and central sets, Combinatorica 14 (3) (1994) 269–277.

[14] V. Bergelson, S. Kasjan, M. Lemańczyk, Polynomial actions of unitary operators and idempotent ultrafilters. submitted for publication. Available at: http://arxiv.org/abs/1401.7869.

[15] V. Bergelson, G. Kolesnik, Y. Son, Uniform distribution of subpolynomial functions along primes and applications, J. Anal. Math., in press. Available at: http://arxiv.org/abs/1503.04960.

[16] V. Bergelson, A. Leibman, Polynomial extensions of van der Waerden's and Szemerédi's theorems, J. Amer. Math. Soc. 9 (3) (1996) 725–753.

[17] V. Bergelson, A. Leibman, Y. Son, Joint ergodicity along generalized linear functions, Ergodic Theory Dynam. Systems (2015) in press. Available on CJO2015. http://dx.doi.org/10.1017/etds.2015.11.

[18] V. Bergelson, A. Leibman, T. Ziegler, The shifted primes and the multidimensional Szemerédi and polynomial van der Waerden theorems, C. R. Math. Acad. Sci. Paris 349 (3–4) (2011) 123–125.

[19] V. Bergelson, E. Lesigne, Van der Corput sets in $\mathbb{Z}^d$, Colloq. Math. 110 (1) (2008) 1–49.

[20] V. Bergelson, R. McCutcheon, An ergodic IP polynomial Szemerédi theorem, Mem. Amer. Math. Soc. 146 (695) (2000) viii+106.

[21] V. Bergelson, R. McCutcheon, Central sets and a non-commutative Roth theorem, Amer. J. Math. 129 (5) (2007) 1251–1275.

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

[23] V. Bergelson, J. Moreira, Ergodic theorem involving additive and multiplicative groups of a field and $\{x + y, xy\}$ patterns, Ergodic Theory Dynam. Systems (2015) in press. Available at: http://journals.cambridge.org/abstract_S0143385715000681.

[24] V. Bergelson, J. Rosenblatt, Joint ergodicity for group actions, Ergodic Theory Dynam. Systems 8 (3) (1988) 351–364.

[25] V. Bergelson, D. Robertson, Polynomial recurrence with large intersection over countable fields, Israel J. Math., in press. Available at: http://arxiv.org/abs/1409.6774.

[26] A.S. Besicovitch, Almost Periodic Functions, Dover Publications, Inc., New York, 1955.

[27] A. Błaszczyk, S. Plewik, S. Turek, Topological multidimensional van der Waerden theorem, Comment. Math. Univ. Carolin. 30 (4) (1989) 783–787.

[28] M. Boshernitzan, Uniform distribution and Hardy fields, J. Anal. Math. 62 (1994) 225–240.

[29] D. Cantor, On the elementary theory of diophantine approximation over the ring of adeles. I, Illinois J. Math. 9 (1965) 677–700.

[30] J. Cigler, The fundamental theorem of van der Corput on uniform distribution and its generalizations, Compos. Math. 16 (1964) 29–34.

[31] J. Cigler, A characterization of well-distributed sequences, Compos. Math. 17 (1966) 263–267.

[32] J. Cigler, Some remarks on the distribution mod1 of tempered sequences, Nieuw Arch. Wiskd. (3) 16 (1968) 194–196.

[33] van der J.G. Corput, Diophantische Ungleichungen. I. Zur Gleichverteilung Modulo Eins, Acta Math. 56 (1) (1931) 373–456.

[34] L. Corwin, C. Pfeffer, On the density of sets in $(\mathbf{A}/\mathbf{Q})^n$ defined by polynomials, Colloq. Math. 68 (1) (1995) 1–5.

[35] H. Daboussi, H. Delange, On multiplicative arithmetical functions whose modulus does not exceed one, J. Lond. Math. Soc. (2) 26 (2) (1982) 245–264.

[36] H. Daboussi, M. Mendès France, Spectrum, almost-periodicity and equidistribution modulo 1, Studia Sci. Math. Hungar. 9 (1974) 173–180.

[37] A.F.M. ter Elst, V. Müller, A van der Corput-type lemma for power bounded operators. Preprint.

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

[39] H. Furstenberg, Recurrence in Ergodic Theory and Combinatorial Number Theory, Princeton University Press, Princeton, NJ, 1981.

[40] H. Furstenberg, Y. Katznelson, An ergodic Szemerédi theorem for IP-systems and combinatorial theory, J. Anal. Math. 45 (1985) 117–168.

[41] H. Furstenberg, B. Weiss, Topological dynamics and combinatorial number theory, J. Anal. Math. 34 (1978) 61–85.

[42] Hillel Furstenberg, Benjamin Weiss, The finite multipliers of infinite ergodic transformations, in: The Structure of Attractors in Dynamical Systems (Proc. Conf., North Dakota State Univ., Fargo, N.D., 1977), in: Lecture Notes in Math., vol. 668, Springer, Berlin, 1978, pp. 127–132.

[43] B. Green, T. Tao, The Möbius function is strongly orthogonal to nilsequences, Ann. of Math. (2) 175 (2) (2012) 541–566.

[44] F.P. Greenleaf, Ergodic theorems and the construction of summing sequences in amenable locally compact groups, Comm. Pure Appl. Math. 26 (1973) 29–46.

[45] I. Håland, Uniform distribution of generalized polynomials, J. Number Theory 45 (3) (1993) 327–366.

[46] N. Hindman, Finite sums from sequences within cells of a partition of N, J. Combin. Theory Ser. A 17 (1974) 1–11.

[47] N. Hindman, D. Strauss, Algebra in the Stone–Čech Compactification, Walter de Gruyter & Co., Berlin, 1998.

[48] T. Kamae, M. Mendès France, Van der Corput's difference theorem, Israel J. Math. 31 (3–4) (1978) 335–342.

[49] I. Kátai, A remark on a theorem of H. Daboussi, Acta Math. Hungar. 47 (1–2) (1986) 223–225.

[50] J.H.B. Kemperman, On the distribution of a sequence in a compact group, Compos. Math. 16 (1964) 138–157.

[51] L. Kuipers, H. Niederreiter, Uniform Distribution of Sequences, John Wiley & Sons, New York, London, Sydney, 1974.

[52] M. Mendès France, Les suites à spectre vide et la répartition modulo 1, J. Number Theory 5 (1973) 1–15.

[53] von J. Neumann, Proof of the quasi-ergodic hypothesis, Proc. Natl. Acad. Sci. 18 (3) (1932) 70–82.

[54] V. Paulsen, Completely Bounded Maps and Operator Algebras, in: Cambridge Studies in Advanced Mathematics, vol. 78, Cambridge University Press, Cambridge, 2002.

[55] G. Rauzy, Propriétés Statistiques de Suites Arithmétiques, in: Le Mathématicien, Collection SUP, vol. 15, Presses Universitaires de France, Paris, 1976.

[56] I.Z. Ruzsa, Connections between the uniform distribution of a sequence and its differences, in: Topics in Classical Number Theory (Budapest, 1981), in: Colloq. Math. Soc. János Bolyai, vol. 34, North-Holland, Amsterdam, 1984, pp. 1419–1443.

[57] A. Sárközy, On difference sets of sequences of integers. I, Acta Math. Acad. Sci. Hungar. 31 (1–2) (1978) 125–149.

[58] C. Schnell, Idempotent ultrafilters and polynomial recurrence. Available at: http://arxiv.org/abs/0711.0484.

[59] J. Spilker, Die Fastperiodizität der Watson-Funktion, Arch. Math. (Basel) 74 (1) (2000) 26–29.

[60] T. Tao, Higher Order Fourier Analysis, in: Graduate Studies in Mathematics, vol. 142, Amer. Math. Soc., Providence, RI, 2012.

[61] R.J. Taschner, Der Differenzensatz von van der Corput und gleichverteilte Funktionen, J. Reine Angew. Math. 307–308 (1979) 325–340.

[62] van der B.L. Waerden, Beweis einer baudetschen vermutung, Nieuw Arch. Wiskd. 15 (1927) 212–216.

[63] H. Weyl, Über die Gleichverteilung von Zahlen mod. Eins, Math. Ann. 77 (3) (1916) 313–352.