Ergod. Th. & Dynam. Sys. (2019), **39**, 2332–2383 © *Cambridge University Press, 2018*
doi:10.1017/etds.2017.130

Rationally almost periodic sequences, polynomial multiple recurrence and symbolic dynamics

V. BERGELSON†, J. KUŁAGA-PRZYMUS‡, M. LEMAŃCZYK‡ and F. K. RICHTER†

† *Department of Mathematics, Ohio State University, Columbus, OH 43210, USA*
(e-mail: vitaly@math.ohio-state.edu, richter.109@osu.edu)
‡ *Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, Chopina 12/18, 87-100 Toruń, Poland*
(e-mail: joanna.kulaga@gmail.com, mlem@mat.umk.pl)

(*Received* 23 *November* 2016 *and accepted in revised form* 13 *September* 2017)

Abstract. A set $R \subset \mathbb{N}$ is called *rational* if it is well approximable by finite unions of arithmetic progressions, meaning that for every $\epsilon > 0$ there exists a set $B = \bigcup_{i=1}^{r} a_i\mathbb{N} + b_i$, where $a_1, \ldots, a_r, b_1, \ldots, b_r \in \mathbb{N}$, such that

$$\overline{d}(R \triangle B) := \limsup_{N\to\infty} \frac{|(R \triangle B) \cap \{1, \ldots, N\}|}{N} < \epsilon.$$

Examples of rational sets include many classical sets of number-theoretical origin such as the set of squarefree numbers, the set of abundant numbers, or sets of the form $\Phi_x := \{n \in \mathbb{N} : \boldsymbol{\varphi}(n)/n < x\}$, where $x \in [0, 1]$ and $\boldsymbol{\varphi}$ is Euler's totient function. We investigate the combinatorial and dynamical properties of rational sets and obtain new results in ergodic Ramsey theory. Among other things, we show that if $R \subset \mathbb{N}$ is a rational set with $\overline{d}(R) > 0$, then the following are equivalent:

(a) R is divisible, i.e. $\overline{d}(R \cap u\mathbb{N}) > 0$ for all $u \in \mathbb{N}$;
(b) R is an averaging set of polynomial single recurrence;
(c) R is an averaging set of polynomial multiple recurrence.

As an application, we show that if $R \subset \mathbb{N}$ is rational and divisible, then for any set $E \subset \mathbb{N}$ with $\overline{d}(E) > 0$ and any polynomials $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, which satisfy $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$ for all $i \in \{1, \ldots, \ell\}$, there exists $\beta > 0$ such that the set

$$\{n \in R : \overline{d}(E \cap (E - p_1(n)) \cap \cdots \cap (E - p_\ell(n))) > \beta\}$$

has positive lower density.

Ramsey-theoretical applications naturally lead to problems in symbolic dynamics, which involve *rationally almost periodic sequences* (sequences whose level-sets are

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

rational). We prove that if $\mathcal{A}$ is a finite alphabet, $\eta \in \mathcal{A}^{\mathbb{N}}$ is rationally almost periodic, S denotes the left-shift on $\mathcal{A}^{\mathbb{Z}}$ and

$$X := \{y \in \mathcal{A}^{\mathbb{Z}} : \text{each word appearing in } y \text{ appears in } \eta\},$$

then η is a generic point for an S-invariant probability measure ν on X such that the measure-preserving system (X, ν, S) is ergodic and has rational discrete spectrum.

Contents

1 Introduction 2333
2 Rationality and recurrence 2338
2.1 Rational sequences are good weights for polynomial multiple convergence 2338
2.2 Averaging single recurrence 2340
2.3 Averaging multiple recurrence 2343
2.4 Inner regular sets, W-rational sets and $\mathscr{B}$-free numbers 2346
3 Rational dynamical systems 2352
3.1 Definition and examples of rational subshifts 2352
3.2 Rationality along subsequences 2356
3.3 Generalizing Theorem 1.7 2357
3.4 Revisiting Theorem 1.4 2364
4 Applications to combinatorics 2369
5 Rational sequences and Sarnak's conjecture 2371
5.1 Synchronized automata and substitutions 2371
5.2 Orthogonality of RAP and WRAP sequences to the Möbius function 2372
Acknowledgements 2377
A Appendix. Uniformity of polynomial multiple recurrence 2377
References 2381

1. *Introduction*

The celebrated Szemerédi theorem on arithmetic progressions [**58**] states that any set $S \subset \mathbb{N}$ having positive upper density $\overline{d}(S) = \limsup_{N\to\infty} |S \cap \{1, \ldots, N\}|/N > 0$ contains arbitrarily long arithmetic progressions. A (one-dimensional special case of a) polynomial generalization of Szemerédi's theorem proved in [**12**] states that for any $S \subset \mathbb{N}$ with $\overline{d}(S) > 0$ and any polynomials $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, which satisfy $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$ for all $i \in \{1, \ldots, \ell\}$, the set S contains (many) polynomial progressions of the form $\{a, a + p_1(n), \ldots, a + p_\ell(n)\}$. The proof of the polynomial extension of Szemerédi's theorem given in [**12**] is obtained with the help of an ergodic approach introduced by Furstenberg (see [**32**, **33**]). In particular, the one-dimensional polynomial Szemerédi theorem formulated above follows from the fact that for any probability space $(X, \mathcal{B}, \mu)$, any invertible measure-preserving transformation $T : X \to X$, any $A \in \mathcal{B}$ with $\mu(A) > 0$ and any ℓ polynomials $p_i \in \mathbb{Q}[t]$ satisfying $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$, $i \in \{1, \ldots, \ell\}$,

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

there exist arbitrarily large $n \in \mathbb{N}$ such that $\mu(A \cap T^{-p_1(n)}A \cap \cdots \cap T^{-p_\ell(n)}A) > 0$. As a matter of fact, one can show† that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-p_1(n)}A \cap \cdots \cap T^{-p_\ell(n)}A) > 0. \tag{1.1}$$

One of the goals of this paper is to refine (1.1) by considering multiple ergodic averages of the from

$$\lim_{N\to\infty} \frac{1}{|R \cap [1, N]|} \sum_{n=1}^{N} \mathbb{1}_R(n)\mu(A \cap T^{-p_1(n)}A \cap \cdots \cap T^{-p_\ell(n)}A), \tag{1.2}$$

for certain sets R of arithmetic origin called *rational sets*, which were introduced in [**13**] (see Definition 1.1 below). We show that for any rational set R the limit in (1.2) exists. Furthermore, we give necessary and sufficient conditions on R for this limit to be positive. This, in turn, allows us to obtain new refinements of the polynomial Szemerédi theorem, some of which we state at the end of this introduction.

To present the main results of our paper we need to introduce some definitions first.

Definition 1.1. (Rationally almost periodic sequences and rational sets) Let $\mathcal{A}$ be a finite set. We endow the space $\mathcal{A}^{\mathbb{N}}$ with the Besicovitch pseudo-metric d_B (cf. [**15**, **16**]),

$$d_B(x, y) := \limsup_{N\to\infty} \frac{|\{1 \leqslant n \leqslant N : x(n) \neq y(n)\}|}{N}. \tag{1.3}$$

A sequence $x \in \mathcal{A}^{\mathbb{N}}$ is called *(Besicovitch) rationally almost periodic* (RAP) if for every $\varepsilon > 0$ there exists a periodic sequence $y \in \mathcal{A}^{\mathbb{N}}$ such that $d_B(x, y) < \varepsilon$. (A more general definition of the Besicovitch pseudo-metric d_B and of rationally almost periodic sequences will be introduced in §2.1 (see p. 2338) and in §3.2 (see Definition 3.7).)

A set $R \subset \mathbb{N}$ is called *rational* if the sequence $\mathbb{1}_R$ (viewed as a sequence in $\{0, 1\}^{\mathbb{N}}$) is RAP; see [**13**, Definition 2.1].

Here are some examples of rational sets:

- the set Q of *squarefree numbers* (see [**13**, Lemma 2.7]);
- the set $\mathscr{A}$ of *abundant numbers*‡ and the set $\mathscr{D}$ of *deficient numbers* (see Corollary 2.17);
- for any $x \in [0, 1]$, the set $\Phi_x := \{n \in \mathbb{N} : \boldsymbol{\varphi}(n)/n < x\}$, where $\boldsymbol{\varphi}$ is Euler's totient function (see also Corollary 2.17).

† We remark that the original proof in [**12**] established only that

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-p_1(n)}A \cap \cdots \cap T^{-p_\ell(n)}A) > 0,$$

whereas the existence of the limit in (1.1) was obtained later; see [**38**, **45**].

‡ Let $\sigma(n) = \sum_{d|n} d$ denote the classical *sum of divisors function*. The set of *abundant numbers* and the set of *deficient numbers* are defined, respectively, as $\mathscr{A} := \{n \in \mathbb{N} : \sigma(n) > 2n\}$ and $\mathscr{D} := \{n \in \mathbb{N} : \sigma(n) < 2n\}$. (The classical set of *perfect numbers* is defined as $\mathscr{P} := \{n \in \mathbb{N} : \sigma(n) = 2n\}$.)

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

The above examples are special cases of *sets of multiples* and sets of $\mathscr{B}$*-free numbers*. For $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$ the corresponding sets of multiples and $\mathscr{B}$-free numbers are defined as $\mathcal{M}_{\mathscr{B}} := \bigcup_{b\in\mathscr{B}} b\mathbb{N}$ and $\mathcal{F}_{\mathscr{B}} := \mathbb{N} \setminus \mathcal{M}_{\mathscr{B}}$, respectively. The abundant numbers (as well as the union of the abundant and perfect numbers) form a set of multiples, the deficient numbers yield an example of a $\mathscr{B}$-free set and Φ_x is a set of multiples. In §2.4 we show that for $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$ the set $\mathcal{F}_{\mathscr{B}}$ is a rational set if and only if the *density* $d(\mathcal{F}_{\mathscr{B}}) := \lim_{N\to\infty} (1/N)|\mathcal{F}_{\mathscr{B}} \cap [1, N]|$ exists (see Corollary 2.16).

A natural way of obtaining rational sets is via level-sets of RAP sequences: if $\mathcal{A} = \{a_1, a_2, \ldots, a_r\}$ is a finite set and $x \in \mathcal{A}^{\mathbb{N}}$ is an RAP sequence then the sets $\{n \in \mathbb{N} : x(n) = a_1\}, \ldots, \{n \in \mathbb{N} : x(n) = a_r\}$ are rational. As a matter of fact, $x \in \mathcal{A}^{\mathbb{Z}}$ is RAP if and only if all its level-sets are rational. Examples of RAP sequences include *regular Toeplitz sequences*, or, more generally, *Weyl rationally almost periodic sequences* (for definitions see §3.1). In particular, *paperfolding sequences*† as well as *automatic sequences coming from synchronized automata* are RAP sequences (see §§3.1 and 5 for definitions and more details).

Definition 1.2. (Cf. [**10**, Definition 1.5]) We say that $R \subset \mathbb{N}$ is an *averaging set of polynomial multiple recurrence* if for any invertible measure-preserving system $(X, \mathcal{B}, \mu, T)$, $A \in \mathcal{B}$ with $\mu(A) > 0$, $\ell \in \mathbb{N}$ and any polynomials $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, with $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$ for all $i \in \{1, \ldots, \ell\}$, the limit in (1.2) exists and is positive. If $\ell = 1$ then we speak of an *averaging set of polynomial single recurrence*.

An averaging set of (single or multiple) polynomial recurrence $R \subset \mathbb{N}$ must also be a *set of recurrence*, i.e. for each measure-preserving system $(X, \mathcal{B}, \mu, T)$ and each $A \in \mathcal{B}$ with $\mu(A) > 0$ there exists $n \in R$ such that $\mu(A \cap T^{-n}A) > 0$. If we assume that the density $d(R) = \lim_{N\to\infty} (1/N)|R \cap [1, N]|$ exists and is positive then it follows—by considering cyclic rotations on finitely many points—that the density of $R \cap u\mathbb{N}$ also exists and is positive for any positive integer u. This divisibility property is a rather trivial but necessary condition for a positive density set to be 'good' for averaging recurrence. This leads to the following definition.

Definition 1.3. Let $R \subset \mathbb{N}$. We say that R is *divisible* if $d(R \cap u\mathbb{N})$ exists and is positive for all $u \in \mathbb{N}$.

Note that for rational sets the existence of $d(R)$ and $d(R \cap u\mathbb{N})$ is automatic (cf. Lemma 3.14 below). Therefore, to verify divisibility, it suffices to check the positivity of $d(R \cap u\mathbb{N})$ for all $u \in \mathbb{N}$.

One of our main theorems asserts that for rational sets divisibility is not only a necessary but also a sufficient condition for averaging recurrence.

THEOREM 1.4. *Let $R \subset \mathbb{N}$ be a rational set and assume $d(R) > 0$. The following are equivalent:*

(a) *R is divisible;*

† Given an infinite binary sequence $i \in \{0, 1\}^{\mathbb{N}}$, we inductively define the *paperfolding sequence* $t \in \{0, 1\}^{\mathbb{N}}$ with *folding instructions* $i(1), i(2), i(3), \ldots$ as follows: set $t(1) := i(1)$ and, whenever $t(n)$ has already been defined for $n \in \{1, 2, \ldots, 2^k - 1\}$, define $t(n)$ for $n \in \{2^k, 2^k + 1, \ldots, 2^{k+1} - 1\}$ as $t(2^k) := i(k)$ and $t(n) := t(2^{k+1} - n)$ for $2^k < n < 2^{k+1}$. For more information on paperfolding sequences, see [**1**, **22**].

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

(b) *R is an averaging set of polynomial single recurrence;*
(c) *R is an averaging set of polynomial multiple recurrence.*

It was proved in [**13**] that every self-shift of the set Q of squarefree numbers, i.e. any set of the form $Q - r$ for $r \in Q$, is divisible and hence satisfies the hypothesis of Theorem 1.4. Moreover, it follows from [**13**] that a shift $Q - r$ for $r \in \mathbb{N}$ is divisible if and only if $r \in Q$. The following theorem establishes a result of similar nature for sets of $\mathscr{B}$-free numbers.

THEOREM 1.5. *Let $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$ and assume that $d(\mathcal{F}_{\mathscr{B}})$ exists and is positive. Then there exists a set $D \subset \mathcal{F}_{\mathscr{B}}$ with $d(\mathcal{F}_{\mathscr{B}} \setminus D) = 0$ such that the set $\mathcal{F}_{\mathscr{B}} - r$ is an averaging set of polynomial multiple recurrence if and only if $r \in D$.*

Remark 1.6. A detailed discussion of criteria for the existence of $d(\mathcal{F}_{\mathscr{B}})$ will be provided in §2.4; see Definition 2.14 and Theorem 2.15. In §3.4 we obtain a version of Theorem 1.5 for the case where $d(\mathcal{F}_{\mathscr{B}})$ does not necessarily exist; see Theorem 3.27.

In §2.4 we also show that in Theorem 1.5 one has $D = \mathcal{F}_{\mathscr{B}}$ if and only if the set $\mathscr{B}$ is *taut*; see Definition 2.19 and Theorem 2.26.

Theorem 1.4 motivates closer interest in RAP sequences as independent objects. In §3 we take a dynamical approach to study RAP sequences more closely. To formulate our results in this direction, let us first recall some basic notions of symbolic dynamics.

As before, let $\mathcal{A}$ be a finite set (alphabet) and let $S\colon \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ denote the left-shift on $\mathcal{A}^{\mathbb{Z}}$, i.e. $Sx = y$ where $x \in \mathcal{A}^{\mathbb{Z}}$ and $y(n) = x(n+1)$ for all $n \in \mathbb{Z}$. For $x \in \mathcal{A}^{\mathbb{Z}}$ (or $x \in \mathcal{A}^{\mathbb{N}}$) and $n < m$ we call $x[n, m] = (x(n), x(n+1), \ldots, x(m))$ a *word appearing in x*. Given $\eta \in \mathcal{A}^{\mathbb{N}}$, let

$$\begin{aligned} X_\eta &:= \{x \in \mathcal{A}^{\mathbb{Z}} : (\forall n < m)(\exists k \in \mathbb{N})\, x[n, m] = \eta[k, k+m-n-1]\} \\ &= \{x \in \mathcal{A}^{\mathbb{Z}} : \text{each word appearing in } x \text{ appears in } \eta\}. \end{aligned}$$

Clearly, X_η is a closed and S-invariant subset of $\mathcal{A}^{\mathbb{Z}}$ (usually referred to as the *subshift determined by η*)†. A sequence $\eta \in \mathcal{A}^{\mathbb{N}}$ is called *generic* for an S-invariant Borel probability measure μ on $\mathcal{A}^{\mathbb{Z}}$ if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(S^n \tilde{\eta}) = \int_{\mathcal{A}^{\mathbb{Z}}} f \, d\mu$$

for all continuous functions $f \in C(\mathcal{A}^{\mathbb{Z}})$, where $\tilde{\eta} \in \mathcal{A}^{\mathbb{Z}}$ denotes any two-sided sequence extending $\eta \in \mathcal{A}^{\mathbb{N}}$. Note that the above definition does not depend on the choice of the two-sided extension $\tilde{\eta}$ of η.

For an RAP sequence $\eta \in \mathcal{A}^{\mathbb{N}}$ we call the corresponding symbolic dynamical system (X_η, S) a *rational subshift*. We show in §3 that any rational sequence η is generic for an ergodic measure ν such that (X_η, ν, S) has rational discrete spectrum (i.e. the span of all eigenfunctions of T is dense in $L^2(X, \mathcal{B}, \mu)$ and all the corresponding eigenvalues are roots of unity), a result which we believe is of independent interest.

that the of has

† When η is (topologically) recurrent, that is, any finite word appearing in η reappears infinitely often, then there is $\widetilde{\eta} \in \mathcal{A}^{\mathbb{Z}}$ such that $\widetilde{\eta}[1, \infty) = \eta$ and $X_\eta = \overline{\{S^k \widetilde{\eta} : k \in \mathbb{Z}\}}$ (cf. [**25**, pp. 189–190]). Note, however, that not all RAP sequences are recurrent.

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

THEOREM 1.7. *Let $\eta \in \mathcal{A}^{\mathbb{N}}$ be RAP. Then there exists an S-invariant Borel probability measure ν on X_η such that η is generic for ν and the measure-preserving system (X_η, ν, S) is ergodic and has rational discrete spectrum.*

In light of Theorem 1.7, the following result (obtained in §3.4) can be viewed as a 'dynamical' generalization of Theorem 1.4.

THEOREM 1.8. *Let $R \subset \mathbb{N}$ with $d(R) > 0$ and suppose $\eta := \mathbb{1}_R$ is generic for a Borel probability measure ν on $X_\eta \subset \{0, 1\}^{\mathbb{Z}}$ such that (X_η, ν, S) has rational discrete spectrum. Then there exists an increasing sequence of natural numbers $(N_k)_{k\geqslant 1}$ such that the following are equivalent.*

(A) *R is divisible along $(N_k)_{k\geqslant 1}$, that is, for all $u \in \mathbb{N}$,*

$$d^{(N_k)}(R \cap u\mathbb{N}) := \lim_{k\to\infty} \frac{|R \cap u\mathbb{N} \cap \{1, \ldots, N_k\}|}{N_k} > 0.$$

(B) *R is an averaging set of polynomial multiple recurrence along $(N_k)_{k\geqslant 1}$, that is, for all invertible measure-preserving systems $(X, \mathcal{B}, \mu, T)$, $\ell \in \mathbb{N}$, $A \in \mathcal{B}$ with $\mu(A) > 0$ and for all polynomials $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, with $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$ for $i \in \{1, \ldots, \ell\}$,*

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{n=1}^{N_k} \mathbb{1}_R(n)\mu(A \cap T^{-p_1(n)}A \cap \cdots \cap T^{-p_\ell(n)}A) > 0.$$

In §§2.4 and 3.4 we give various examples of (classes of) rational sets for which Theorems 1.4 and 1.8 hold.

With the help of Furstenberg's correspondence principle (see Proposition 4.1) we have the following combinatorial corollary of Theorem 1.4.

THEOREM 1.9. *Let $R \subset \mathbb{N}$ be rational and divisible. Then for any set $E \subset \mathbb{N}$ with $\overline{d}(E) > 0$ and any polynomials $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, which satisfy $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$ for all $i \in \{1, \ldots, \ell\}$, there exists $\beta > 0$ such that the set*

$$\{n \in R : \overline{d}(E \cap (E - p_1(n)) \cap \cdots \cap (E - p_\ell(n))) > \beta\}$$

has positive lower density.

We note that Theorem 1.5 also yields combinatorial corollaries in the spirit of Theorem 1.9, which are formulated and proved in §4.

We conclude this introduction by stating an amplified version of Theorem 1.9, a proof of which is also contained in §4.

THEOREM 1.10. *Let $R \subset \mathbb{N}$ be rational and divisible. Then for any $E \subset \mathbb{N}$ with $\overline{d}(E) > 0$ and any polynomials $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, which satisfy $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$, for all $i \in \{1, \ldots, \ell\}$, there exists a subset $R' \subset R$ satisfying $\overline{d}(R') > 0$ and such that for any finite subset $F \subset R'$,*

$$\overline{d}\Big(\bigcap_{n\in F} (E \cap (E - p_1(n)) \cap \cdots \cap (E - p_\ell(n)))\Big) > 0.$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Structure of the paper. Section 2 is divided into four subsections. In §2.1 we show that RAP sequences are good weights for polynomial multiple convergence. In §2.2, we prove the equivalence (a) $\Leftrightarrow$ (b) of Theorem 1.4. In §2.3 we give a proof of the equivalence (a) $\Leftrightarrow$ (c). Finally, in §2.4, we provide more examples of rational sets, and discuss some of their properties. This includes a discourse on $\mathscr{B}$-free numbers and a proof of Theorem 1.5.

In §3 we define rational subshifts and study their dynamical properties. In particular, §3 contains a proof of Theorem 1.7.

In §3.4 we give a proof of a strengthening of Theorem 1.4, and in §4 we provide various combinatorial applications of it via Furstenberg's correspondence principle.

In §5 we prove that systems generated by Weyl rationally almost periodic sequences (see p. 2354 for the definition) satisfy Sarnak's conjecture.

Finally, in the appendix we establish a uniform version of the polynomial multiple recurrence theorem obtained in [**12**], which is needed for the proof of Theorem 1.4.

2. *Rationality and recurrence*

2.1. *Rational sequences are good weights for polynomial multiple convergence.*

The purpose of this subsection is to show that for rational sets R with $d(R) > 0$, the limit in (1.2) always exists.

First, we make the following observation. If $d(R)$ exists and is positive then the limit in (1.2) exists and is positive if and only if the limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_R(n)\mu(A \cap T^{-p_1(n)}A \cap \cdots \cap T^{-p_\ell(n)}A) \tag{2.1}$$

exists and is positive. Since throughout this paper we mostly consider sets R for which $d(R)$ exists (except in §3.4) and is positive, it suffices to study the ergodic averages given by (2.1) instead of (1.2).

For the special case where $\ell = 1$ and $p_1(t) = t$, the existence of the limit in (2.1) follows from the work of Bellow and Losert in [**6**]. To better describe what is known in this case, we need to introduce first the following extended form of Definition 1.1.

Definition 2.1. Given $x, y\colon \mathbb{N} \to \mathbb{C}$, we define

$$d_B(x, y) := \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} |x(n) - y(n)|. \tag{2.2}$$

A sequence $x\colon \mathbb{N} \to \mathbb{C}$ is called *Besicovitch almost periodic* (BAP) [**15**, **16**] if for every $\varepsilon > 0$ there exists a trigonometric polynomial $P(t) = \sum_{j=1}^{M} c_j e^{2\pi i \lambda_j t}$ with $c_1, \ldots, c_M \in \mathbb{C}$ and $\lambda_1, \ldots, \lambda_M \in \mathbb{R}$ such that $d_B(x, P) = d_B((x(n))_{n\in\mathbb{N}}, (P(n))_{n\in\mathbb{N}}) < \varepsilon$. If, for each $\varepsilon > 0$, one can choose $\lambda_1, \ldots, \lambda_M \in \mathbb{Q}$ – which is equivalent to the assertion that the sequence $(P(n))$ is periodic – then we call x *(Besicovitch) rationally almost periodic*, or RAP. In particular, RAP sequences are a special type of BAP sequences.

It is shown in [**6**, §3] that for any bounded BAP sequence $x\colon \mathbb{N} \to \mathbb{C}$, the ergodic averages

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} x(n)T^n f$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

converge almost everywhere for any function $f \in L^1(X, \mathcal{B}, \mu)$. From this, the existence of the limit in (2.1) for $\ell = 1$ and $p_1(t) = t$ follows immediately.

Definition 2.2. A sequence $x \in \{0, 1\}^{\mathbb{N}}$ is called a *good weight for polynomial multiple convergence* if for every invertible measure-preserving system $(X, \mathcal{B}, \mu, T)$, for all $f_1, \ldots, f_\ell \in L^\infty(X, \mu)$ and for all polynomials $p_i \in \mathbb{Q}[t]$, $p_i(\mathbb{Z}) \subset \mathbb{Z}$, $i \in \{1, \ldots, \ell\}$, the limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} x(n) \prod_{i=1}^{\ell} T^{p_i(n)} f_i \tag{2.3}$$

exists in $L^2(X, \mathcal{B}, \mu)$.

The following proposition shows that the limit in (2.1) exists in general.

PROPOSITION 2.3. *Let $x \in \{0, 1\}^{\mathbb{N}}$ be RAP. Then x is a good weight for polynomial multiple convergence.*

Proof. It follows from the results of Host and Kra [**38**] and Leibman [**45**] that the sequence

$$\frac{1}{N} \sum_{n=1}^{N} T^{q_1(n)} f_1 \cdot \ldots \cdot T^{q_\ell(n)} f_\ell, \quad N \geqslant 1,$$

converges in L^2, for any $q_i \in \mathbb{Q}[t]$, $q_i(\mathbb{Z}) \subset \mathbb{Z}$, $i = 1, \ldots, \ell$. In particular, given arbitrary $a \in \mathbb{N}$, $b \in \mathbb{Z}$, the averages

$$\frac{1}{N} \sum_{n=1}^{N} T^{p_1(an+b)} f_1 \cdot \ldots \cdot T^{p_\ell(an+b)} f_\ell$$

converge in L^2 as $N \to \infty$. Equivalently, the limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{a\mathbb{N}+b}(n) T^{p_1(n)} f_1 \cdot \ldots \cdot T^{p_\ell(n)} f_\ell \tag{2.4}$$

exists. Observe that any periodic sequence can be written as a finite linear combination of infinite arithmetic progressions $\mathbb{1}_{a\mathbb{N}+b}$. Therefore, it follows from (2.4) that for any periodic sequence $y \in \{0, 1\}^{\mathbb{N}}$ the limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} y(n) T^{p_1(n)} f_1 \cdot \ldots \cdot T^{p_\ell(n)} f_\ell$$

also exists in L^2.

Since any RAP sequence x can be approximated by periodic sequences, we can find periodic sequences y_m, $m \in \mathbb{N}$, satisfying $d_B(y_m, x) \to 0$ as $m \to \infty$. Define

$$L_m := \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} y_m(n) T^{p_1(n)} f_1 \cdot \ldots \cdot T^{p_\ell(n)} f_\ell.$$

Then

$$\|L_{m_1} - L_{m_2}\|_{L^2} \leqslant d_B(y_{m_1}, y_{m_2}) \|f_1\|_{L^\infty} \cdot \ldots \cdot \|f_\ell\|_{L^\infty},$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

which shows that (L_m) is a Cauchy sequence, whence the limit $L := \lim_{m\to\infty} L_m$ exists. Moreover,

$$\limsup_{N\to\infty} \left\| L_m - \frac{1}{N}\sum_{n=1}^{N} x(n) T^{p_1(n)} f_1 \cdot \ldots \cdot T^{p_\ell(n)} f_\ell \right\|_{L^2}$$

can be bounded from above by $d_B(x, y_m)\|f_1\|_{L^\infty} \cdots \|f_\ell\|_{L^\infty}$, which converges to zero as $m \to \infty$. Therefore, the limit in (2.3) exists and equals L. □

2.2. *Averaging single recurrence.* In this subsection we provide a proof of the equivalence (a) ⇔ (b) in Theorem 1.4. Of course, this equivalence is a special case of the more general equivalence (a) ⇔ (c). We include a separate proof of this simpler case because, on the one hand, this proof is more elementary and self-contained and, on the other hand, it contains in embryonic form the ideas needed for the proof of the general case. Let us state the non-trivial implication, namely (a) ⇒ (b), as an independent theorem.

THEOREM 2.4. *Assume that $R \subset \mathbb{N}$ is rational and divisible. Then R is an averaging set of polynomial single recurrence.*

The proof of Theorem 2.4 is comprised of two parts. First, we prove the assertion for totally ergodic systems. Recall that $(X, \mathcal{B}, \mu, T)$ is called *totally ergodic* if T^m is ergodic for all $m \in \mathbb{N}$. Equivalently, T is ergodic and the spectrum of the unitary operator associated with T contains no non-trivial roots of unity.

Lemma 2.5 below, which is the second ingredient in the proof of Theorem 2.4, allows us to reduce the case of general ergodic systems to those which are totally ergodic. This is done by replacing $T^{p(n)}$ with $T^{p(un)}$ for a highly divisible natural number u. Since $p(0) = 0$, this allows us to identify $T^{p(un)}$ with $T^{q(n)}$ for some other polynomial q. This procedure annihilates the rational part of the spectrum in the sense that will be made precise below.

In the following, we use $\mathcal{K}_{\text{rat}}$ to denote the *rational Kronecker factor* of $(X, \mathcal{B}, \mu, T)$, which is defined as the smallest sub-σ-algebra of $\mathcal{B}$ for which all eigenfunctions with roots of unity as eigenvalues are measurable. Equivalently, the rational Kronecker factor is the largest factor of T which has rational discrete spectrum. It is also a characteristic factor for ergodic averages along polynomials. This means that for any function $f \in L^2$ and any polynomial $p \in \mathbb{Q}[t]$, $p(\mathbb{Z}) \subset \mathbb{Z}$,

$$\lim_{N\to\infty} \left\| \frac{1}{N}\sum_{n=1}^{N} (T^{p(n)} f - T^{p(n)} \mathbb{E}(f|\mathcal{K}_{\text{rat}})) \right\|_{L^2} = 0, \tag{2.5}$$

where $\mathbb{E}(f|\mathcal{K}_{\text{rat}})$ denotes the *conditional expectation of f with respect to $\mathcal{K}_{\text{rat}}$*, i.e. the unique function in $L^2(X, \mathcal{B}, \mu)$ such that $\mathbb{E}(f|\mathcal{K}_{\text{rat}})$ is $\mathcal{K}_{\text{rat}}$-measurable and $\int_A \mathbb{E}(f|\mathcal{K}_{\text{rat}})\, d\mu = \int_A f\, d\mu$ for all $A \in \mathcal{K}_{\text{rat}}$. A proof of (2.5) can be found in [**9**, §2].

LEMMA 2.5. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible measure-preserving system and let $R \subset \mathbb{N}$ with $d(R) > 0$. Also, let $p \in \mathbb{Q}[t]$ satisfy $p(\mathbb{Z}) \subset \mathbb{Z}$ and $p(0) = 0$. Assume that for each real-valued $g \in L^2(X, \mathcal{B}, \mu)$ with $\mathbb{E}(g|\mathcal{K}_{\text{rat}}) = \int g\, d\mu > 0$ there exists some $\delta > 0$ such that*

$$\overline{d}(D_\delta(g) \cap u\mathbb{N}) > 0 \quad \textit{for all } u \in \mathbb{N}, \tag{2.6}$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

where $D_\delta(g) := \{n \in R : \langle T^{p(n)}g, g\rangle > \delta\}$. *Then for all non-negative* $f \in L^2(X, \mathcal{B}, \mu)$ *with* $\int_X f\, d\mu > 0$,

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_R(n)\langle T^{p(n)}f, f\rangle > 0. \tag{2.7}$$

Proof. Fix $f \in L^2(X, \mathcal{B}, \mu)$, $f \geqslant 0$ with $\int_X f\, d\mu > 0$. Then the function

$$g^{(1)} := f - \mathbb{E}(f|\mathcal{K}_{\text{rat}}) + \int_X f\, d\mu$$

is real-valued and satisfies $\mathbb{E}(g^{(1)}|\mathcal{K}_{\text{rat}}) = \int_X g^{(1)}\, d\mu > 0$. Therefore, we can find some $\delta > 0$ such that (2.6) holds for $g = g^{(1)}$. Pick $0 < \epsilon < \sqrt{\delta}$. Let $\mathcal{K}_u$ stand for the factor of $\mathcal{K}_{\text{rat}}$ that is generated by eigenfunctions corresponding to eigenvalues which are roots of unity of degree at most u. Note that $\mathcal{K}_{1!} \subset \mathcal{K}_{2!} \subset \mathcal{K}_{3!} \subset \ldots$ and

$$\mathcal{K}_{\text{rat}} = \bigvee_{m\in\mathbb{N}} \mathcal{K}_{m!}.$$

Hence, using Doob's martingale convergence theorem (see [**54**, §3.4]), we can find $m \in \mathbb{N}$ such that $\|\mathbb{E}(f|\mathcal{K}_{\text{rat}}) - \mathbb{E}(f|\mathcal{K}_{m!})\|_{L^2} < \epsilon$. Take $u = m!$ Define

$$g^{(2)} = \mathbb{E}(f|\mathcal{K}_u) - \int_X f\, d\mu,$$
$$g^{(3)} = \mathbb{E}(f|\mathcal{K}_{\text{rat}}) - \mathbb{E}(f|\mathcal{K}_u),$$

so that $f = g^{(1)} + g^{(2)} + g^{(3)}$. A simple calculation shows that

$$\langle T^n g^{(i)}, g^{(j)}\rangle = 0 \quad \text{for all } i, j \in \{1, 2, 3\} \text{ with } i \neq j \text{ and for all } n \in \mathbb{Z}.$$

It follows that

$$\begin{aligned}\langle T^{p(n)}f, f\rangle &= \langle T^{p(n)}(g^{(1)} + g^{(2)} + g^{(3)}), g^{(1)} + g^{(2)} + g^{(3)}\rangle \\ &= \langle T^{p(n)}g^{(1)}, g^{(1)}\rangle + \langle T^{p(n)}g^{(2)}, g^{(2)}\rangle + \langle T^{p(n)}g^{(3)}, g^{(3)}\rangle.\end{aligned}$$

Then, using $T^{p(n)}g^{(2)} = g^{(2)}$ for all $n \in u\mathbb{N}$ and $\|g^{(3)}\|_{L^2} < \varepsilon$, we get that for every $n \in D_\delta(g^{(1)}) \cap u\mathbb{N}$,

$$\begin{aligned}\langle T^{p(n)}f, f\rangle &= \langle T^{p(n)}g^{(1)}, g^{(1)}\rangle + \langle T^{p(n)}g^{(2)}, g^{(2)}\rangle + \langle T^{p(n)}g^{(3)}, g^{(3)}\rangle \\ &\geqslant \langle T^{p(n)}g^{(1)}, g^{(1)}\rangle + \langle g^{(2)}, g^{(2)}\rangle - \varepsilon^2 \\ &\geqslant \delta - \varepsilon^2 > 0.\end{aligned}$$

To complete the proof, it suffices to notice that

$$\begin{aligned}\limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_R(n)\langle T^{p(n)}f, f\rangle &\geqslant \limsup_{N\to\infty} \frac{1}{N} \sum_{n\in D_\delta(g^{(1)})\cap u\mathbb{N}\cap\{1,\ldots,N\}} \mathbb{1}_R(n)\langle T^{p(n)}f, f\rangle \\ &\geqslant (\delta - \varepsilon^2)\, \overline{d}(D_\delta(g^{(1)}) \cap u\mathbb{N})\end{aligned}$$

and apply (2.6) for $g^{(1)}$. □

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Proof of Theorem 2.4. Suppose $R \subset \mathbb{N}$ with $d(R) > 0$ is both rational and divisible. We want to show that R is a set of averaging polynomial single recurrence, i.e. we want to show that for any invertible measure-preserving system $(X, \mathcal{B}, \mu, T)$, $p \in \mathbb{Q}[t]$, $p(\mathbb{Z}) \subset \mathbb{Z}$, with $p(0) = 0$ and $A \in \mathcal{B}$ with $\mu(A) > 0$, the limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_R(n)\mu(A \cap T^{-p(n)}A) \tag{2.8}$$

is positive. Note that the limit in (2.8) exists by Proposition 2.3.

In view of Lemma 2.5, to show that (2.8) is positive it suffices to show that (2.6) holds for all real-valued $g \in L^2(X, \mathcal{B}, \mu)$ with $\mathbb{E}(g|\mathcal{K}_{\text{rat}}) = \int_X g\, d\mu > 0$. However, for any such g, it follows from (2.5) that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle T^{q(n)}g, g\rangle = \left(\int_X g\, d\mu\right)^2$$

for all polynomials $q \in \mathbb{Q}[t]$, $q(\mathbb{Z}) \subset \mathbb{Z}$. In particular, we can pick $q(n) = p(u(an + b))$ and obtain

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle T^{p(u(an+b))}g, g\rangle = \left(\int_X g\, d\mu\right)^2 \quad \text{for all } a, u \in \mathbb{N},\ b \in \mathbb{N} \cup \{0\}.$$

This can be rewritten as

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{a\mathbb{N}+b}(n)\langle T^{p(un)}g, g\rangle = \frac{1}{a}\left(\int_X g\, d\mu\right)^2 \quad \text{for all } a, u \in \mathbb{N},\ b \in \mathbb{N} \cup \{0\}. \tag{2.9}$$

Now, if $E \subset \mathbb{N}$ is a finite union of infinite arithmetic progressions then $\mathbb{1}_E$ can be written as a finite linear combination of functions of the form $\mathbb{1}_{a\mathbb{N}+b}$ and it follows from (2.9) that for any such set E,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_E(n)\langle T^{p(un)}g, g\rangle = d(E)\left(\int_X g\, d\mu\right)^2.$$

Finally, since $R \cap u\mathbb{N}$ is rational for all $u \in \mathbb{N}$ and every rational set can be approximated by finite unions of infinite arithmetic progressions, we deduce that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{R\cap u\mathbb{N}}(n)\langle T^{p(un)}g, g\rangle = d(R \cap u\mathbb{N})\left(\int_X g\, d\mu\right)^2 \quad \text{for all } u \in \mathbb{N}. \tag{2.10}$$

Choose $\delta > 0$ so that $\delta(1 + \|g\|_{L^2}^2) < (\int_X g\, d\mu)^2$. It is now an immediate consequence of (2.10) that

$$\overline{d}(\{n \in R \cap u\mathbb{N} : \langle T^{p(n)}g, g\rangle > \delta\}) \geqslant d(R \cap u\mathbb{N})\delta.$$

From this it follows that (2.6) holds. □

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

2.3. *Averaging multiple recurrence.* In this subsection we prove (a) $\Rightarrow$ (c) in Theorem 1.4. Since (c) $\Rightarrow$ (a) is trivial, this will complete the proof of Theorem 1.4. Let us state the implication that we want to prove as a separate theorem.

THEOREM 2.6. *Assume $R \subset \mathbb{N}$ is rational and divisible. Then R is an averaging set of polynomial multiple recurrence.*

For the proof of Theorem 2.6, we rely on a series of known results. We recall first some fundamental properties of nilsystems.

Let G be a nilpotent Lie group and let Γ be a uniform and discrete subgroup of G. The compact manifold $X := G/\Gamma$ is called a *nilmanifold*. G acts naturally on X. More precisely, if $g,\ y \in G$ and $x = y\Gamma \in X$ then $T_g x$ is defined as $(gy)\Gamma$. For a fixed $g \in G$ the topological dynamical system (X, T_g) is called a *nilsystem*. Every nilmanifold $X = G/\Gamma$ possesses a unique G-invariant probability measure μ_X, called the *Haar measure* of X.

A bounded function $\phi\colon \mathbb{N} \to \mathbb{C}$ is called a *basic nilsequence* if there exist a nilmanifold $X = G/\Gamma$, a point $x \in X$, an element $g \in G$ and a continuous function $f \in C(X)$ such that $\phi(n) = f(T_g^n x)$ for all $n \in \mathbb{N}$. Here, $T_g^n x$ coincides with $T_{g^n} x$. A function $\psi\colon \mathbb{N} \to \mathbb{C}$ is called a *nilsequence* if for each $\varepsilon > 0$ there exists a basic nilsequence $(\phi(n))$ such that $|\psi(n) - \phi(n)| < \varepsilon$ for all $n \in \mathbb{N}$.

An important tool in the proof of Theorem 2.6 is a theorem of Leibman that allows us to replace multiple ergodic averages along polynomials with Birkhoff sums of nilsequences.

THEOREM 2.7. (Cf. [**47**, Theorem 4.1] and [**46**, Proposition 3.14]) *Assume that $(X, \mathcal{B}, \mu, T)$ is an invertible measure-preserving system, let $f_1, \ldots, f_\ell \in L^\infty(X, \mathcal{B}, \mu)$, $p_1, \ldots, p_\ell \in \mathbb{Q}[t]$ $(p_i(\mathbb{Z}) \subset \mathbb{Z}$ for $i = 1, \ldots, \ell)$ and set $\varphi(n) := \int_X T^{p_1(n)} f_1 \cdot \ldots \cdot T^{p_\ell(n)} f_\ell \, d\mu$, $n \in \mathbb{Z}$. Then there exists a nilsequence $(\psi(n))$ such that*

$$\limsup_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} |\varphi(n) - \psi(n)| = 0.$$

In particular, $d_B(\varphi, \psi) = 0$.

If $(x_n)_{n\geqslant 1}$ is a sequence of points from a nilmanifold $X = G/\Gamma$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(x_n) = \int_X f \, d\mu_X$$

for all continuous functions $f \in C(X)$, then we call such a sequence *uniformly distributed*. If $(x_n)_{n\geqslant 1}$ has the property that $(x_{an+b})_{n\in\mathbb{N}}$ is uniformly distributed for all $a, b \in \mathbb{N}$, then we call this sequence *totally equidistributed*. It is well known that for any nilsystem (X, T_g) the following are equivalent (see, for instance, [**3**, **52**] in the case of connected G and [**46**] in the general case):

- the sequence $(T_g^n x)_{n\in\mathbb{N}}$ is totally equidistributed for all $x \in X$;
- the system (X, μ_X, T_g) is totally ergodic.

Any nilmanifold has finitely many connected components (and each such component is a sub-nilmanifold). Moreover, since any ergodic nilrotation T_g permutes these components in a cyclical fashion, we deduce that for some $u \in \mathbb{N}$ the nilrotation T_{g^u} fixes each

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

connected component. The next proposition asserts that in this case the action of T_{g^u} on each of these connected components is totally ergodic.

PROPOSITION 2.8. (See [**29**, Proposition 2.1]) *Let $X = G/\Gamma$ be a nilmanifold, $g \in G$ and assume that the nilrotation T_g is ergodic. Fix $x \in X$ and let Y denote the connected component of X containing x. Then there exists $u \in \mathbb{N}$ such that Y is T_{g^u}-invariant and (Y, μ_Y, T_{g^u}) is totally ergodic.*

The next lemma is important for the proof of Theorem 1.4 and asserts that linear sequences coming from totally ergodic nilrotations (or equivalently, totally equidistributed sequences) do not correlate with RAP sequences.

LEMMA 2.9. *Suppose $R \subset \mathbb{N}$ is rational and T_g is a totally ergodic nilrotation on a nilmanifold $X = G/\Gamma$. Then, for all $x \in X$ and $f \in C(X)$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_R(n) f(T_g^n x) = d(R) \int_X f \, d\mu_X.$$

Proof. Since T_g is totally ergodic, we deduce that the sequence $(T_g^n x)_{n\in\mathbb{N}}$ is totally equidistributed for each $x \in X$. Therefore, for all $a \in \mathbb{N}$ and $b \in \mathbb{N} \cup \{0\}$, we obtain

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T_g^{an+b} x) = \int_X f \, d\mu_X.$$

This can be rewritten as

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{a\mathbb{N}+b}(n) f(T_g^n x) = \frac{1}{a} \int_X f \, d\mu_X. \tag{2.11}$$

If $E \subset \mathbb{N}$ is a finite union of infinite arithmetic progressions then $\mathbb{1}_E$ can be written as a finite linear combination of functions of the form $\mathbb{1}_{a\mathbb{N}+b}$. It now follows directly from (2.11) that for any such set E,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_E(n) f(T_g^n x) = d(E) \int_X f \, d\mu_X. \tag{2.12}$$

Finally, since R is rational, it can be approximated in the d_B pseudo-metric by finite unions of infinite arithmetic progressions and so, using (2.12), we obtain

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_R(n) f(T_g^n x) = d(R) \int_X f \, d\mu_X. \qquad \square$$

Proof of Theorem 2.6. Let $(X, \mathcal{B}, \mu, T)$ be an invertible measure-preserving system and assume that $R \subset \mathbb{N}$ is rational and divisible. Take any $A \in \mathcal{B}$ with $\mu(A) > 0$ and let $p_1, \ldots, p_\ell \in \mathbb{Q}[t]$ with $p_i(\mathbb{Z}) \subset \mathbb{Z}$, $p_i(0) = 0$, $i = 1, \ldots, \ell$, be arbitrary. We will show that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_R(n) \varphi(n) > 0, \tag{2.13}$$

where $\varphi(n) = \mu(A \cap T^{-p_1(n)} A \cap \cdots \cap T^{-p_\ell(n)} A)$. This, in view of (2.1), suffices to conclude that R is an averaging set of polynomial multiple recurrence. The existence

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

of the limit in (2.13) follows immediately from Proposition 2.3, hence it only remains to show its positivity.

By Theorem A.2, there exists $\delta > 0$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \varphi(un) > \delta \quad \text{for all } u \in \mathbb{N}. \tag{2.14}$$

Using Theorem 2.7, we can find a nilsequence $(\psi(n))$ such that (2.13) holds if and only if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_R(n)\psi(n) > 0. \tag{2.15}$$

Moreover, since $d_B(\varphi, \psi) = 0$, it follows from (2.14) that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \psi(un) > \delta \quad \text{for all } u \in \mathbb{N}. \tag{2.16}$$

By definition, every nilsequence can be uniformly approximated by basic nilsequences. For us this means that there exist a nilpotent Lie group G, a uniform and discrete subgroup Γ, $x \in X = G/\Gamma$ and $f \in C(X)$ such that $|\psi(n) - f(T_g^n x)| \leqslant \delta/4$ for all $n \in \mathbb{N}$. We can assume without loss of generality that T_g is ergodic and, since $\varphi(n) \in [0, 1]$ and $d_B(\varphi, \psi) = 0$, that $0 \leqslant f \leqslant 1$. It follows from (2.16) and the inequalities $|\psi(n) - f(T_g^n x)| \leqslant \delta/4$, $n \in \mathbb{N}$, that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T_g^{un} x) > \frac{3\delta}{4} \quad \text{for all } u \in \mathbb{N}. \tag{2.17}$$

Using Proposition 2.8, we can find $u \in \mathbb{N}$ and a sub-nilmanifold $Y \subset X$ containing x such that (Y, μ_Y, T_{g^u}) is totally ergodic. In the following, we identify f with $f|_Y$. Since R is rational, it is straightforward that the set $R/u := \{n \in \mathbb{N} : nu \in R\}$ is also rational. Thus, we can invoke Lemma 2.9 and obtain

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{R/u}(n) f(T_{g^u}^n x) = d(R/u) \int_Y f \, d\mu_Y. \tag{2.18}$$

Finally, combining (2.18) with (2.17) (together with the ergodic theorem) and the fact that $|\psi(un) - f(T_{g^u}^n x)| \leqslant \delta/4$ for all $n \in \mathbb{N}$, we obtain

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_R(n)\psi(n) &\geqslant \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{R\cap u\mathbb{N}}(n)\psi(n) \\
&= \frac{1}{u}\left(\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{R/u}(n)\psi(un) \right) \\
&\geqslant \frac{1}{u}\left(\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{R/u}(n) f(T_{g^u}^n x) - \frac{\delta}{4} d(R/u) \right) \\
&\geqslant \frac{1}{u}\left(\frac{3\delta}{4} d(R/u) - \frac{\delta}{4} d(R/u) \right) > 0.
\end{aligned}$$

This completes the proof. □

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

We would like to pose the following question describing one possible way of extending Theorem 1.4 to a more general version involving several commuting measure-preserving transformations.

Question 2.10. Assume $R \subset \mathbb{N}$ is rational and $d(R) > 0$. Are the following equivalent?

(α) R is divisible.

(β) For all probability spaces $(X, \mathcal{B}, \mu)$, all $\ell \in \mathbb{N}$, all ℓ-tuples of commuting invertible measure-preserving transformations $T_1, \ldots, T_\ell$ on $(X, \mathcal{B}, \mu)$, all $A \in \mathcal{B}$ with $\mu(A) > 0$ and for all polynomials $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, with $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_R(n)\mu(A \cap T_1^{-p_1(n)}A \cap \cdots \cap T_\ell^{-p_\ell(n)}A) > 0.$$

2.4. *Inner regular sets, W-rational sets and $\mathcal{B}$-free numbers.* The set Q of squarefree numbers is rational (see Corollary 2.16 below) but it is not divisible, as $Q \cap p^2\mathbb{N} = \emptyset$ for all primes p. In particular, Q is not a set of recurrence. However, as was mentioned in §1, it follows from results obtained in [**13**] that $Q - r$ is divisible (and hence—by virtue of Theorem 1.4—an averaging set of polynomial multiple recurrence) if and only if $r \in Q$.

This raises the question whether every rational set can be shifted to become divisible. In general, the answer to this question is negative. For example, one can show that for a carefully chosen increasing sequence $a_0, a_1, a_2, \ldots \in \mathbb{N}$, the set $S = \mathbb{N} \setminus \bigcup_{n\geqslant 0}(a_n\mathbb{N} + n)$ is rational. On the other hand, for any integer $n \geqslant 0$, one has $(S - n) \cap a_n\mathbb{N} = \emptyset$ (cf. [**36**, Theorem 11.6] and [**5**, Theorem 2.20]).

We will now introduce a rather natural family of rational sets with the property that for any set in this family there is a shift that is divisible.

Definition 2.11. We define the *Weyl pseudo-metric* d_W on $\{0, 1\}^{\mathbb{N}}$ as

$$d_W(x, y) = \limsup_{N\to\infty} \sup_{\ell\geqslant 1} \frac{1}{N} |\{\ell \leqslant n \leqslant \ell + N : x(n) \neq y(n)\}|. \tag{2.19}$$

A set $R \subset \mathbb{N}$ is called *W-rational* if $\mathbb{1}_R \in \{0, 1\}^{\mathbb{N}}$ can be approximated by periodic sequences in the d_W pseudo-metric.

Note that every W-rational set is a rational set.

In §3.1 below we will extend the definition of the d_W pseudo-metric from $\{0, 1\}^{\mathbb{N}}$ to $\mathcal{A}^{\mathbb{N}}$ for arbitrary finite subsets $\mathcal{A}$ and we also introduce the related notion of Weyl rationally almost periodic sequences (see page 2354).

PROPOSITION 2.12. *Suppose $D \subset \mathbb{N}$ is W-rational and $d(D) > 0$. There exists a shift of D which is divisible.*

Proof. Assume that no translation of D is divisible. Hence, for each $n \geqslant 0$, there exists $w_n \geqslant 1$ such that if $C_n := \{s \in \mathbb{N} : n + sw_n \in D^c\}$ then

$$d(C_n) = 1. \tag{2.20}$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Fix $K \geqslant 1$. Then by (2.20), also

$$d(\{s \geqslant 0 : n + sw_1 \cdot \cdots \cdot w_K \in D^c \text{ for each } n = 0, 1, \ldots, K-1\}) = 1.$$

It follows that for every $K \geqslant 1$ there exists $s \geqslant 1$ such that $\mathbb{1}_D(n + sw_1 \cdot \cdots \cdot w_K) = 0$ for each $n = 0, 1, \ldots, K-1$. In other words, in the sequence $\mathbb{1}_D$ there appear arbitrarily long blocks of consecutive zeros. This implies that the only periodic sequence that approximates $\mathbb{1}_D$ in the d_W pseudo-metric is the sequence $(0, 0, 0, \ldots)$, which contradicts $d(D) > 0$. □

In order to give more examples of rational sets that possess shifts that are divisible, we will now recall the notion of inner regular sets (see [**13**, Definition 2.3]). A subset $R \subset \mathbb{N}$ is called *inner regular* if for each $\varepsilon > 0$ there exists $m \geqslant 1$ such that for each $a \in \mathbb{N} \cup \{0\}$ the intersection $R \cap (m\mathbb{Z} + a)$ is either empty or has lower density greater than $(1 - \varepsilon)/m$. It follows immediately that every inner regular set is rational. Also, it is shown in [**13**, Lemma 2.7] that the set of squarefree numbers Q is inner regular.

PROPOSITION 2.13. *Assume that $\emptyset \neq R \subset \mathbb{N}$ is inner regular. Then, for each $r \in R$, the set $R - r$ is divisible.*

Proof. Suppose $u \geqslant 1$ is arbitrary. Fix $\varepsilon > 0$ with $\varepsilon < 1/u$. We can find $m \geqslant 1$ so that for every $a \in \mathbb{N} \cup \{0\}$ the intersection $R \cap (m\mathbb{Z} + a)$ is either empty or has lower density greater than $(1 - \varepsilon)/m$. Since $r \in R$, the intersection $R \cap (m\mathbb{Z} + r)$ is not empty. This means that the set $\{k \in \mathbb{N} : mk + r \in R\} = \{k \in \mathbb{N} : mk \in R - r\}$ has lower density greater than $1 - \varepsilon$. Since $\varepsilon < 1/u$, the set $\{k \in \mathbb{N} : mk \in R - r\} \cap u\mathbb{N}$ has positive lower density. This means the set $(R - r) \cap um\mathbb{N}$ has positive lower density and the assertion follows. □

We move the discussion now on to sets of $\mathscr{B}$-free numbers. The purpose of the remainder of this section is to prove a general form of Theorem 1.5 formulated in §1.

Given $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$, we consider its *set of multiples* $\mathcal{M}_{\mathscr{B}} := \bigcup_{b \in \mathscr{B}} b\mathbb{N}$ and the corresponding set of *$\mathscr{B}$-free numbers* $\mathcal{F}_{\mathscr{B}} := \mathbb{N} \setminus \mathcal{M}_{\mathscr{B}}$, i.e. the set of integers without a divisor in $\mathscr{B}$. Without loss of generality, we can assume that $\mathscr{B}$ is *primitive*, i.e. no b divides b' for distinct $b, b' \in \mathscr{B}$. Indeed, for a general set $\mathscr{B}$ one can find a primitive subset $\mathscr{B}_0 \subset \mathscr{B}$ such that $\mathcal{M}_{\mathscr{B}} = \mathcal{M}_{\mathscr{B}_0}$ and $\mathcal{F}_{\mathscr{B}} = \mathcal{F}_{\mathscr{B}_0}$ (cf. [**35**, Ch. 0]). Note that if we take $\mathscr{B} = \{p^2 : p \text{ is prime}\}$, then the set of $\mathscr{B}$-free numbers equals the set Q of squarefree numbers.

Sets of $\mathscr{B}$-free numbers make good candidates for rational sets. Unfortunately, not every set of $\mathscr{B}$-free numbers is a rational set, since the density $d(\mathcal{F}_{\mathscr{B}})$ of $\mathcal{F}_{\mathscr{B}}$ need not exist. An example of a set $\mathscr{B}$ for which the density of $\mathcal{M}_{\mathscr{B}}$ and $\mathcal{F}_{\mathscr{B}}$ does not exist was given by Besicovitch in [**15**]. This leads to the following definition.

Definition 2.14. (Cf. [**35**]) We say that $\mathscr{B}$ is *Besicovitch* if $d(\mathcal{M}_{\mathscr{B}})$ exists. (This is equivalent to the existence of $d(\mathcal{F}_{\mathscr{B}})$.)

Davenport and Erdős proved that the *logarithmic density*

$$\delta(\mathcal{M}_{\mathscr{B}}) := \lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \mathbb{1}_{\mathcal{M}_{\mathscr{B}}}(n)$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

exists for all $\mathscr{B} = \{b_1, b_2, \ldots\} \subset \mathbb{N} \setminus \{1\}$. This, of course, implies that the logarithmic density $\boldsymbol{\delta}(\mathcal{F}_{\mathscr{B}})$ exists. For $m \geqslant 1$, consider the sets $\mathscr{B}(m) = \{b_1, b_2, \ldots, b_m\}$ and let $\mathcal{M}_{\mathscr{B}(m)}$ and $\mathcal{F}_{\mathscr{B}(m)}$ denote the corresponding set of multiples of $\mathscr{B}(m)$ and set of $\mathscr{B}(m)$-free numbers, respectively.

THEOREM 2.15. (See [**20**, **21**]) *For each $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$, the logarithmic density $\boldsymbol{\delta}(\mathcal{M}_{\mathscr{B}})$ of $\mathcal{M}_{\mathscr{B}}$ exists. Moreover,*

$$\boldsymbol{\delta}(\mathcal{M}_{\mathscr{B}}) = \underline{d}(\mathcal{M}_{\mathscr{B}}) = \lim_{m\to\infty} d(\mathcal{M}_{\mathscr{B}(m)}).$$

In particular, if $\mathscr{B}$ is Besicovitch then $d(\mathcal{M}_{\mathscr{B}}) = \lim_{m\to\infty} d(\mathcal{M}_{\mathscr{B}(m)})$. Analogous results hold for $\mathcal{F}_{\mathscr{B}}$ instead of $\mathcal{M}_{\mathscr{B}}$.

From Theorem 2.15, we obtain two useful corollaries.

COROLLARY 2.16. *Let $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$. Then $\mathcal{M}_{\mathscr{B}}$ and $\mathcal{F}_{\mathscr{B}}$ are rational if and only if $\mathscr{B}$ is Besicovitch.*

Proof. Note that for any $m \geqslant 1$, the sequence $\mathbb{1}_{\mathcal{M}_{\mathscr{B}(m)}}$ is periodic. Hence if $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$ is Besicovitch, then by Theorem 2.15 the sequence $\mathbb{1}_{\mathcal{M}_{\mathscr{B}}}$ can be approximated in the d_B-pseudo-metric by $\mathbb{1}_{\mathcal{M}_{\mathscr{B}(m)}}$ as $m \to \infty$, which proves that $\mathcal{M}_{\mathscr{B}}$ is rational. An analogous argument applies to $\mathcal{F}_{\mathscr{B}}$.

On the other hand, if $\mathcal{M}_{\mathscr{B}}$ is rational then the density of $\mathcal{M}_{\mathscr{B}}$ exists and hence, by definition, the set $\mathscr{B}$ is Besicovitch. □

In the following, let $\mathscr{A}$ denote the set of *abundant numbers*, $\mathscr{P}$ the set of *perfect numbers* and $\mathscr{D}$ the set of *deficient numbers* (for definitions, see footnote ‡ on page 2334).

COROLLARY 2.17. *Let $A \subset \mathbb{N} \setminus \{1\}$. Suppose A satisfies the following two conditions:*
(1) *$d(A)$ exists, and*
(2) *$nA \subset A$ for all $n \in \mathbb{N}$.*
Then A is a rational set. In particular, the set of abundant numbers $\mathscr{A}$, the set of deficient numbers $\mathscr{D}$ and, for any $x \in [0, 1]$, the set $\Phi_x := \{n \in \mathbb{N} : \boldsymbol{\varphi}(n)/n < x\}$ are rational sets.

Proof. Set $\mathscr{B} := A$. It follows from property (2) that $\mathcal{M}_{\mathscr{B}} = A$. Also, $\mathscr{B}$ is Besicovitch because $d(\mathcal{M}_{\mathscr{B}}) = d(A)$ exists according to property (1). Hence, in view of Corollary 2.16, the set $A = \mathcal{M}_{\mathscr{B}}$ is rational.

We now turn our attention to the set of abundant numbers. First, note that $n\mathscr{A} \subset \mathscr{A}$ for all $n \in \mathbb{N}$. Also, the fact that $d(\mathscr{A})$ exists was proven by Davenport in [**19**]. Therefore $\mathscr{A}$ is rational. Moreover, since $\mathbb{N} = \mathscr{A} \,\dot\cup\, \mathscr{P} \,\dot\cup\, \mathscr{D}$ and $d(\mathscr{P}) = 0$ (cf. [**37**]), we conclude that $\mathscr{D}$ is also rational.

Finally, for any $x \in [0, 1]$, the set $\Phi_x := \{n \in \mathbb{N} : \boldsymbol{\varphi}(n)/n < x\}$ satisfies $n\Phi_x \subset \Phi_x$ and it was first shown in [**57**] that $d(\Phi_x)$ exists. Hence Φ_x is rational. □

As mentioned above, a shift of the set of squarefree numbers, $Q - r$, is an averaging set of polynomial multiple recurrence if and only if $r \in Q$. Our next goal is to show that a similar result holds for other sets of $\mathscr{B}$-free numbers. Note that if $r \notin \mathcal{F}_{\mathscr{B}}$, i.e. $r \in \mathcal{M}_{\mathscr{B}}$, then $\mathcal{F}_{\mathscr{B}} - r$ is not a set of recurrence. Indeed, if it were a set of recurrence then, by

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

considering the cyclic rotation on r points, for some $u \geqslant 1$ we would have $ur \in \mathcal{F}_{\mathscr{B}} - r$ and therefore $(u+1)r \in \mathcal{F}_{\mathscr{B}}$, which is a contradiction. Hence, $r \in \mathcal{F}_{\mathscr{B}}$ is a necessary condition for $\mathcal{F}_{\mathscr{B}} - r$ to be good for recurrence. As for the other direction, we have the following result.

THEOREM 2.18. *Suppose $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$ is Besicovitch. Then 'almost every' self-shift of $\mathcal{F}_{\mathscr{B}}$ is an averaging set of polynomial multiple recurrence. More precisely, there exists a set $D \subset \mathcal{F}_{\mathscr{B}}$ with $d(\mathcal{F}_{\mathscr{B}} \setminus D) = 0$ such that for all $r \in \mathbb{N}$ the following are equivalent:*

- *$r \in D$;*
- *$\mathcal{F}_{\mathscr{B}} - r$ is divisible;*
- *$\mathcal{F}_{\mathscr{B}} - r$ is an averaging set of polynomial multiple recurrence.*

Note that Theorem 1.5 is now an immediate consequence of Theorem 2.18. We give a proof of Theorem 2.18 at the end of this subsection. Let us remark that in most cases one can actually take $D = \mathcal{F}_{\mathscr{B}}$. To distinguish between sets of $\mathscr{B}$-free numbers for which $D = \mathcal{F}_{\mathscr{B}}$ and for which $D \subsetneq \mathcal{F}_{\mathscr{B}}$, we introduce the following notions.

Definition 2.19. (Cf. [**35**]) Let $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$. We call $\mathscr{B}$ *Behrend* if $\boldsymbol{\delta}(\mathcal{M}_{\mathscr{B}}) = 1$ (this is equivalent to the existence of the density of $\mathcal{M}_{\mathscr{B}}$ with $d(\mathcal{M}_{\mathscr{B}}) = 1$). We call $\mathscr{B}$ *taut* if for every $b \in \mathscr{B}$, one has $\boldsymbol{\delta}(\mathcal{M}_{\mathscr{B}}) > \boldsymbol{\delta}(\mathcal{M}_{\mathscr{B}\setminus\{b\}})$.

If $\mathscr{B}$ is Behrend then $D = \emptyset$, because in this case $\mathcal{F}_{\mathscr{B}}$ (and each of its translations) has zero density. Behrend sets are not taut (see Lemma 2.20 below) and it will be clear from the proof of Theorem 2.18 that in the statement of Theorem 2.18 one can take $D = \mathcal{F}_{\mathscr{B}}$ if and only if $\mathscr{B}$ is taut (see Theorem 2.26 below).

The remainder of this subsection is dedicated to proving Theorem 2.18. We start with a series of lemmas.

LEMMA 2.20. [**35**, Corollary 0.14] *$\mathscr{A} \cup \mathscr{B}$ is Behrend if and only if at least one of $\mathscr{A}$ and $\mathscr{B}$ is Behrend. In particular, Behrend sets are not taut.*

LEMMA 2.21. [**35**, Corollary 0.19] *$\mathscr{B}$ is taut if and only if it is primitive and does not contain a set of the form $c\mathscr{A}$, where $c \in \mathbb{N}$ and $\mathscr{A} \subset \mathbb{N} \setminus \{1\}$ is Behrend.*

LEMMA 2.22. (Cf. the proof of [**4**, Lemma 6.5]) *Let $\mathscr{C} \subset \mathbb{N}$. For any $u \in \mathbb{N}$ and $a \in \mathbb{N} \cup \{0\}$ the logarithmic densities of $\mathcal{M}_{\mathscr{C}} \cap (u\mathbb{N} + a)$ and $\mathcal{F}_{\mathscr{C}} \cap (u\mathbb{N} + a)$ exist and satisfy*

$$\boldsymbol{\delta}(\mathcal{M}_{\mathscr{C}} \cap (u\mathbb{N} + a)) = \underline{d}(\mathcal{M}_{\mathscr{C}} \cap (u\mathbb{N} + a)) = \lim_{m\to\infty} d(\mathcal{M}_{\mathscr{C}(m)} \cap (u\mathbb{N} + a)),$$
$$\boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}} \cap (u\mathbb{N} + a)) = \overline{d}(\mathcal{F}_{\mathscr{C}} \cap (u\mathbb{N} + a)) = \lim_{m\to\infty} d(\mathcal{F}_{\mathscr{C}(m)} \cap (u\mathbb{N} + a)).$$

Proof. The assertion concerning $\mathcal{M}_{\mathscr{C}} \cap (u\mathbb{N} + a)$ was covered in the proof of [**4**, Lemma 6.5]. The remaining part follows immediately, as $\mathcal{F}_{\mathscr{C}} = \mathbb{N} \setminus \mathcal{M}_{\mathscr{C}}$. □

LEMMA 2.23. *Let $\mathscr{C} \subset \mathbb{N} \setminus \{1\}$ and let $a \in \mathbb{N} \cup \{0\}$. If $u \in \mathbb{N}$ is coprime to each element of $\mathscr{C}$ then*

$$\boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}} \cap (u\mathbb{N} + a)) = \frac{1}{u} \cdot \boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}}).$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Proof. Suppose $\mathscr{C} = \{b_1, b_2, \ldots\}$ and let $\mathscr{C}(m) := \{b_1, \ldots, b_m\}$. The assertion of the lemma is clearly equivalent to $\boldsymbol{\delta}(\mathcal{M}_{\mathscr{C}} \cap (u\mathbb{N} + a)) = (1/u) \cdot \boldsymbol{\delta}(\mathcal{M}_{\mathscr{C}})$. Since u is coprime to each element of $\mathscr{C}$, it follows by the Chinese remainder theorem that for any $m \geqslant 1$ and $r \in \mathbb{N} \cup \{0\}$ there exists $r' \in \mathbb{N} \cup \{0\}$ such that

$$(\operatorname{lcm}(b_1, \ldots, b_m)\mathbb{N} + r) \cap (u\mathbb{N} + a) = u \cdot \operatorname{lcm}(b_1, \ldots, b_m)\mathbb{N} + r'.$$

In particular,

$$d((\operatorname{lcm}(b_1, \ldots, b_m)\mathbb{N} + r) \cap (u\mathbb{N} + a)) = \frac{1}{u \cdot \operatorname{lcm}(b_1, \ldots, b_m)}.$$

It follows that

$$d(\mathcal{M}_{\mathscr{C}(m)} \cap (u\mathbb{N} + a)) = \frac{1}{u} \cdot d(\mathcal{M}_{\mathscr{C}(m)}) \tag{2.21}$$

since $\mathcal{M}_{\mathscr{C}(m)}$ is periodic with period $\operatorname{lcm}(b_1, \ldots, b_m)$. Using Lemma 2.22, (2.21) and Theorem 2.15, we obtain

$$\begin{aligned}\boldsymbol{\delta}(\mathcal{M}_{\mathscr{C}} \cap (u\mathbb{N} + a)) &= \lim_{m\to\infty} d(\mathcal{M}_{\mathscr{C}(m)} \cap (u\mathbb{N} + a)) \\ &= \lim_{m\to\infty} \frac{1}{u} \cdot d(\mathcal{M}_{\mathscr{C}(m)}) = \frac{1}{u} \cdot \boldsymbol{\delta}(\mathcal{M}_{\mathscr{C}}),\end{aligned}$$

which completes the proof. □

LEMMA 2.24. *Suppose $\mathscr{C} \subset \mathbb{N} \setminus \{1\}$ is taut. If $a \in \mathcal{F}_{\mathscr{C}}$ then, for every $u \in \mathbb{N}$, one has $\boldsymbol{\delta}((\mathcal{F}_{\mathscr{C}} - a)/u) > 0$.*

Proof. Define

$$\mathscr{C}'(a) := \left\{ \frac{b}{\gcd(b, a)} : b \in \mathscr{C} \right\}.$$

Notice that

$$\gcd(a, c) = 1 \quad \text{for each } c \in \mathscr{C}'(a). \tag{2.22}$$

Moreover, $\mathcal{M}_{\mathscr{C}'(a)} \supset \mathcal{M}_{\mathscr{C}}$, whence

$$\mathcal{F}_{\mathscr{C}'(a)} \subset \mathcal{F}_{\mathscr{C}}. \tag{2.23}$$

Since $\gcd(b, a)$ takes only finitely many values as $b \in \mathscr{C}$ varies,

$$\mathscr{C}'(a) = \bigcup_{d|a} \left\{ \frac{b}{d} : b \in \mathscr{C}, \gcd(b, a) = d \right\}.$$

Suppose that $1 \in \mathscr{C}'(a)$. Then for some $b \in \mathscr{C}$, $\gcd(b, a) = b$, whence $a \in \mathcal{M}_{\mathscr{C}}$, a contradiction. It follows that

$$1 \notin \mathscr{C}'(a). \tag{2.24}$$

Suppose that $\mathscr{C}'(a)$ is Behrend. Then, by Lemma 2.20, for some $d_0 \mid a$, the set $\mathscr{A} := \{b/d_0 : b \in \mathscr{C}, \gcd(b, a) = d_0\}$ is Behrend and we have $d_0\mathscr{A} \subset \mathscr{C}$. However, this and (2.24), in view of Lemma 2.21, contradict the tautness of $\mathscr{C}$. Therefore, $\mathscr{C}'(a)$ cannot be Behrend, i.e. $c := \boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}'(a)}) > 0$. We will use this constant to prove that $\boldsymbol{\delta}((\mathcal{F}_{\mathscr{C}} - a)/u) > c$ for all $u \in \mathbb{N}$.

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

By Lemma 2.22 and (2.23),

$$\begin{aligned}\boldsymbol{\delta}((\mathcal{F}_{\mathscr{C}} - a)/u) &= u \cdot \boldsymbol{\delta}((\mathcal{F}_{\mathscr{C}} - a) \cap u\mathbb{N}) \geqslant u \cdot \boldsymbol{\delta}((\mathcal{F}_{\mathscr{C}'(a)} - a) \cap u\mathbb{N}) \\ &= u \cdot \boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}'(a)} \cap (u\mathbb{N} + a)).\end{aligned}$$

Hence, it suffices to show that

$$\boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}'(a)} \cap (u\mathbb{N} + a)) \geqslant \frac{1}{u} \cdot \boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}'(a)}).$$

In order to verify this last claim, let us divide $\mathscr{C}'(a)$ into two pieces:

$$\begin{aligned}\mathscr{C}'_1(a, u) &:= \{c \in \mathscr{C}'(a) : \gcd(u, c) > 1\}, \\ \mathscr{C}'_2(a, u) &:= \{c \in \mathscr{C}'(a) : \gcd(u, c) = 1\}.\end{aligned}$$

We claim that

$$\mathcal{F}_{\mathscr{C}'_1(a,u)} \cap (u\mathbb{N} + a) = u\mathbb{N} + a. \tag{2.25}$$

Indeed, take any $c \in \mathscr{C}'_1(a, u)$. Then $\gcd(u, c) > 1$ and since (2.22) holds, $\gcd(u, c)$ does not divide a. Hence $c\mathbb{N} \cap (u\mathbb{N} + a) = \emptyset$. It follows that $\mathcal{M}_{\mathscr{C}'_1(a,u)} \cap (u\mathbb{N} + a) = \emptyset$ and (2.25) follows.

Therefore, using (2.25) and additionally Lemma 2.23, we obtain

$$\boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}'(a)} \cap (u\mathbb{N} + a)) = \boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}'_2(a)} \cap (u\mathbb{N} + a)) = \frac{1}{u} \cdot \boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}'_2(a)}) \geqslant \frac{1}{u} \cdot \boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}'(a)})$$

and the result follows. □

Before we present the proof of Theorem 2.18, one more theorem needs to be quoted.

THEOREM 2.25. [**4**, Theorem 4.5 and the proof of Lemma 4.11] *Let $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$. Then there exists a taut set $\mathscr{C} \subset \mathbb{N} \setminus \{1\}$ such that $\mathcal{F}_{\mathscr{C}} \subset \mathcal{F}_{\mathscr{B}}$ and $\boldsymbol{\delta}(\mathcal{F}_{\mathscr{C}}) = \boldsymbol{\delta}(\mathcal{F}_{\mathscr{B}})$. Moreover, if $\mathscr{B}$ is Besicovitch, then $\mathscr{C}$ is Besicovitch.*

Proof of Theorem 2.18. Let $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$ be Besicovitch. If $\mathscr{B}$ is Behrend then $\mathcal{F}_{\mathscr{B}}$ has zero density and so no shift of $\mathcal{F}_{\mathscr{B}}$ is divisible or good for averaging polynomial recurrence. In this case we can put $D = \emptyset$ and we are done. Thus, let us assume that $\mathscr{B}$ is not Behrend. In view of Theorem 1.4, it suffices to find a set $D \subset \mathcal{F}_{\mathscr{B}}$ with $d(\mathcal{F}_{\mathscr{B}} \setminus D) = 0$ and such that $\mathcal{F}_{\mathscr{B}} - r$ is divisible if and only if $r \in D$. Pick $\mathscr{C} \subset \mathbb{N} \setminus \{1\}$ taut with $\mathcal{F}_{\mathscr{C}} \subset \mathcal{F}_{\mathscr{B}}$ and $d(\mathcal{F}_{\mathscr{C}}) = d(\mathcal{F}_{\mathscr{B}})$; the existence of $\mathscr{C}$ is guaranteed by Theorem 2.25.

We make the claim that one can choose $D := \mathcal{F}_{\mathscr{C}}$. In particular, if $\mathscr{B}$ is taut then one can choose $D = \mathcal{F}_{\mathscr{B}}$.

To verify this claim, we invoke Lemma 2.24, which tells us that $\boldsymbol{\delta}((\mathcal{F}_{\mathscr{C}} - r)/u) > 0$ if and only if $r \in \mathcal{F}_{\mathscr{C}}$. Since $\mathscr{C}$ is Besicovitch, we can replace logarithmic density with density and conclude that $\mathcal{F}_{\mathscr{C}} - r$ is divisible if and only if $r \in \mathcal{F}_{\mathscr{C}}$. Finally, to finish the proof, we observe that $d(\mathcal{F}_{\mathscr{B}} \setminus \mathcal{F}_{\mathscr{C}}) = 0$ and therefore $\mathcal{F}_{\mathscr{B}} - r$ is divisible if and only if $r \in \mathcal{F}_{\mathscr{C}}$. □

THEOREM 2.26. (Corollary of the proof of Theorem 2.18) *In the statement of Theorem 2.18 one has $D = \mathcal{F}_{\mathscr{B}}$ if and only if $\mathscr{B}$ is taut.*

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

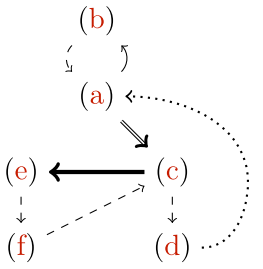


FIGURE 1. Logical connections between statements (a)–(f) in Remark 2.27. Dashed arrows: trivial implications which hold for any $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$. Dotted arrow: this implication was explained in the paragraph before Theorem 2.18. Short plain arrow: implication follows from [**4**]. Thick arrow: this implication follows from Theorem 1.4. Double arrow: implication proved in Lemma 2.24.

Remark 2.27. Let $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$ be Besicovitch and taut (hence $d(\mathcal{F}_{\mathscr{B}}) > 0$). Here is a summary of equivalent conditions that we obtained in this section.

(a) $a \in \mathcal{F}_{\mathscr{B}}$.
(b) $d(\mathcal{M}_{\mathscr{B}} \cup a\mathbb{N}) > d(\mathcal{M}_{\mathscr{B}})$.
(c) $\mathcal{F}_{\mathscr{B}} - a$ is divisible.
(d) $(\mathcal{F}_{\mathscr{B}} - a) \cap u\mathbb{N} \neq \emptyset$ for all $u \in \mathbb{N}$.
(e) $\mathcal{F}_{\mathscr{B}} - a$ is an averaging set of polynomial multiple recurrence.
(f) $\mathcal{F}_{\mathscr{B}} - a$ is an averaging set of polynomial single recurrence.

Figure 1 describes the logical connections between the above statements.

3. *Rational dynamical systems*

The purpose of this section is to give a proof of (slightly more general versions of) Theorems 1.7 and 1.8. In §3.1 we define rational and W-rational subshifts and we give a variety of examples. In §3.2 we extend the notion of rational subshifts to the notion of rational subshifts along increasing subsequences. Finally, in §§3.3 and 3.4, we formulate and prove extensions of Theorems 1.7 and 1.8.

3.1. *Definition and examples of rational subshifts.*

In this section we define and give examples of symbolic dynamical systems determined by RAP sequences. We will refer to such systems as *rational subshifts*.

Consider the product space $\mathcal{A}^{\mathbb{Z}}$, where $\mathcal{A}$ is a finite set (*alphabet*). We endow $\mathcal{A}$ with the discrete metric ρ and $\mathcal{A}^{\mathbb{Z}}$ with the product topology induced by $(\mathcal{A}, \rho)$; in particular, $\mathcal{A}^{\mathbb{Z}}$ is compact and metrizable. Let $S\colon \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ be the left shift, i.e. $S((x(n))_{n\in\mathbb{Z}}) = (y(n))_{n\in\mathbb{Z}}$, where $y(n) = x(n+1)$ for each $n \in \mathbb{Z}$.

Recall that for any closed and S-invariant subset $X \subset \mathcal{A}^{\mathbb{Z}}$, the system (X, S) is called a *subshift* of $(\mathcal{A}^{\mathbb{Z}}, S)$. Recall that for $x \in \mathcal{A}^{\mathbb{Z}}$ (or $x \in \mathcal{A}^{\mathbb{N}}$) and $n < m$, $x[n, m] = (x(n), x(n+1), \ldots, x(m))$ is said to be a word appearing in x.

Given $\eta \in \mathcal{A}^{\mathbb{N}}$, the set

$$X_\eta := \{x \in \mathcal{A}^{\mathbb{Z}} : (\forall n < m)(\exists k \in \mathbb{N})\, x[n, m] = \eta[k, k+m-n-1]\}$$

is closed and S-invariant. It is the *subshift determined by* η.

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Recall that in Definition 1.1 we introduced the Besicovitch pseudo-metric,

$$d_B(x, y) := \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \rho(x(n), y(n)) \tag{3.1}$$

and defined *rationally almost periodic* sequences, which are sequences that can be approximated in the d_B-pseudo-metric by periodic sequences.

Definition 3.1. A subshift (X, S) of $(\mathcal{A}^{\mathbb{Z}}, S)$ is called *rational* if there exists an RAP sequence $\eta \in \mathcal{A}^{\mathbb{N}}$ such that $X = X_\eta$.

We now present some examples of rational subshifts.

The squarefree subshift. Consider the set Q of squarefree numbers and let $X_Q := X_{\mathbb{1}_Q} \subset \{0, 1\}^{\mathbb{Z}}$. The resulting topological dynamical system (X_Q, S) is called the *squarefree subshift* and has been studied in [**17**, **53**, **56**]. Naturally, many combinatorial properties of Q are encoded in the dynamics of (X_Q, S), which further motivates the study of this system. This line of investigation is related to Sarnak's conjecture; see [**27**, **56**] and §5.2.

$\mathscr{B}$-free subshifts. For $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$ let $\mathcal{F}_{\mathscr{B}}$ denote the set of $\mathscr{B}$-free numbers and let $X_{\mathcal{F}_{\mathscr{B}}} := X_{\mathbb{1}_{\mathcal{F}_{\mathscr{B}}}} \subset \{0, 1\}^{\mathbb{Z}}$. The system $(X_{\mathcal{F}_{\mathscr{B}}}, S)$ is called the *$\mathscr{B}$-free subshift*. Such subshifts have recently been studied in [**4**, **28**, **43**]. If the set $\mathscr{B}$ is Besicovitch (see Definition 2.14) then it follows from Corollary 2.16 that $(X_{\mathcal{F}_{\mathscr{B}}}, S)$ is a rational subshift. Note that the squarefree subshift (X_Q, S) is an example of a $\mathscr{B}$-free subshift.

Toeplitz systems. Following [**40**], a sequence $\eta \in \mathcal{A}^{\mathbb{N}}$ is called *Toeplitz*, if for each $n \geqslant 0$ there is $p \geqslant 1$ such that

$$\eta(n) = \eta(n + sp) \quad \text{for all } s \in \mathbb{N}. \tag{3.2}$$

In this case the subshift (X_η, S) is called a *Toeplitz system*. In [**24**], Downarowicz characterized Toeplitz dynamical systems as being exactly all symbolic, minimal and almost one-to-one extensions of odometers.

If additionally

$$\limsup_{p\to\infty} \overline{d}(\{n \in \mathbb{N} : \eta(n) = \eta(n + sp) \text{ for all } s \in \mathbb{N}\}) = 1 \tag{3.3}$$

then the Toeplitz sequence η is called *regular*. It follows from (3.3) that any regular Toeplitz sequence is RAP (in fact, it is straightforward to check that a Toeplitz sequence is regular if and only if it is RAP). Therefore Toeplitz systems coming from regular Toeplitz sequences are rational subshifts.

Weyl almost periodic sequences and Weyl rational subshifts. We recall the definition of the *Weyl pseudo-metric* d_W (see Definition 2.11),

$$d_W(x, y) = \limsup_{N\to\infty} \sup_{\ell\geqslant 1} \frac{1}{N} |\{\ell \leqslant n \leqslant \ell + N : x(n) \neq y(n)\}|.$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Any d_W-limit of periodic sequences is called a *Weyl rationally almost periodic sequence* (WRAP). The subshift (X_η, S) determined by a WRAP sequence $\eta \in \mathcal{A}^{\mathbb{N}}$ is called *W-rational*†. Each WRAP sequence is RAP and hence any W-rational subshift is a rational subshift. Note that the reverse implication does not hold. For example, the indicator function $\mathbb{1}_Q$ of the squarefree numbers Q is a sequence that is RAP but not WRAP. It should also be mentioned that any regular Toeplitz sequence is WRAP, which can be shown easily using the definition of regular Toeplitz sequences.

Sequences generated by synchronized automata. Paperfolding sequences, which were introduced in §1 (see footnote † on page 2335), provide examples of rational sequences generated by so-called synchronized automata.

Let $k \in \mathbb{N}$, let $\mathcal{B} := \{0, 1, \ldots, k-1\}$, let $\mathcal{A}$ be a finite alphabet, let $Q := \{q_0, \ldots, q_r\}$ be a finite set, let $\tau : Q \to \mathcal{A}$ and let $\delta\colon Q \times \mathcal{B} \to Q$. The quintuple $M = (Q, \mathcal{B}, \delta, q_0, \tau)$ is called a *complete deterministic finite automaton* with *set of states* Q, *input alphabet* $\mathcal{B}$, *output alphabet* $\mathcal{A}$, *transition function* δ, *initial state* q_0 and *output mapping* τ. Let $\mathcal{B}^*$ denote the collection of all finite words in letters from $\mathcal{B}$. There is a natural way of extending $\delta\colon Q \times \mathcal{B} \to Q$ to $\delta\colon Q \times \mathcal{B}^* \to Q$: for the empty word $\epsilon \in \mathcal{B}^*$ we define $\delta(q, \epsilon) := q$, $q \in Q$, and for a non-empty word $w = w_1 \ldots w_n \in \mathcal{B}^*$ we define recursively $\delta(q, w_1 \ldots w_n) := \delta(\delta(q, w_1 \ldots w_{n-1}), w_n)$, $q \in Q$. This way, we can associate to each word $w \in \mathcal{B}^*$ an element $a \in \mathcal{A}$ via $a = \tau(\delta(q_0, w))$. For more details on deterministic finite automata, see [**2**, §4.1].

Given $n \in \mathbb{N}$, we consider its expansion in base k, i.e.

$$n = \sum_{j \geqslant 0} \varepsilon_j k^j \quad \text{where } \varepsilon_j \in \mathcal{B},\ j \geqslant 0.$$

In this representation $\varepsilon_j = 0$ for all but finitely many $j \geqslant 0$. Let j_n be the largest index such that $\varepsilon_{j_n} \neq 0$. We then set $[n]_k := (\varepsilon_{j_n}, \varepsilon_{j_n-1}, \ldots, \varepsilon_0)$; note that $[n]_k \in \mathcal{B}^*$ for all $n \in \mathbb{N}$.

Definition 3.2. Following [**55**], we say that a sequence $a \in \mathcal{A}^{\mathbb{N}}$ is *automatic* if there exists a complete deterministic finite automaton M as above such that $a(n) = \tau(\delta(q_0, [n]_k))$ for all $n \in \mathbb{N}$.

Definition 3.3. Following [**14**, Part 4], an automaton $M = (Q, \mathcal{B}, \delta, q_0, \tau)$ is called *synchronized* if there exists a word $w \in \mathcal{B}^*$ such that $\delta(q, w) = \delta(q_0, w)$ for all $q \in Q$. In this case, the word w is called a *synchronizing word*.

While not every automatic sequence is RAP‡, any automatic sequence coming from a synchronized automaton is not only RAP but also WRAP. This result, which we state as a proposition below, has been shown implicitly in [**23**]. For the sake of completeness, we give a proof of it in §5.

† W-rational subshifts are a special kind of *Weyl almost periodic systems*; see [**26**, **39**].
‡ The Thue–Morse sequence (which was independently discovered by Thue [**59**, **60**] and Morse [**50**, **51**]) is known to be automatic (see [**2**]), but it is not RAP as it is a generic point for a measure such that the corresponding dynamical system has no purely discrete spectrum [**41**]; see Theorem 1.7 above.

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

PROPOSITION 3.4. (See [**23**]) *Each automatic sequence generated by a synchronized automaton is WRAP.*

As RAP sequences can be approximated by periodic sequences, one may be tempted to believe that the dynamics of rational subshifts is similar to the dynamics of certain low-complexity systems such as translations on compact groups. However, rational subshifts exhibit a wide variety of dynamical properties, such as the following.

- Rational subshifts can have positive topological entropy (for the definition of topological entropy, see, for instance, [**54**, §6.3]); for example, rational $\mathscr{B}$-free subshifts can have positive topological entropy (see [**4**, **28**, **53**]). Moreover, they can have many invariant measures [**43**].
- Rational subshifts can be proximal†; in fact, the squarefree subshift is an example of a proximal rational subshift [**56**].
- Rational subshifts can be topologically mixing (see Remark 3.20 for a proof).
- If (X, S) is a rational subshift of positive entropy, then there is $y \in X$ which is not RAP. As a matter of fact, no generic point for a measure of positive entropy is RAP (this follows from Theorem 3.12 below).

On the other hand, W-rational subshifts have much more regular properties than general rational subshifts. This is illustrated by the following two propositions.

PROPOSITION 3.5. [**26**, **40**] *If x is WRAP then (X_x, S) is uniquely ergodic and has zero topological entropy‡. In particular, any non-trivial W-rational subshift is not proximal.*

PROPOSITION 3.6. (Cf. [**39**, Lemma 4]) *Let $x \in \mathcal{A}^{\mathbb{N}}$ be WRAP and suppose $z \in X_x$. Let $z|_{\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ denote the restriction of $z \in \mathcal{A}^{\mathbb{Z}}$ to $\mathbb{N}$. Then $z|_{\mathbb{N}}$ is also WRAP.*

Proof. Let $z \in X_x$ and let $\varepsilon > 0$. Pick any periodic sequence $y \in \mathcal{A}^{\mathbb{N}}$ with $d_W(x, y) \leqslant \varepsilon/2$. Let M denote the period of y. Let ℓ_K be such that $x(n + \ell_K) = z(n)$ for $n \leqslant K$. We can assume without loss of generality that there exists $0 \leqslant i_0 < M$ with $\ell_K \equiv i_0 \bmod M$ for all $K \in \mathbb{N}$.

Let N_0 be such that for all $N \geqslant N_0$,

$$\sup_{\ell \geqslant 0} \frac{1}{N} \sum_{n=1}^{N} \rho(x(n + \ell), y(n + \ell)) \leqslant \varepsilon. \tag{3.4}$$

Fix $N \geqslant N_0$ and $\ell \in \mathbb{N}$, and let $K \geqslant N + \ell$. Then, by the choice of ℓ_K,

$$z(n + \ell) = x(n + \ell + \ell_K) \quad \text{for all } n \leqslant N. \tag{3.5}$$

Moreover, since y is M-periodic and since $\ell_K \equiv i_0 \bmod M$,

$$S^{i_0} y(n + \ell) = y(n + \ell + \ell_K) \quad \text{for each } n \in \mathbb{N}. \tag{3.6}$$

† A dynamical system (Y, T) is *proximal* if any pair $y, z \in Y$ is proximal, i.e. there exists a sequence (n_k) such that $d(T^{n_k} y, T^{n_k} z) \to 0$. In particular, in any such system y and Ty are proximal and it follows that proximal systems have exactly one fixed point to which all other points are proximal.
‡ It is shown in [**26**] that for subshifts on a finite alphabet the entropy function is d_W-continuous.

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Using (3.5), (3.6) and (3.4), we conclude that

$$\frac{1}{N}\sum_{n=1}^{N}\rho(z(n+\ell),\, S^{i_0}y(n+\ell)) = \frac{1}{N}\sum_{n=1}^{N}\rho(x(n+\ell+\ell_K),\, S^{i_0}y(n+\ell))$$
$$= \frac{1}{N}\sum_{n=1}^{N}\rho(x(n+\ell+\ell_K),\, y(n+\ell+\ell_K)) \leqslant \varepsilon$$

and the result follows. □

3.2. *Rationality along subsequences.* In this short subsection we introduce and discuss a useful generalization of RAP sequences.

Definition 3.7. Let $(N_k)_{k\geqslant 1}$ be an increasing sequence of natural numbers and assume that $\mathcal{A}$ is a finite set endowed with the discrete metric ρ.

- For $x,\, y \in \mathcal{A}^{\mathbb{N}}$, we define (cf. (1.3), (2.2) and (3.1))

$$d_B^{(N_k)}(x,\, y) := \limsup_{k\to\infty}\frac{1}{N_k}\sum_{n=1}^{N_k}\rho(x(n),\, y(n)).$$

  We say that $x \in \mathcal{A}^{\mathbb{N}}$ is *rationally almost periodic along* $(N_k)_{k\geqslant 1}$ (RAP along $(N_k)_{k\geqslant 1}$) if x is a $d_B^{(N_k)}$-limit of periodic sequences. A subset $R \subset \mathbb{N}$ is called *rational along* $(N_k)_{k\geqslant 1}$, if $\mathbb{1}_R$ is RAP along $(N_k)_{k\geqslant 1}$.
- Let $(X,\, S)$ be a subshift of $(\mathcal{A}^{\mathbb{Z}},\, S)$. We call the topological dynamical system $(X,\, S)$ *rational along* $(N_k)_{k\geqslant 1}$ if $X = X_\eta$ for some $\eta \in \mathcal{A}^{\mathbb{N}}$ that is RAP along $(N_k)_{k\geqslant 1}$. If $N_k = k$ for all k then, clearly, $(X,\, S)$ is a rational subshift (see Definition 3.1).

Remark 3.8. Let $R \subset \mathbb{N}$ and suppose that R is rational along $(N_k)_{k\geqslant 1}$. For $u \geqslant 1$, let $R/u := \{n \in \mathbb{N} : nu \in R\}$. Then R/u is rational along $(N_k/u)_{k\geqslant 1}$ for any $u \geqslant 1$.

Clearly, any sequence that is RAP is also RAP along $(N_k)_{k\geqslant 1}$. The following example shows that, in general, the converse does not hold.

Example 3.9. Given an increasing sequence $(b_k) \subset \mathbb{N}$, we define

$$x = x^{(b_k)} := \underbrace{010101\ldots01}_{b_1}\,\underbrace{101010\ldots10}_{b_2}\,\underbrace{010101\ldots01}_{b_3}\ldots$$

Note that if $(b_k)_{k\in\mathbb{N}}$ is increasing sufficiently fast then x is RAP along $(N_{2k})_{k\geqslant 1}$, where $N_k := b_1 + \cdots + b_k$, $k \geqslant 1$. We will now show that for no choice of increasing (b_k) is the sequence x RAP. Suppose that there is $(b_k)_{k\in\mathbb{N}}$ that yields an RAP sequence x. Then the sequence

$$x' := \underbrace{0\ldots0}_{b_1/2}\,\underbrace{1\ldots1}_{b_2/2}\,\underbrace{0\ldots0}_{b_3/2}\ldots$$

must also be RAP and therefore the density $d(A)$ of $A := \{n \in \mathbb{N} : x'(n) = 1\}$ exists. Notice that

$$|\{1 \leqslant n \leqslant N_{2k}/2 : x'(n) = 1\}| > 1/2 \cdot N_{2k}/2,$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

so, in particular, $d(A) \geqslant 1/2 > 0$. Since (b_k) is increasing, $d(A) = 1/2$. Let $0 < \varepsilon < 1/4$ and let y be a periodic sequence with $d_B(x', y) < \varepsilon$. Then $|\{1 \leqslant n \leqslant R : y(n) = 1\}| = (\frac{1}{4} \pm \varepsilon)R$, where R is a period of y. We will now estimate $d_B(x', y)$ from below. Fix $1 \leqslant i \leqslant R$. If $y(i) = 0$ then

$$\frac{1}{N}|\{1 \leqslant n \leqslant N : x'(Rn + i) \neq y(i)\}| = \frac{1}{N}|\{1 \leqslant n \leqslant N : x'(Rn + i) = 1\}| \to \frac{1}{R} \cdot \frac{1}{2},$$

and similarly, if $y(i) = 1$ then

$$\frac{1}{N}|\{1 \leqslant n \leqslant N : x'(Rn + i) \neq y(i)\}| = \frac{1}{N}|\{1 \leqslant n \leqslant N : x'(Rn + i) = 0\}| \to \frac{1}{R} \cdot \frac{1}{2}.$$

It follows that

$$\begin{aligned}
\varepsilon \geqslant d_B(x', y) &= \limsup_{N\to\infty} \frac{1}{N}|\{1 \leqslant n \leqslant N : x'(n) \neq y(n)\}| \\
&= \limsup_{N\to\infty} \frac{1}{N}\Bigg|\bigcup_{0\leqslant i<R} \{1 \leqslant n \leqslant N : x'(Rn + i) \neq y(i)\}\Bigg| \\
&\geqslant \liminf_{N\to\infty} \frac{1}{N}\Bigg|\bigcup_{0\leqslant i<R} \{1 \leqslant n \leqslant N : x'(Rn + i) \neq y(i)\}\Bigg| \\
&\geqslant \sum_{0\leqslant i<R} \liminf_{N\to\infty} \frac{1}{N}|\{1 \leqslant n \leqslant N : x'(Rn + i) \neq y(i)\}| \\
&\geqslant \left(\frac{1}{2} - \varepsilon\right)R \cdot \frac{1/2}{R} + \left(\frac{1}{2} - \varepsilon\right)R \cdot \frac{1/2}{R} = \frac{1}{2} - \varepsilon.
\end{aligned}$$

This yields a contradiction with our choice of ε.

Remark 3.10. Notice that for x as in Example 3.9 the density of the set $\{n \in \mathbb{N} : x(n) = 1\}$ equals $\frac{1}{2}$. This shows that there are sets that have density, are not rational but are rational along some subsequence. In fact, one can show that x is a generic point for the measure $\frac{1}{2}(\delta_{101010\ldots} + \delta_{010101\ldots})$.

Example 3.11. It was shown in §2.4 (see Corollary 2.16) that for any set $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$ the set of $\mathscr{B}$-free numbers $\mathcal{F}_{\mathscr{B}}$ is rational if and only if $\mathscr{B}$ is Besicovitch. In particular, if $\mathscr{B}$ is not Besicovitch then $\mathcal{F}_{\mathscr{B}}$ is not rational. However, if $(N_k)_{k\geqslant 1}$ is an increasing sequence such that $\underline{d}(\mathcal{F}_{\mathscr{B}}) = \lim_{k\to\infty} |\mathcal{F}_{\mathscr{B}} \cap \{1, \ldots, \mathbb{N}_k\}|/N_k$, then it follows from Theorem 2.15 that $\mathcal{F}_{\mathscr{B}}$ is rational along $(N_k)_{k\geqslant 1}$.

3.3. *Generalizing Theorem 1.7.* Let $\mathcal{A}$ be a finite alphabet and let (X, S) be a subshift of the full shift $(\mathcal{A}^{\mathbb{Z}}, S)$. Denote by $\mathcal{P}(X, S)$ the set of all S-invariant Borel probability measures on X and by $\mathcal{P}^e(X, S)$ the subset of $\mathcal{P}(X, S)$ of ergodic measures. By the Krylov–Bogolyubov theorem [**42**], $\mathcal{P}(X, S)$ is non-empty.

Given an increasing sequence $(N_k)_{k\geqslant 1}$ of natural numbers, we say that $x \in \mathcal{A}^{\mathbb{N}}$ is *quasi-generic for* $\mu \in \mathcal{P}(X, S)$ *along* $(N_k)_{k\geqslant 1}$ if

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} f(S^n\tilde{x}) = \int_X f \, d\mu$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

for all continuous functions $f \in C(\mathcal{A}^{\mathbb{Z}})$, where $\tilde{x} \in \mathcal{A}^{\mathbb{Z}}$ is any two-sided sequence that *extends* x, i.e. $\tilde{x}(n) = x(n)$ for all $n \in \mathbb{N}$. If $N_k = k$ for all k then x is *generic* for μ (as was defined in §1). Note that if x is quasi-generic for μ then $\mu \in \mathcal{P}(X_x, S)$.

The goal of this subsection is to show that for any RAP sequence $x \in \mathcal{A}^{\mathbb{N}}$ there exists a measure μ for which x is generic and the corresponding measure-preserving system (X_x, μ, S) has rational discrete spectrum. As a matter of fact, we will prove a slightly more general theorem; see Theorem 3.12 below.

A partly related result was proved by Iwanik in [**39**, Theorem 2]. He showed that any sequence $x \in \mathcal{A}^{\mathbb{N}}$ in the class of so-called *Weyl almost periodic sequences* (which is a class that contains all WRAP sequences) is generic for a measure μ and the measure-preserving system (X_x, μ, S) has discrete spectrum (but not necessarily rational discrete spectrum). We refer the reader to [**39**] for the definitions of Weyl almost periodic sequences. Our variant of Iwanik's result regarding RAP sequences seems to be new. The authors would like to thank B. Weiss for fruitful discussions on the subject.

THEOREM 3.12. *Let $x \in \mathcal{A}^{\mathbb{N}}$ be RAP along $(N_k)_{k\geqslant 1}$. There exists a measure μ for which x is quasi-generic along $(N_k)_{k\geqslant 1}$ and the corresponding measure-preserving system (X_x, μ, S) is ergodic and has rational discrete spectrum.*

Note that Theorem 1.7 follows immediately from Theorem 3.12, if we put $N_k = k$ for all $k \geqslant 1$.

Remark 3.13. Some special cases of Theorem 3.12 are known. It is shown in [**17**, **56**]—in the context of the squarefree subshift (X_Q, S)—that the characteristic function of the set of squarefree numbers is generic for a measure which yields a measure-preserving system with rational discrete spectrum. This result has been generalized to arbitrary sets of $\mathscr{B}$-free numbers $\mathcal{F}_{\mathscr{B}}$ (see [**4**, **28**]). Also, for regular Toeplitz sequences it is shown in [**40**, §4] that the corresponding Toeplitz system has rational discrete spectrum (with respect to the unique invariant measure).

The proof of Theorem 3.12 hinges on four lemmas. The first two lemmas, namely Lemmas 3.14 and 3.15, are needed to prove that any RAP sequence is generic for an invariant probability measure. Lemma 3.16 shows that the measure obtained this way is ergodic. Finally, Lemma 3.17 proves that the corresponding system has rational discrete spectrum. We conclude this subsection by combining these four lemmas to give a proof of Theorem 3.12.

Given $(\alpha_1, \ldots, \alpha_\ell) \in \mathcal{A}^\ell$ and $n_1 < \cdots < n_\ell$, we define the corresponding cylinder set

$$C = C^{\alpha_1,\ldots,\alpha_\ell}_{n_1,\ldots,n_\ell} := \{x \in \mathcal{A}^{\mathbb{Z}} : x(n_j) = \alpha_j \text{ for } j = 1, \ldots, \ell\}.$$

Each cylinder set is a clopen subset of $\mathcal{A}^{\mathbb{Z}}$ and the family of cylinder sets forms a basis of topology on $\mathcal{A}^{\mathbb{Z}}$. More generally, for every subshift (X, S) a basis of topology is given by the clopen sets of the from $C^{\alpha_1,\ldots,\alpha_\ell}_{n_1,\ldots,n_\ell} \cap X$, $(\alpha_1, \ldots, \alpha_\ell) \in \mathcal{A}^\ell$ and $n_1 < \cdots < n_\ell \in \mathbb{Z}$.

LEMMA 3.14. *Let $x, y \in \mathcal{A}^{\mathbb{N}}$ and let $C = C^{\alpha_1,\ldots,\alpha_\ell}_{n_1,\ldots,n_\ell}$, where $n_1, \ldots, n_\ell \in \mathbb{Z}$ and $\alpha_1, \ldots, \alpha_\ell \in \mathcal{A}$. Then, for any two-sided sequences $\tilde{x}, \tilde{y} \in \mathcal{A}^{\mathbb{Z}}$ extending x and y,*

$$\limsup_{k\to\infty} \frac{1}{N_k} \sum_{n=1}^{N_k} |\mathbb{1}_C(S^n\tilde{x}) - \mathbb{1}_C(S^n\tilde{y})| \leqslant \ell d_B^{(N_k)}(x, y). \tag{3.7}$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Proof. Let $\mathcal{A}^{\Delta} := \{(a, a) : a \in \mathcal{A}\}$. For any $k \geqslant 1$,

$$\frac{1}{N_k}|\{1 \leqslant n \leqslant N_k : x(n) \neq y(n)\}| = \frac{1}{N_k}\sum_{n=1}^{N_k} \rho(x(n), y(n)).$$

It follows that

$$\limsup_{k\to\infty} \frac{1}{N_k}\sum_{n=1}^{N_k} \mathbb{1}_{\mathcal{A}^2\setminus\mathcal{A}^{\Delta}}(x(n), y(n)) \leqslant d_B^{(N_k)}(x, y). \tag{3.8}$$

In view of (3.8), the left-hand side of (3.7) can be estimated by

$$\begin{aligned}
&\limsup_{k\to\infty} \frac{1}{N_k}\sum_{n=1}^{N_k} |\mathbb{1}_A(S^n\tilde{x}) - \mathbb{1}_A(S^n\tilde{y})| \\
&\quad\leqslant \limsup_{k\to\infty} \frac{1}{N_k}\sum_{n=1}^{N_k}\sum_{i=1}^{\ell} \mathbb{1}_{\mathcal{A}^2\setminus\mathcal{A}^{\Delta}}(x(n+n_i), y(n+n_i)) \\
&\quad\leqslant \sum_{i=1}^{\ell} \limsup_{k\to\infty} \frac{1}{N_k}\sum_{n=1}^{N_k} \mathbb{1}_{\mathcal{A}^2\setminus\mathcal{A}^{\Delta}}(x(n+n_i), y(n+n_i)) \\
&\quad\leqslant \sum_{i=1}^{\ell} d_B^{(N_k)}(x, y).
\end{aligned}$$

This completes the proof of (3.7). □

LEMMA 3.15. *Let $x, x_n \in \mathcal{A}^{\mathbb{N}}$, $n \in \mathbb{N}$, and suppose $\lim_{n\to\infty} d_B^{(N_k)}(x_n, x) = 0$. If x_n is quasi-generic along $(N_k)_{k\geqslant 1}$ for all $n \in \mathbb{N}$, then x is quasi-generic along $(N_k)_{k\geqslant 1}$.*

Proof. To show that x is quasi-generic, it suffices to show that for all continuous functions $f : \mathcal{A}^{\mathbb{Z}} \to \mathbb{R}$ the limit

$$\lim_{k\to\infty} \frac{1}{N_k}\sum_{n=1}^{N_k} f(S^n\tilde{x}) \tag{3.9}$$

exists, where $\tilde{x} \in \mathcal{A}^{Z}$ is any two-sided sequence extending $x \in \mathcal{A}^{\mathbb{N}}$. Note that any continuous function $f : \mathcal{A}^{\mathbb{Z}} \to \mathbb{R}$ can be approximated uniformly by linear combinations of characteristic functions of cylinder sets $A = C_{n_1,\dots,n_\ell}^{\alpha_1,\dots,\alpha_\ell}$. Hence, we can assume without loss of generality that the function f in (3.9) is given by the indicator function of such a cylinder set.

Let $\varepsilon > 0$ be arbitrary and pick $m \geqslant 1$ such that $d_B^{(N_k)}(x_m, x) < \varepsilon$. Let $\tilde{x}_m \in \mathcal{A}^{\mathbb{Z}}$ be any two-sided sequence extending $x_m \in \mathcal{A}^{\mathbb{N}}$. Then, from Lemma 3.14, we deduce that the difference

$$\limsup_{k\to\infty} \frac{1}{N_k}\sum_{n=1}^{N_k} f(S^n\tilde{x}) - \liminf_{k\to\infty} \frac{1}{N_k}\sum_{n=1}^{N_k} f(S^n\tilde{x})$$

is bounded from above by

$$\limsup_{k\to\infty} \frac{1}{N_k}\sum_{n=1}^{N_k} f(S^n\tilde{x}_m) - \liminf_{k\to\infty} \frac{1}{N_k}\sum_{n=1}^{N_k} f(S^n\tilde{x}_m) + 2\ell d_B^{(N_k)}(x, x_m).$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

But x_m is quasi-generic along $(N_k)_{k\geqslant 1}$ and therefore

$$\limsup_{k\to\infty}\frac{1}{N_k}\sum_{n=1}^{N_k} f(S^n\tilde{x}_m)=\liminf_{k\to\infty}\frac{1}{N_k}\sum_{n=1}^{N_k} f(S^n\tilde{x}_m).$$

This implies that

$$\limsup_{k\to\infty}\frac{1}{N_k}\sum_{n=1}^{N_k} f(S^n\tilde{x})-\liminf_{k\to\infty}\frac{1}{N_k}\sum_{n=1}^{N_k} f(S^n\tilde{x})\leqslant 2\ell\epsilon.$$

Since $\varepsilon>0$ was arbitrary, this shows that the limit in (3.9) exists. □

Next, we need a slight generalization of [**63**, Proposition 4.6]. We include the proof for the convenience of the reader.

LEMMA 3.16. (See [**63**, Proposition 4.6] for the case $N_k=k$) *Suppose* $x, x_j\in\mathcal{A}^{\mathbb{N}}$, $j\in\mathbb{N}$, *and* $\lim_{j\to\infty} d_B^{(N_k)}(x_j,x)=0$. *If each* x_j *is quasi-generic along* $(N_k)_{k\geqslant 1}$ *for an ergodic measure, then* x *is quasi-generic along* $(N_k)_{k\geqslant 1}$ *for an ergodic measure.*

Proof. By Lemma 3.15, we know that x is quasi-generic along $(N_k)_{k\geqslant 1}$ for an invariant measure $\mu\in\mathcal{P}(\mathcal{A}^{\mathbb{Z}},S)$. It only remains to show that μ is ergodic. It suffices to show that for all $f,g\in L^\infty(\mathcal{A}^{\mathbb{Z}},\mu)$,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int_{\mathcal{A}^{\mathbb{Z}}} S^n f\cdot g\,d\mu=\int_{\mathcal{A}^{\mathbb{Z}}} f\,d\mu\int_{\mathcal{A}^{\mathbb{Z}}} g\,d\mu. \tag{3.10}$$

Similarly to the argument in the proof of Lemma 3.15, it suffices to prove (3.10) for the special case where f and g are the indicator functions of cylinder sets. In other words, we can assume without loss of generality that $f=\mathbb{1}_A$ and $g=\mathbb{1}_B$, where $A=C^{\alpha_1,\dots,\alpha_\ell}_{n_1,\dots,n_\ell}$ and $B=C^{\beta_1,\dots,\beta_r}_{m_1,\dots,m_r}$.

Fix $\varepsilon>0$. Let

$$\begin{aligned}C_n&:=S^{-n}A\cap B\\&=\{z\in\mathcal{A}^{\mathbb{Z}}: z(n+n_i)=\alpha_i,\ z(m_j)=\beta_j \text{ for } i=1,\dots,\ell,\ j=1,\dots,r\}.\end{aligned}$$

Then (3.10) can be rewritten as

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu(C_n)=\mu(A)\mu(B). \tag{3.11}$$

The sequence x is quasi-generic for μ along (N_k), so

$$\mu(C_n)=\lim_{k\to\infty}\frac{1}{N_k}\sum_{i=1}^{N_k}\mathbb{1}_{C_n}(S^i\tilde{x}),$$

for all $\tilde{x}\in\mathcal{A}^{\mathbb{Z}}$ that extend $x\in\mathcal{A}^{\mathbb{N}}$. Fix $\varepsilon>0$. In view of Lemma 3.14, for j sufficiently large, we obtain

$$\limsup_{k\to\infty}\frac{1}{N_k}\sum_{i=1}^{N_k}|\mathbb{1}_{C_n}(S^i\tilde{x})-\mathbb{1}_{C_n}(S^i\tilde{x}_j)|\leqslant\varepsilon, \tag{3.12}$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

for all $\tilde{x}_j \in \mathcal{A}^{\mathbb{Z}}$ that extend $x_j \in \mathcal{A}^{\mathbb{N}}$. Here, it is important that ϵ appearing in (3.12) does not depend on n. Denote by μ_j the measure for which the sequence x_j is quasi-generic along $(N_k)_{k\geqslant 1}$. It then follows from (3.12) that $|\mu(C_n) - \mu_j(C_n)| \leqslant \varepsilon$ for all n. A similar argument shows that if j is sufficiently large then $|\mu(A) - \mu_j(A)| \leqslant \varepsilon$ and $|\mu(B) - \mu_j(B)| \leqslant \varepsilon$.

Now, since μ_j is ergodic, we have that (3.11) holds with μ replaced by μ_j. Then, using the triangle inequality and the fact that $\mu_j(C_n)$, $\mu_j(A)$ and $\mu_j(B)$ are ε-close to $\mu(C_n)$, $\mu(A)$ and $\mu(B)$, respectively, we obtain that

$$\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{N}\mu(C_n) - \mu(A)\mu(B)\right| \leqslant 3\varepsilon.$$

Since ε was chosen arbitrarily, this proves (3.11). □

For the statement of the next lemma, we need to recall the definition of a joining. Let $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, R)$ be ergodic measure-preserving systems and let λ be a $(T \times R)$-invariant measure on $(X \times Y, \mathcal{B} \otimes \mathcal{C})$. We say that $(X \times Y, \mathcal{B} \otimes \mathcal{C}, \lambda, T \times R)$ is a *joining* of $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, R)$ if $\lambda|_X = \mu$ and $\lambda|_Y = \nu$ [**31**]. We will write $J((X, \mathcal{B}, \mu, T), (Y, \mathcal{C}, \nu, R))$ for the set of all joinings of $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, R)$. The subset of ergodic joinings will be denoted by $J^e((X, \mathcal{B}, \mu, T), (Y, \mathcal{C}, \nu, R))$. The definition of a joining extends naturally to any finite or countably infinite family of systems.

LEMMA 3.17. *Assume that $x, x^{(n)} \in \mathcal{A}^{\mathbb{N}}$ are quasi-generic along $(N_k)_{k\geqslant 1}$ for ergodic measures μ, μ_n $(n \geqslant 1)$, respectively. Assume, moreover, that $x^{(n)} \to x$ in $d_B^{(N_k)}$. Then (X_x, μ, S) is a factor of $((\mathcal{A}^{\mathbb{Z}})^{\times\infty}, \nu, S^{\times\infty})$ for some $\nu \in J^e((\mathcal{A}^{\mathbb{Z}}, \mu_1, S), (\mathcal{A}^{\mathbb{Z}}, \mu_2, S), \ldots)$.*

Proof. Consider

$$z := (x, x^{(1)}, x^{(2)}, \ldots) \in \mathcal{A}^{\mathbb{N}} \times (\mathcal{A}^{\mathbb{N}})^{\times\infty}.$$

Then z is quasi-generic along a subsequence of $(N_k)_{k\geqslant 1}$ for an invariant measure $\overline{\nu}$, i.e. for an increasing sequence $(k_\ell)_{\ell\geqslant 1}$, we have

$$\int_{\mathcal{A}^{\mathbb{Z}}\times(\mathcal{A}^{\mathbb{Z}})^{\times\infty}} f\, d\overline{\nu} = \lim_{\ell\to\infty}\frac{1}{N_{k_\ell}}\sum_{n=1}^{N_{k_\ell}} f((S \times (S^{\times\infty}))^n(\tilde{z})) \tag{3.13}$$

for all $f \in C(\mathcal{A}^{\mathbb{Z}} \times (\mathcal{A}^{\mathbb{Z}})^{\times\infty})$ and all $\tilde{z} \in \mathcal{A}^{\mathbb{Z}} \times (\mathcal{A}^{\mathbb{Z}})^{\times\infty}$ that extend z. Using the assumption of quasi-genericity along (N_k), we have

$$\overline{\nu} \in J((\mathcal{A}^{\mathbb{Z}}, \mu, S), (\mathcal{A}^{\mathbb{Z}}, \mu_1, S), (\mathcal{A}^{\mathbb{Z}}, \mu_2, S), \ldots).$$

We use $\mathcal{B}(\mathcal{A}^{\mathbb{Z}})$ and $\mathcal{B}((\mathcal{A}^{\mathbb{Z}})^{\times\infty})$ to denote the Borel σ-algebra on $\mathcal{A}^{\mathbb{Z}}$ and $(\mathcal{A}^{\mathbb{Z}})^{\times\infty}$, respectively. We claim now that $(\mathcal{A}^{\mathbb{Z}}, \mu, S)$ is a factor of $((\mathcal{A}^{\mathbb{Z}})^{\times\infty}, \overline{\nu}|_{(\mathcal{A}^{\mathbb{Z}})^{\times\infty}}, S^{\times\infty})$, i.e. that up to $\overline{\nu}$-measure zero sets the σ-algebra $\mathcal{B}(\mathcal{A}^{\mathbb{Z}}) \otimes \{\emptyset, (\mathcal{A}^{\mathbb{Z}})^{\times\infty}\}$ is contained in the σ-algebra $\{\emptyset, \mathcal{A}^{\mathbb{Z}}\} \otimes \mathcal{B}((\mathcal{A}^{\mathbb{Z}})^{\times\infty})$. Notice first that it is enough to show that for each $\alpha \in \mathcal{A}$,

$$C_0^\alpha = \{u \in \mathcal{A}^{\mathbb{Z}} : u(0) = \alpha\} \in \{\emptyset, \mathcal{A}^{\mathbb{Z}}\} \otimes \mathcal{B}((\mathcal{A}^{\mathbb{Z}})^{\times\infty}) \bmod \overline{\nu}, \tag{3.14}$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

as $\{C_0^\alpha : \alpha \in \mathcal{A}\}$ is a generating partition. To obtain (3.14), we note that for each $n \geqslant 1$,

$$\overline{\nu}((C_0^\alpha \times (\mathcal{A}^{\mathbb{Z}})^{\times\infty})\triangle(\mathcal{A}^{\mathbb{Z}} \times (\mathcal{A}^{\mathbb{Z}} \times \cdots \times \mathcal{A}^{\mathbb{Z}} \times \underbrace{C_0^\alpha}_{n\text{th position}} \times \mathcal{A}^{\mathbb{Z}} \times \cdots)))$$

$$\leqslant d_B^{(N_k)}(x, x^{(n)}). \tag{3.15}$$

Indeed, if $\pi_{0,n} : \mathcal{A}^{\mathbb{Z}} \times (\mathcal{A}^{\mathbb{Z}})^{\times\infty} \to \mathcal{A}^{\mathbb{Z}} \times \mathcal{A}^{\mathbb{Z}}$ denotes the projection

$$\pi_{0,n}(y, y^{(1)}, \ldots, y^{(n-1)}, y^{(n)}, y^{(n+1)}, \ldots) := (y, y^{(n)})$$

and $\overline{\nu}_{0,n}$ denotes the push-forward of $\overline{\nu}$ under $\pi_{0,n}$, then we obtain

$$\begin{aligned}
&\overline{\nu}((C_0^\alpha \times (\mathcal{A}^{\mathbb{Z}})^{\times\infty})\triangle(\mathcal{A}^{\mathbb{Z}} \times (\mathcal{A}^{\mathbb{Z}} \times \cdots \times \mathcal{A}^{\mathbb{Z}} \times \underbrace{C_0^\alpha}_{n} \times \mathcal{A}^{\mathbb{Z}} \times \cdots)))\\
&\quad = \overline{\nu}_{0,n}((C_0^\alpha \times \mathcal{A}^{\mathbb{Z}})\triangle(\mathcal{A}^{\mathbb{Z}} \times C_0^\alpha)) = \overline{\nu}_{0,n}(C_0^\alpha \times (C_0^\alpha)^c) + \overline{\nu}_{0,n}((C_0^\alpha)^c \times C_0^\alpha)\\
&\quad = \lim_{\ell\to\infty} \frac{1}{N_{k_\ell}} |\{0 \leqslant t \leqslant N_{k_\ell} - 1 : (S \times S)^t(x, x^{(n)}) \in (C_0^\alpha \times (C_0^\alpha)^c) \cup ((C_0^\alpha)^c \times C_0^\alpha)\}|\\
&\quad \leqslant d_B^{(N_k)}(x, x^{(n)}) \to 0 \quad \text{when } n \to \infty.
\end{aligned}$$

(We used the fact that the sets under consideration are clopen and hence (3.13) applies.) Since (3.15) holds, also (3.14) holds as any (complete) σ-algebra is closed in the metric $\overline{\nu}(\cdot\triangle\cdot)$.

We have shown that $(\mathcal{A}^{\mathbb{Z}}, \mu, S)$ is a measure-theoretical factor of the system

$$\mathcal{Z} := ((\mathcal{A}^{\mathbb{Z}})^{\times\infty}, \overline{\nu}|_{(\mathcal{A}^{\mathbb{Z}})^{\times\infty}}, S^{\times\infty}),$$

i.e. $(\mathcal{A}^{\mathbb{Z}}, \mu, S)$ is represented by an $S^{\times\infty}$-invariant sub-σ-algebra $\mathcal{C}$ of $(\mathcal{A}^{\mathbb{Z}})^{\times\infty}$. Consider the ergodic decomposition of $\mathcal{Z}$:

$$\overline{\nu}|_{(\mathcal{A}^{\mathbb{Z}})^{\times\infty}} = \int \kappa \, dQ(\kappa).$$

After the restriction to $\mathcal{C}$, we obtain

$$\mu = (\overline{\nu}|_{(\mathcal{A}^{\mathbb{Z}})^{\times\infty}})|_{\mathcal{C}} = \int \kappa|_{\mathcal{C}} \, dQ(\kappa).$$

Since, by Lemma 3.16, the system $(\mathcal{A}^{\mathbb{Z}}, \mu, S)$ is ergodic, it follows by the uniqueness of ergodic decomposition that $\kappa|_{\mathcal{C}} = \mu$ for Q-almost every κ. In other words, $(\mathcal{A}^{\mathbb{Z}}, \mu, S)$ is a measure-theoretical factor of almost every ergodic component of $\mathcal{Z}$. Moreover, because of the ergodicity of μ_n, $n \geqslant 1$, such an ergodic component is an ergodic joining of the family $\{(\mathcal{A}^{\mathbb{Z}}, \mu_n, S)\}_{n\in\mathbb{N}}$. It follows that a typical ergodic component ν satisfies the assertion of the lemma. □

Proof of Theorem 3.12. Suppose $x \in \mathcal{A}^{\mathbb{N}}$ is RAP along $(N_k)_{k\geqslant 1}$. By definition, we can find periodic points $x_n \in \mathcal{A}^{\mathbb{N}}$, $n \in \mathbb{N}$, such that x_n converges to x in the $d_B^{(N_k)}$ pseudo-metric.

Each x_n is generic for a cyclic rotation and hence, by Lemma 3.15, x is quasi-generic along $(N_k)_{k\geqslant 1}$ for some invariant measure $\mu \in \mathcal{P}(X_x, S)$. Following Lemma 3.16, we deduce that (X_x, μ, S) is ergodic.

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Finally, any ergodic joining of cyclic rotations exhibits rational discrete spectrum and therefore any factor of an ergodic joining of cyclic rotations also has rational discrete spectrum. However, it follows from Lemma 3.17 that $(\mathcal{A}^{\mathbb{Z}}, \mu, S)$, which is isomorphic to (X_x, μ, S), is a factor of a system given by such a joining, hence it has rational discrete spectrum. □

It is natural to inquire whether Theorem 3.12 characterizes RAP sequences. The answer is negative. In the following example we construct a subshift of $\{0, 1\}^{\mathbb{Z}}$ containing a point that is not RAP but is transitive and generic for an ergodic measure yielding a dynamical system with rational discrete spectrum.

Example 3.18. We will define a uniquely ergodic model for the cyclic rotation on two points (in such a model each point is generic for the unique invariant measure). The subshift $X \subset \{0, 1\}^{\mathbb{Z}}$ will consists of three orbits:

- the periodic point $a = \ldots 01.010101 \ldots$;
- the orbit of the point b which arises from the point a by erasing one '1';
- the orbit of the point c which arises from the point a by erasing infinitely many '1's so that the distance between the consecutive erased '1's goes to infinity.

It follows that in the orbit of c longer and longer ('periodic') words $0101 \ldots 01$ are approaching the periodic orbit of a from odd and even positions—this makes the point c non-rational (cf. Example 3.9). Since c is generic for the measure given by a, our claim follows.

Even though the system constructed in Example 3.18 contains a transitive point that is not RAP, it also contains an abundance of transitive points that are in fact RAP. As the following proposition shows, this is no coincidence.

PROPOSITION 3.19. *Let $\nu \in \mathcal{P}^e(\{0, 1\}^{\mathbb{Z}}, S)$ be such that $(\{0, 1\}^{\mathbb{Z}}, \nu, S)$ has rational discrete spectrum. Then ν-almost every $x \in \{0, 1\}^{\mathbb{Z}}$ is RAP. Moreover, if X is the topological support of ν, then there exists a transitive point $\eta \in X$ for which $\eta|_{[1,\infty)}$ is RAP.*

Proof. (Cf. [**6**, Theorem 3.19]) By assumption, the spectrum of $(\{0, 1\}^{\mathbb{Z}}, \nu, S)$ consists of roots of unity of degree n_t, with $n_t \mid n_{t+1}$, $t \geqslant 0$. Note that if $f_t \circ S = e^{2\pi i/n_t} \cdot f_t$, then $f_t^{n_t} \circ S = f_t^{n_t}$ and hence, by ergodicity, we can assume that f_t takes its values in the group $\{e^{2\pi ij/n_t} : j = 0, \ldots, n_t - 1\}$. By setting $D_0^t := \{x \in \{0, 1\}^{\mathbb{Z}} : f_t(x) = 1\}$, we obtain the partition

$$D_t := \{D_0^t, SD_0^t, \ldots, S^{n_t-1}D_0^t\}$$

of the space $\{0, 1\}^{\mathbb{Z}}$. Since $\{f_t : t \geqslant 0\}$ forms an orthonormal basis of $L^2(\nu)$, it follows that for each $t \geqslant 0$ there is a partition $Q^t = \{E_0^t, E_1^t\}$ of $\{0, 1\}^{\mathbb{Z}}$ such that E_i^t is a union of elements of the partition D_t $(i = 0, 1)$ satisfying

$$\nu(C_0^0 \triangle E_0^t) + \nu(C_0^1 \triangle E_1^t) \to 0 \quad \text{when } t \to \infty.$$

Since $S^{n_t} D_0^t = D_0^t$, the sequence $(\mathbb{1}_{E_1^t}(S^k x))_k$ is periodic of period n_t, for ν-almost every $x \in \{0, 1\}^{\mathbb{Z}}$. Moreover, ν-almost every point $x \in \{0, 1\}^{\mathbb{Z}}$ satisfies the ergodic theorem for

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

all sets $C_0^i \triangle E_i^t$, $t \geqslant 0$, $i = 0, 1$. Hence, the first part of the assertion follows from the pointwise ergodic theorem.

If, additionally, ν has full topological support then the orbit of ν-almost every point has to intersect any open set belonging to a countable basis of open sets. In other words, ν-almost every point is transitive, so the the second assertion follows from the first one. □

Remark 3.20. If a subshift (X, S) is a strictly ergodic and topologically mixing model of an odometer (note that such models exist due to [**44**]), then by Proposition 3.19 almost every $x \in X$ is an RAP point generating a topologically mixing subshift. Hence, RAP points can generate topologically mixing systems.

3.4. *Revisiting Theorem 1.4.* The purpose of this subsection is to give a proof of the following generalized form of Theorem 1.8 and discuss some applications thereof (see Theorem 3.27).

Definition 3.21. A set $R \subset \mathbb{N}$ is called an *averaging set of polynomial multiple recurrence along* $(M_k)_{k\geqslant 1}$ if for all invertible measure-preserving systems $(X, \mathcal{B}, \mu, T)$, $\ell \in \mathbb{N}$, $A \in \mathcal{B}$ with $\mu(A) > 0$ and for all polynomials $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, with $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$ for $i \in \{1, \ldots, \ell\}$,

$$\lim_{k\to\infty} \frac{1}{M_k} \sum_{n=1}^{M_k} \mathbb{1}_R(n)\mu(A \cap T^{-p_1(n)}A \cap \cdots \cap T^{-p_\ell(n)}A) > 0.$$

THEOREM 3.22. *Let $R \subset \mathbb{N}$ with $d^{(N_k)}(R) > 0$ and suppose $\eta := \mathbb{1}_R$ is quasi-generic along (N_k) for a measure ν such that (X_η, ν, S) has rational discrete spectrum. Then there exists a subsequence $(M_k)_{k\geqslant 1}$ of $(N_k)_{k\geqslant 1}$ such that the following are equivalent:*

(I) *R is divisible along $(M_k)_{k\geqslant 1}$, that is,*

$$d^{(M_k)}(R \cap u\mathbb{N}) := \lim_{k\to\infty} \frac{|R \cap u\mathbb{N} \cap \{1, \ldots, M_k\}|}{M_k} > 0$$

for all $u \in \mathbb{N}$;

(II) *R is an averaging set of polynomial multiple recurrence along $(M_k)_{k\geqslant 1}$.*

Remark 3.23. If R is RAP along $(N_k)_{k\geqslant 1}$ then it follows from the proof of Theorem 3.22 given below that in the statement of the theorem one can take $M_k = N_k$ for all $k \in \mathbb{N}$. On the other hand, if R is not RAP along $(N_k)_{k\geqslant 1}$ then this is not necessarily true. For instance, take $R \subset \mathbb{N}$ such that the sequence $\mathbb{1}_R$ equals the sequence x from Example 3.9. Then $\mathbb{1}_R$ is generic for a measure ν such that (X_η, ν, S) has rational discrete spectrum (see Remark 3.10). However, the set R is not divisible, since $d(R \cap 2\mathbb{Z})$ does not exist. For the same reason, $\mathbb{1}_R$ will not be a good weight for polynomial multiple convergence (see Definition 2.2). Therefore it is indeed necessary to pass to a subsequence $(M_k)_{k\geqslant 1}$ of $(N_k)_{k\geqslant 1}$ in Theorem 3.22.

For the proof of Theorem 3.22, we need the following variant of Lemma 2.9.

LEMMA 3.24. *Let $R \subset \mathbb{N}$ and suppose $\mathbb{1}_R = \eta$ is quasi-generic along $(N_k)_{k\geqslant 1}$ for a measure ν_η on $X_\eta \subset \{0, 1\}^{\mathbb{Z}}$ such that (X_η, ν_η, S) has rational discrete spectrum.*

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Let (X, T) be a topological dynamical system and let $\mu \in \mathcal{P}^e(X, T)$ be a measure with a generic point $x \in X$. If (X, μ, T) is totally ergodic then

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{n=1}^{N_k} \mathbb{1}_R(n) f(T^n x) = d^{(N_k)}(R) \int_X f \, d\mu$$

for each $f \in C(X)$.

Proof. Consider $(\tilde{\eta}, x) \in X_\eta \times X$, where $\tilde{\eta} \in \{0, 1\}^{\mathbb{Z}}$ is defined as $\tilde{\eta}(n) = \eta(n)$ for all $n \in \mathbb{N}$ and $\tilde{\eta}(n) = 0$ for all $n \in \mathbb{Z} \setminus \mathbb{N}$. Since (X, μ, T) is totally ergodic and (X_η, ν_η, S) has rational discrete spectrum, it follows that (X, μ, T) and (X_η, ν_η, S) are spectrally disjoint [**34**]. In particular, the only joining of these two systems is given by the product measure $\nu_\eta \otimes \mu$, i.e. $J((X_\eta, \nu_\eta, S), (X, \mu, T)) = \{\nu_\eta \otimes \mu\}$. It follows that

$$(\tilde{\eta}, x) \text{ is quasi-generic along } (N_k)_{k\geqslant 1} \text{ for the product measure } \nu_\eta \otimes \mu. \tag{3.16}$$

Fix $f \in C(X)$ and let $F\colon X_\eta \to \{0, 1\}$ be given by $F(z) = z(0)$. Then $F \in C(X_\eta)$ and, since η is quasi-generic along $(N_k)_{k\geqslant 1}$ for ν_η, we obtain

$$\int_{X_\eta} F \, d\nu_\eta = \lim_{k\to\infty} \frac{1}{N_k} \sum_{n=1}^{N_k} F(S^n \tilde{\eta}) = \lim_{k\to\infty} \frac{1}{N_k} \sum_{n=1}^{N_k} \mathbb{1}_R(n) = d^{(N_k)}(R).$$

In view of (3.16), we have

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{n=1}^{N_k} F \otimes f((S \times T)^n(\tilde{\eta}, x)) = \int F \otimes f \, d(\nu_\eta \otimes \mu) = \int_{X_\eta} F \, d\nu_\eta \int_X f \, d\mu.$$

Finally, we only need to observe that

$$\frac{1}{N_k} \sum_{n=1}^{N_k} F \otimes f((S \times T)^n(\tilde{\eta}, x)) = \frac{1}{N_k} \sum_{n=1}^{N_k} \mathbb{1}_R(n) f(T^n x)$$

and the proof is complete. □

LEMMA 3.25. *Let $R \subset \mathbb{N}$ and suppose $\mathbb{1}_R = \eta$ is quasi-generic along $(N_k)_{k\geqslant 1}$ for a measure ν on $X_\eta \subset \{0, 1\}^{\mathbb{Z}}$ such that (X_η, ν, S) has rational discrete spectrum. For $u \in \mathbb{N}$ and $j \in \{0, 1, \ldots, u-1\}$, let $(R-j)/u$ denote the set $\{n \in \mathbb{N} : nu + j \in R\}$. Then there exists a subsequence $(M_k)_{k\geqslant 1}$ of $(N_k)_{k\geqslant 1}$ with the property that for every $u \in \mathbb{N}$ and $j \in \{0, 1, \ldots, u-1\}$ the point $\eta_{u,j} := \mathbb{1}_{(R-j)/u}$ is quasi-generic along $(M_k/u)_{k\geqslant 1}$ for a measure $\nu_{u,j}$ such that $(X_{\eta_{u,j}}, \nu_{u,j}, S)$ has rational discrete spectrum.*

Proof. By applying a standard diagonalization method, choose a subsequence $(M_k)_{k\geqslant 1}$ of $(N_k)_{k\geqslant 1}$ such that for every $u \in \mathbb{N}$ the point η is quasi-generic along $(M_k/u)_{k\geqslant 1}$ for a measure μ_u with respect to the transformation S^u. In other words, for every $u \in \mathbb{N}$ and every continuous function $f \in C(\{0, 1\}^{\mathbb{Z}})$, the limit

$$\lim_{k\to\infty} \frac{u}{M_k} \sum_{n=1}^{M_k/u} f(S^{un} \tilde{\eta})$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

exists and equals $\int f\,d\mu_u$, where $\tilde{\eta} \in \{0,1\}^{\mathbb{Z}}$ is any two-sided sequence that extends $\eta \in \{0,1\}^{\mathbb{N}}$. Define $\mu_{u,j} := S^j \mu_u$ and note that $\mu_{u,j}$ is S^u-invariant. Since

$$\begin{aligned}\frac{1}{u}\sum_{j=0}^{u-1}\int f\,d\mu_{u,j} &= \frac{1}{u}\sum_{j=0}^{u-1}\int S^j f\,d\mu_u\\ &= \frac{1}{u}\sum_{j=0}^{u-1}\lim_{k\to\infty}\frac{u}{M_k}\sum_{n=1}^{M_k/u} f(S^{un+j}\tilde{\eta})\\ &= \lim_{k\to\infty}\frac{1}{M_k}\sum_{n=1}^{M_k} f(S^n\tilde{\eta})\\ &= \int f\,d\nu,\end{aligned}$$

we deduce that

$$\frac{1}{u}\sum_{j=0}^{u-1}\mu_{u,j} = \nu. \tag{3.17}$$

In particular, we have that for each Borel-measurable function g on X_η,

$$\|g\|^2_{L^2(\nu)} = \frac{1}{u}\sum_{j=0}^{u-1}\|g\|^2_{L^2(S^j\mu_u)} \geqslant \max_{0\leqslant j<u}\frac{1}{u}\|g\|^2_{L^2(\mu_{u,j})}. \tag{3.18}$$

We deduce from (3.17) that $\mu_{u,j}$ is absolutely continuous with respect to ν, that is, any set that has zero measure with respect to ν also has zero measure with respect to $\mu_{u,j}$. Therefore, any eigenfunction of the system (X_η, ν, S) with eigenvalue λ is an eigenfunction of the system $(X_\eta, \mu_{u,j}, S^u)$ with eigenvalue λ^u. The system (X_η, ν, S) has rational discrete spectrum and so the span of eigenfunctions with rational eigenvalue is dense in $L^2(\nu)$. However, if a class of bounded measurable functions is dense in $L^2(\nu)$, then, by (3.18), it is also dense in $L^2(\mu_{u,j})$. Hence, the span of eigenfunctions with rational eigenvalue is dense in $L^2(\mu_{u,j})$, which proves that $(X_\eta, \mu_{u,j}, S^u)$ has rational discrete spectrum.

Let $\Phi : \{0,1\}^{\mathbb{Z}} \to \{0,1\}^{\mathbb{Z}}$ denote the map defined by the rule $\Phi(x)(n) = x(un+j)$ for all $x \in \{0,1\}^{\mathbb{Z}}$. It is straightforward to verify that $\Phi(X_\eta) = X_{\eta_{u,j}}$ and that Φ satisfies

$$\Phi \circ S^u = S \circ \Phi. \tag{3.19}$$

Let $\nu_{u,j}$ denote the push-forward of $\mu_{u,j}$ under Φ. Since $S^j\eta$ is quasi-generic along $(M_k/u)_{k\geqslant 1}$ for $\nu_{u,j}$ under the transformation S^u, it follows from (3.19) that $\eta_{u,j}$ is generic for the measure $\nu_{u,j}$ along $(M_k/u)_{k\geqslant 1}$. Finally, observe that $(X_{\eta_{u,j}}, \nu_{u,j}, S)$ has rational discrete spectrum because $(X_\eta, \mu_{u,j}, S^u)$ has rational discrete spectrum. □

THEOREM 3.26. *Let $R \subset \mathbb{N}$ and suppose $\eta := \mathbb{1}_R$ is quasi-generic along (N_k) for a measure ν such that (X_η, ν, S) has rational discrete spectrum. Then there exists a subsequence $(M_k)_{k\geqslant 1}$ of $(N_k)_{k\geqslant 1}$ such that R is an averaging set of polynomial multiple convergence along $(M_k)_{k\geqslant 1}$, that is, for all invertible measure-preserving*

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

systems $(X, \mathcal{B}, \mu, T)$, $\ell \in \mathbb{N}$, $A \in \mathcal{B}$ *with* $\mu(A) > 0$ *and for all polynomials* $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, *with* $p_i(\mathbb{Z}) \subset \mathbb{Z}$ *for* $i \in \{1, \ldots, \ell\}$, *the limit*

$$\lim_{k\to\infty} \frac{1}{M_k} \sum_{n=1}^{M_k} \mathbb{1}_R(n)\mu(A \cap T^{-p_1(n)}A \cap \cdots \cap T^{-p_\ell(n)}A) \tag{3.20}$$

exists (cf. Definition 2.2).

Proof. By applying Lemma 3.25 we can find a subsequence $(M_k)_{k\geqslant 1}$ of $(N_k)_{k\geqslant 1}$ such that, for every $u \in \mathbb{N}$ and every $j \in \{0, 1, \ldots, u-1\}$, the point $\eta_{u,j} := \mathbb{1}_{(R-j)/u}$ is quasi-generic along $(M_k/u)_{k\geqslant 1}$ for a measure $\nu_{u,j}$ such that $(X_{\eta_{u,j}}, \nu_{u,j}, S)$ has rational discrete spectrum. Let $(X, \mathcal{B}, \mu, T)$, $\ell \in \mathbb{N}$, $A \in \mathcal{B}$ with $\mu(A) > 0$ and $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, with $p_i(\mathbb{Z}) \subset \mathbb{Z}$ for $i \in \{1, \ldots, \ell\}$ be arbitrary. Define

$$\varphi(n) := \mu(A \cap T^{-p_1(n)}A \cap \cdots \cap T^{-p_\ell(n)}A).$$

In view of Theorem 2.7, we can find for every $\epsilon > 0$ a basic nilsequence $(f(T_g^n x))$, where T_g is an ergodic nilrotation on some nilmanifold $X = G/\Gamma$, $f \in C(X)$ and $x \in X$, such that

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} |\varphi(n) - f(T_g^n x)| \leqslant \epsilon.$$

It thus suffices to show that the limit

$$\lim_{k\to\infty} \frac{1}{M_k} \sum_{n=1}^{M_k} \mathbb{1}_R(n) f(T_g^n x) \tag{3.21}$$

exists, because from this it follows that

$$\limsup_{k\to\infty} \frac{1}{M_k} \sum_{n=1}^{M_k} \mathbb{1}_R(n)\varphi(n) - \liminf_{k\to\infty} \frac{1}{M_k} \sum_{n=1}^{M_k} \mathbb{1}_R(n)\varphi(n) \leqslant 2\epsilon,$$

from which we can deduce that the limit in (3.20) exists, as ϵ was chosen arbitrarily.

Let $X_0, X_1, \ldots, X_{u-1}$ denote the connected components of the nilmanifold X. Since T_g is ergodic, it cyclically permutes the connected components of X. We can therefore assume without loss of generality that $T_g X_j = X_{j+1 \bmod u}$. In particular, $T_g^u X_j = X_j$ and, according to Proposition 2.8, the nilsystem (X_j, μ_{X_j}, T_g^u) is totally ergodic. Note that

$$\lim_{k\to\infty} \frac{1}{M_k} \sum_{n=1}^{M_k} \mathbb{1}_R(n) f(T_g^n x) = \sum_{j=0}^{u-1} \lim_{k\to\infty} \frac{u}{M_k} \sum_{n=1}^{M_k/u} \mathbb{1}_{(R-j)/u}(n) f(T_g^{un+j} x),$$

where the limit on the left-hand side in the above equation exists if all the limits for $j = 0, 1, \ldots, u-1$ on the right-hand side exist. It remains to show that for every $j \in \{0, 1, \ldots, u-1\}$ the limit

$$\lim_{k\to\infty} \frac{u}{M_k} \sum_{n=1}^{M_k/u} \mathbb{1}_{(R-j)/u}(n) f(T_g^{un+j} x)$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

exists. Suppose $T_g^j x \in X_{j_0}$ for some $j_0 \in \{0, 1, \ldots, u-1\}$. Since $(X_{j_0}, \mu_{X_j}, T_g^u)$ is totally ergodic (and uniquely ergodic), it follows from Lemma 3.24 that

$$\lim_{k\to\infty} \frac{u}{M_k} \sum_{n=1}^{M_k/u} \mathbb{1}_{(R-j)/u}(n) f(T_g^{un+j} x) = d^{(M_k)}((R-j)/u) \int f \, d\mu_{X_{j_0}}.$$

This finishes the proof. □

The proof of Theorem 3.22 hinges on Lemmas 3.24 and 3.25 and is a modification of the proof of Theorem 1.4 given in §2.3.

Proof of Theorem 3.22. Let $(X, \mathcal{B}, \mu, T)$ be an invertible measure-preserving system, let $R \subset \mathbb{N}$ with $d^{(N_k)}(R) > 0$ and assume $\eta := \mathbb{1}_R$ is quasi-generic along (N_k) for a measure ν such that (X_η, ν, S) has rational discrete spectrum. Take any $A \in \mathcal{B}$ with $\mu(A) > 0$ and let $p_1, \ldots, p_\ell \in \mathbb{Q}[t]$ with $p_i(\mathbb{Z}) \subset \mathbb{Z}$, $p_i(0) = 0$, $i = 1, \ldots, \ell$, be arbitrary. Choose a subsequence $(M_k)_{k\geqslant 1}$ of $(N_k)_{k\geqslant 1}$ such that the conclusions of both Theorem 3.26 and Lemma 3.25 hold. We will show that

$$\lim_{k\to\infty} \frac{1}{M_k} \sum_{n=1}^{M_k} \mathbb{1}_R(n)\varphi(n) > 0, \tag{3.22}$$

where $\varphi(n) = \mu(A \cap T^{-p_1(n)}A \cap \cdots \cap T^{-p_\ell(n)}A)$. The existence of the limit in (3.22) follows from Theorem 3.26. It remains to show that the limit in (3.22) is positive.

Arguing as in the proof of Theorem 2.6, we can assume without loss of generality that $(\varphi(n))$ is a nilsequence. By Theorem A.2 (see the appendix), there exists $\delta > 0$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \varphi(un) > \delta \quad \text{for all } u \in \mathbb{N}. \tag{3.23}$$

We can approximate $(\varphi(n))$ by a basic nilsequence $(f(T_g^n x))$, where T_g is a nilrotation on some nilmanifold $X = G/\Gamma$, $f \in C(X)$ and $x \in X$, such that $|\varphi(n) - f(T_g^n x)| \leqslant \delta/4$ for all $n \in \mathbb{N}$.

Using Proposition 2.8, we can find $u \in \mathbb{N}$ and a sub-nilmanifold $Y \subset X$ containing x such that (Y, μ_Y, T_{g^u}) is totally ergodic. It follows from Lemma 3.25 that $\mathbb{1}_{R/u}$ is quasi-generic along $(M_k/u)_{k\in\mathbb{N}}$ for a measure ν' such that the system $(\{0,1\}^{\mathbb{Z}}, \nu', S)$ has rational discrete spectrum. It now follows from Lemma 3.24 that

$$\lim_{k\to\infty} \frac{u}{M_k} \sum_{n=1}^{M_k/u} \mathbb{1}_{R/u}(n) f(T_{g^u}^n x) = d^{(M_k/u)}(R/u) \int_Y f \, d\mu_Y. \tag{3.24}$$

Finally, combining (3.23) and (3.24) and $|\varphi(un) - f(T_{g^u}^n x)| \leqslant \delta/4$, we obtain

$$\begin{aligned}\lim_{k\to\infty} \frac{1}{M_k} \sum_{n=1}^{M_k} \mathbb{1}_R(n)\psi(n) &\geqslant \lim_{k\to\infty} \frac{1}{M_k} \sum_{n=1}^{M_k} \mathbb{1}_{R\cap u\mathbb{N}}(n)\psi(n) \\ &= \frac{1}{u}\left(\lim_{k\to\infty} \frac{u}{M_k} \sum_{n=1}^{M_k/u} \mathbb{1}_{R/u}(n)\psi(un)\right)\end{aligned}$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

$$\geqslant \frac{1}{u}\left(\lim_{k\to\infty}\frac{u}{M_k}\sum_{n=1}^{M_k/u}\mathbb{1}_{R/u}(n)\,f(T_{g^u}^n x)-\frac{\delta}{4}d^{(M_k/u)}(R/u)\right)$$
$$\geqslant \frac{1}{u}\left(\frac{3\delta}{4}d^{(M_k/u)}(R/u)-\frac{\delta}{4}d^{(M_k/u)}(R/u)\right)>0.$$

This completes the proof. □

As an application of Theorem 3.22 together with Remark 3.23, we obtain a strengthening of Theorem 2.18 that also applies to $\mathscr{B}$-free numbers for a general set $\mathscr{B}\subset\mathbb{N}\setminus\{1\}$.

THEOREM 3.27. *Suppose $\mathscr{B}\subset\mathbb{N}\setminus\{1\}$. Then there exist an increasing sequence of positive integers $(N_k)_{k\geqslant 1}$ and a set $D\subset\mathcal{F}_{\mathscr{B}}$ with $d^{(N_k)}(\mathcal{F}_{\mathscr{B}}\setminus D)=0$ such that for all $r\in\mathbb{N}$ the following are equivalent:*

- *$r\in D$;*
- *$\mathcal{F}_{\mathscr{B}}-r$ is divisible along $(N_k)_{k\geqslant 1}$;*
- *$\mathcal{F}_{\mathscr{B}}-r$ is an averaging set of polynomial multiple recurrence along $(N_k)_{k\geqslant 1}$.*

Proof of Theorem 3.27. Let $\mathscr{B}\subset\mathbb{N}$ be arbitrary. If $\mathscr{B}$ is Behrend then $\mathcal{F}_{\mathscr{B}}$ has zero density and we can put $D=\emptyset$. Thus, let us assume that $\mathscr{B}$ is not Behrend. Therefore, the logarithmic density of $\delta(\mathcal{F}_{\mathscr{B}})$ is positive. Moreover, by Theorem 2.15, $\delta(\mathcal{F}_{\mathscr{B}})=d^{(N_k)}(\mathcal{F}_{\mathscr{B}})$ for some increasing sequence $(N_k)_{k\geqslant 1}$. We now repeat word for word the proof of Theorem 2.18 with density and divisibility replaced by density along $(N_k)_{k\geqslant 1}$ and divisibility along $(N_k)_{k\geqslant 1}$, respectively. □

4. *Applications to combinatorics*

In this section we show how the results obtained in the previous sections allow us to derive new refinements of the polynomial Szemerédi theorem. In particular, prove Theorems 1.9 and 1.10.

First, let us recall Furstenberg's correspondence principle.

PROPOSITION 4.1. (Furstenberg correspondence principle; see [**8**, **9**]) *Let $E\subset\mathbb{N}$ be a set with positive upper density $\overline{d}(E)>0$. Then there exist an invertible measure-preserving system $(X,\mathcal{B},\mu,T)$ and a set $A\in\mathcal{B}$ with $\mu(A)\geqslant\overline{d}(E)$ such that for all $n_1,\ldots,n_\ell\in\mathbb{N}$,*

$$\overline{d}(E\cap(E-n_1)\cap\cdots\cap(E-n_\ell))\geqslant\mu(A\cap T^{-n_1}A\cap\cdots\cap T^{-n_\ell}A). \tag{4.1}$$

We now have the following result regarding averaging sets of polynomial multiple recurrence along $(N_k)_{k\geqslant 1}$.

PROPOSITION 4.2. *Let $(N_k)_{k\geqslant 1}$ be an increasing sequence and let $R\subset\mathbb{N}$ be an averaging set of polynomial multiple recurrence along $(N_k)_{k\geqslant 1}$. Then for any set $E\subset\mathbb{N}$ with $\overline{d}(E)>0$ and any polynomials $p_1,\ldots,p_\ell\in\mathbb{Q}[t]$ which satisfy $p_i(\mathbb{Z})\subset\mathbb{Z}$ and $p_i(0)=0$ for all $i\in\{1,\ldots,\ell\}$, there exists $\beta>0$ such that the set*

$$\{n\in R:\overline{d}(E\cap(E-p_1(n))\cap\cdots\cap(E-p_\ell(n)))>\beta\}$$

has positive lower density (with respect to $(N_k)_{k\geqslant 1}$).

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

PROPOSITION 4.3. *Let $R \subset \mathbb{N}$ be an averaging set of polynomial multiple recurrence (along $(N_k)_{k\geqslant 1}$). Then for any $E \subset \mathbb{N}$ with $\overline{d}(E) > 0$ and any polynomials $p_1, \ldots, p_\ell \in \mathbb{Q}[t]$ which satisfy $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$ for all $i \in \{1, \ldots, \ell\}$, there exists a subset $R' \subset R$ satisfying $\overline{d}(R') > 0$ such that for any finite subset $F \subset R'$,*

$$\overline{d}\bigg(\bigcap_{n\in F}(E \cap (E - p_1(n)) \cap \cdots \cap (E - p_\ell(n)))\bigg) > 0.$$

By combining Proposition 4.2 with Theorem 1.4, we immediately obtain a proof of Theorem 1.9. Likewise, by combining Proposition 4.3 with Theorem 1.4, we immediately obtain a proof of Theorem 1.10.

We can also get a slight generalization of Theorem 1.9: we can replace the notions 'rational' and 'divisible' with 'rational along $(N_k)_{k\geqslant 1}$' and 'divisible along $(N_k)_{k\geqslant 1}$' for any increasing sequence $(N_k)_{k\geqslant 1}$ and, in virtue of Theorem 3.22, the statement of Theorem 1.9 remains valid.

Proposition 4.2 is an immediate consequence of Furstenberg's correspondence principle and of the definition of averaging sets of polynomial multiple recurrence.

For the proof of Proposition 4.3 we need the following theorem.

THEOREM 4.4. (See [**7**, Theorem 1.1]) *Let $(X, \mathcal{B}, \mu)$ be a probability space and suppose $A_n \in \mathcal{B}$, $\mu(A_n) \geqslant \delta > 0$, for $n = 1, 2, \ldots$. Then there exists a set $P \subset \mathbb{N}$ with $\overline{d}(P) \geqslant \delta$ such that for any finite subset $F \subset P$,*

$$\mu\bigg(\bigcap_{n\in F} A_n\bigg) > 0.$$

Proof of Proposition 4.3. Let $R \subset \mathbb{N}$ be an averaging set of polynomial multiple recurrence along $(N_k)_{k\geqslant 1}$. Let $E \subset \mathbb{N}$ with $\overline{d}(E) > 0$ and let $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, with $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$, for all $i \in \{1, \ldots, \ell\}$.

By applying Proposition 4.1, we can find an invertible measure-preserving system $(X, \mathcal{B}, \mu, T)$ and a set $A \in \mathcal{B}$ with $\mu(A) \geqslant \overline{d}(E)$ such that (4.1) is satisfied. Next, since R is an averaging set of polynomial multiple recurrence along $(N_k)_{k\geqslant 1}$, we can find some $\delta > 0$ such that the set

$$D := \{n \in R : \mu(A \cap T^{-p_1(n)}A \cap \cdots \cap T^{-p_\ell(n)}A) > \delta\}$$

has positive lower density, i.e. $\underline{d}(D) = \liminf_{N\to\infty} |D \cap \{1, \ldots, N\}|/N > 0$. Let $n_1, n_2, n_3, \ldots$ be an enumeration of D and let $A_i \in \mathcal{B}$ denote the set

$$A_i := A \cap T^{-p_1(n_i)}A \cap \cdots \cap T^{-p_\ell(n_i)}A.$$

Then, according to Theorem 4.4, we can find a set $P \subset \mathbb{N}$ with $\overline{d}(P) \geqslant \delta$ such that for any finite subset $F \subset P$,

$$\mu\bigg(\bigcap_{n\in F} A_n\bigg) > 0. \tag{4.2}$$

Let $R' := \{n_i : i \in P\}$. Then $R' \subset R$ and it is straightforward to show that $\overline{d}(R') > 0$. Moreover, combining (4.2) with (4.1), for any finite subset $\{n_1, \ldots, n_k\} \subset R'$, we obtain

$$\overline{d}\bigg(\bigcap_{i=1}^{r}(E \cap (E - p_1(n_i)) \cap \cdots \cap (E - p_\ell(n_i)))\bigg) > 0.$$

From this the claim follows immediately. □

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

For the special case of $\mathscr{B}$-free numbers, we have the following combinatorial corollary of the above results.

THEOREM 4.5. *For $\mathscr{B} \subset \mathbb{N} \setminus \{1\}$ let $\mathcal{F}_{\mathscr{B}}$ denote the set of $\mathscr{B}$-free numbers and let $(N_k)_{k\geqslant 1}$ be any sequence of increasing positive integers such that $d^{(N_k)}(\mathcal{F}_{\mathscr{B}})$ exists and is positive. Then there exists a set $D \subset \mathcal{F}_{\mathscr{B}}$ with $d^{(N_k)}(\mathcal{F}_{\mathscr{B}} \setminus D) = 0$ and such that for all $r \in D$, for all $E \subset \mathbb{N}$ with $\overline{d}(E) > 0$ and any polynomials $p_i \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, which satisfy $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$, for all $i \in \{1, \ldots, \ell\}$, there exists $\beta > 0$ such that the set*

$$\{n \in \mathcal{F}_{\mathscr{B}} - r : \overline{d}(E \cap (E - p_1(n_i)) \cap \cdots \cap (E - p_\ell(n_i))) > \beta\}$$

has positive lower density with respect to $(N_k)_{k\geqslant 1}$. If, additionally, $\mathscr{B}$ is taut then one can take $D = \mathcal{F}_{\mathscr{B}}$.

A proof of Theorem 4.5 follows immediately by combining Proposition 4.2 and Theorem 3.27.

5. *Rational sequences and Sarnak's conjecture*

Section 5 is divided into two subsections. In §5.1 we give a proof of Proposition 3.4, which states that any automatic sequence generated by a synchronized automaton is WRAP. In §5.2 we use Proposition 3.4 to strengthen a result obtained by Deshouillers, Drmota and Müllner in [**23**], which states that sequences given by synchronized automata satisfy Sarnak's conjecture.

5.1. *Synchronized automata and substitutions.*

We begin with a proof of Proposition 3.4. For the convenience of the reader, we restate the proposition here.

PROPOSITION 3.4. *Each automatic sequence given by a synchronized automaton M is WRAP.*

Proof. Let $M = (Q, \mathcal{B}, \delta, q_0, \tau)$ be a synchronized complete deterministic automaton with set of states $Q := \{q_0, \ldots, q_r\}$, input alphabet $\mathcal{B} := \{0, 1, \ldots, k-1\}$, finite output alphabet $\mathcal{A}$, transition function $\delta\colon Q \times \mathcal{B} \to Q$, initial state q_0 and output mapping $\tau : Q \to \mathcal{A}$. For $n \in \mathbb{N}$, let $[n]_k \in \mathcal{B}^*$ be defined as in §3.1. Let $a(n) = \tau(\delta(q_0, [n]_k))$, $n \in \mathbb{N}$, denote the automatic sequence generated by the synchronized automaton M.

Fix $\varepsilon > 0$. Let n_1 be such that at least $k^{n_1}(1-\varepsilon)$ words of length n_1 are synchronizing (see Definition 3.3). In other words, if we set

$$K := \{0 \leqslant m < k^{n_1} : [m]_k \text{ is synchronizing}\}$$

then we have $|K| \geqslant k^{n_1}(1-\varepsilon)$ (note that if w is a synchronizing word then so is every one of its extensions). Notice that

$$a(n) = a(m) \text{ whenever } n \equiv m \bmod k^{n_1} \text{ for some } m \in K \tag{5.1}$$

as $[n]_k$ and $[m]_k$ share the last k^{n_1} digits. Consider a' given by

$$a'(n) := \begin{cases} a(n) & \text{if } n \bmod k^{n_1} \text{ belongs to } K, \\ 0 & \text{otherwise.} \end{cases}$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Notice that a' is periodic of period k^{n_1}: for $0 \leqslant m < K$, $j \geqslant 0$,

$$a'(m + jk^{n_1}) = \begin{cases} a(m) & \text{if } m \in K, \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, using (5.1), we obtain

$$\begin{aligned} d_W(a, a') &= \limsup_{N\to\infty} \sup_{\ell} \frac{1}{N} |\{1 \leqslant n \leqslant N : a_{n+\ell} \neq a'_{n+\ell}\}| \\ &\leqslant \limsup_{N\to\infty} \sup_{\ell} \frac{1}{N} |\{1 \leqslant n \leqslant N : n + \ell \bmod k^{n_1} \notin K\}| \\ &= \frac{k^{n_1} - |K|}{k^{n_1}} \leqslant \varepsilon \end{aligned}$$

and the result follows. □

5.2. *Orthogonality of RAP and WRAP sequences to the Möbius function.* Let (X, T) be a topological system, that is, X is a compact metric space and $T: X \to X$ a homeomorphism. Let $\boldsymbol{\mu}$ denote the classical *Möbius function*, i.e. for all $n \in \mathbb{N}$,

$$\boldsymbol{\mu}(n) = \begin{cases} (-1)^k & \text{if there exist } k \text{ distinct prime numbers } p_1, \ldots, p_k, \\ & \text{such that } n = p_1 \cdot \cdots \cdot p_k, \\ 0 & \text{otherwise.} \end{cases}$$

We write $(X, T) \perp \boldsymbol{\mu}$ whenever $\lim_{N\to\infty} (1/N) \sum_{n=1}^{N} f(T^n x)\boldsymbol{\mu}(n) = 0$ for all $f \in C(X)$ and $x \in X$. Sarnak's conjecture [**56**] states that

$$(X, T) \perp \boldsymbol{\mu} \text{ whenever the topological entropy of } T \text{ is zero.} \tag{5.2}$$

If $x \in \mathcal{A}^{\mathbb{N}}$ is an automatic sequence generated by a synchronized automata then its subword complexity is at most linear (see, for example, [**2**, Theorem 10.3.1]), which implies that the entropy of the dynamical system (X_x, S) is zero. It is therefore natural to ask if systems generated by such automatic sequences satisfy Sarnak's conjecture. This question was answered affirmatively in [**23**].

The next theorem states that any W-rational system satisfies Sarnak's conjecture. In view of Proposition 3.4, our result can be viewed as an extension of the main result in [**23**].

THEOREM 5.1. *Let $x \in \mathcal{A}^{\mathbb{N}}$ be WRAP. Then for all $f \in C(X_x)$ and $z \in X_x$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(S^n z)\boldsymbol{\mu}(n) = 0. \tag{5.3}$$

Equivalently, $(X_x, S) \perp \boldsymbol{\mu}$.

For the proof of Theorem 5.1 we need two lemmas. The first lemma is a slight modification of Lemma 3.14 involving the Weyl pseudo-metric d_W instead of the Besicovitch pseudo-metric d_B.

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

LEMMA 5.2. *Let* $x, y \in \mathcal{A}^{\mathbb{N}}$, $n_1, \ldots, n_\ell \in \mathbb{Z}$ *and* $\alpha_1, \ldots, \alpha_\ell \in \mathcal{A}$. *Then, for* $C = C^{\alpha_1,\ldots,\alpha_\ell}_{n_1,\ldots,n_\ell}$,

$$\limsup_{H\to\infty} \sup_{m\in\mathbb{N}} \frac{1}{H} \sum_{m\leqslant h<m+H} |\mathbb{1}_C(S^h\tilde{x}) - \mathbb{1}_C(S^h\tilde{y})| \leqslant \ell d_W(x, y),$$

where $\tilde{x}, \tilde{y} \in \mathcal{A}^{\mathbb{Z}}$ *are any two-sided sequences extending* x *and* y, *respectively.*

The proof of Lemma 5.2 is very similar to the proof of Lemma 3.14 and is omitted.

The next lemma, which is also needed for the proof of Theorem 5.1, states that RAP sequences are orthogonal to the Möbius function $\boldsymbol{\mu}$.

LEMMA 5.3. *Suppose* $x \in \mathcal{A}^{\mathbb{N}}$ *is RAP and* $f \in C(\mathcal{A}^{\mathbb{Z}})$. *Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(S^n\tilde{x})\boldsymbol{\mu}(n) = 0. \tag{5.4}$$

Proof. Since any continuous function $f \in C(X_x)$ can be approximated uniformly by cylinder sets $C = C^{\alpha_1,\ldots,\alpha_\ell}_{n_1,\ldots,n_\ell}$, it suffices to show (5.4) for the special case where $f = \mathbb{1}_C = \mathbb{1}_{C^{\alpha_1,\ldots,\alpha_\ell}_{n_1,\ldots,n_\ell}}$ for any $n_1, \ldots, n_\ell \in \mathbb{Z}$ and $\alpha_1, \ldots, \alpha_\ell \in \mathcal{A}$.

Hence, let $\ell \in \mathbb{N}$, $n_1, \ldots, n_\ell \in \mathbb{Z}$ and $\alpha_1, \ldots, \alpha_\ell \in \mathcal{A}$ be arbitrary. Fix $\epsilon > 0$. Since x is RAP we can find a periodic sequence $y \in \mathcal{A}^{\mathbb{N}}$ such that $d_B(x, y) \leqslant \epsilon/\ell$. Let $\tilde{y} \in \mathcal{A}^{\mathbb{Z}}$ be a two-sided periodic sequence that extends y. Then, using Lemma 3.14, we get

$$\limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_C(S^n\tilde{x})\boldsymbol{\mu}(n) - \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_C(S^n\tilde{y})\boldsymbol{\mu}(n) \right| \leqslant \ell d_B(x, y) = \epsilon. \tag{5.5}$$

It is a well-known fact that Dirichlet's prime number theorem along arithmetic progressions is equivalent to the assertion that for any periodic sequence $a(n)$ one has $\lim_{N\to\infty} (1/N) \sum_{n=1}^{N} a(n)\boldsymbol{\mu}(n) = 0$. In particular, $a(n) = \mathbb{1}_C(S^n\tilde{y})$ is a periodic sequence and hence

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_C(S^n\tilde{y})\boldsymbol{\mu}(n) = 0.$$

Therefore, (5.5) simplifies to

$$\limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_C(S^n\tilde{x})\boldsymbol{\mu}(n) \right| \leqslant \epsilon.$$

Since $\epsilon > 0$ was chosen arbitrarily, the proof of (5.4) is complete. □

Proof of Theorem 5.1. Let $x \in \mathcal{A}^{\mathbb{N}}$ be WRAP and let $f \in C(X_x)$ and $z \in X_x$ be arbitrary. It follows from Proposition 3.6 that $z|_{\mathbb{N}}$ is WRAP and therefore $z|_{\mathbb{N}}$ is also RAP. Hence (5.3) follows directly from (5.4). □

In light of Theorem 5.1 it is natural to inquire about the behavior of averages of the from

$$\frac{1}{H} \sum_{m\leqslant h<m+H} f(S^h z)\boldsymbol{\mu}(n) \tag{5.6}$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

for large values of H and arbitrary $m \in \mathbb{N}$. It is believed that the expression in (5.6) does not converge to 0 (as H approaches ∞) uniformly in m. Indeed, by Chowla's conjecture [**18**] (see also [**27**, **56**]) it follows that for every word $w \in \{0, 1\}^H$ that appears in $\boldsymbol{\mu}^2 = \mathbb{1}_Q$, where Q denotes the set of squarefree numbers, all words in $v \in \{-1, 0, 1\}^H$ with $v^2 = w$ must appear in $\boldsymbol{\mu}$. In particular (assuming Chowla's conjecture), for every $H \geqslant 1$ there is $m \geqslant 1$ such that $\mathbb{1}_Q(h) = \boldsymbol{\mu}(m + h)$ for all $h \in [1, H]$ and therefore (5.6) with $f = 1$ is close to $6/\pi^2$. Nonetheless, using recent results of Matomäki, Radziwiłł and Tao [**49**], we will show that for large H and 'typical' $m \in \mathbb{N}$ the averages in (5.6) are small. Such averages of $\boldsymbol{\mu}$ (or, more generally, of bounded multiplicative functions) over 'short intervals' have also been considered in [**28**, **48**, **49**, **62**]. We obtain the following result in this direction.

THEOREM 5.4. *Let $x \in \mathcal{A}^{\mathbb{N}}$ be WRAP, let $f \in C(X_x)$ and let $z \in X_x$. Then for every $\delta > 0$ there exists $H_0 \in \mathbb{N}$ such that for all $H \geqslant H_0$ the set of all $m \in \mathbb{N}$ for which*

$$\left| \frac{1}{H} \sum_{m \leqslant h < m+H} f(S^h z)\boldsymbol{\mu}(h) \right| < \delta \tag{5.7}$$

has lower density greater than or equal to $1 - \delta$.

It is not clear if Theorem 5.1 can be derived quickly from Theorem 5.4. However, we will see that Theorem 5.4 is a corollary of a stronger result which is a strengthening of Theorem 5.1 and which we state next.

THEOREM 5.5. *Let $x \in \mathcal{A}^{\mathbb{N}}$ be WRAP. Then for all $f \in C(X_x)$ and $z \in X_x$,*

$$\lim_{\substack{H\to\infty \\ H/M\to 0}} \frac{1}{M} \sum_{M \leqslant m < 2M} \left| \frac{1}{H} \sum_{m \leqslant h < m+H} f(S^h z)\boldsymbol{\mu}(h) \right| = 0. \tag{5.8}$$

Before providing a proof of Theorem 5.5, let us show that Theorem 5.5 implies both Theorems 5.1 and 5.4. We will need the following standard lemma, the proof of which is included for the convenience of the reader.

LEMMA 5.6. *For every $H \in \mathbb{N}$, let $x_H : \mathbb{N} \to \mathbb{C}$ be a sequence bounded in modulus by 1. If*

$$\lim_{\substack{H\to\infty \\ H/M\to 0}} \frac{1}{M} \sum_{M \leqslant m < 2M} x_H(m) = 0,$$

then

$$\lim_{\substack{H\to\infty \\ H/N\to 0}} \frac{1}{N} \sum_{n=1}^{N} x_H(n) = 0.$$

Proof. Let H_k and N_k be two sequences such that $\lim_{k\to\infty} H_k = \infty$ and $\lim_{k\to\infty} (H_k/N_k) = 0$. Let $\ell \in \mathbb{N}$ be arbitrary. Then have

$$\left| \frac{1}{N_k} \sum_{n=1}^{N_k} x_{H_k}(n) \right| = \left| \sum_{1 \leqslant j \leqslant \log_2(N_k)} \frac{1}{2^j} \left(\frac{2^j}{N_k} \sum_{N_k/2^j \leqslant m < N_k/2^{j-1}} x_{H_k}(m) \right) + \frac{1}{N_k} x_{H_k}(N_k) \right|$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

$$\leqslant \sum_{1\leqslant j\leqslant \ell} \frac{1}{2^j}\left|\frac{2^j}{N_k}\sum_{N_k/2^j\leqslant m<N_k/2^{j-1}} x_{H_k}(m)\right| + \frac{1}{2^\ell} + \frac{1}{N_k},$$

whenever $\ell \leqslant \log_2(N_k)$. Note that

$$\lim_{k\to\infty} \frac{2^j}{N_k}\sum_{N_k/2^j\leqslant m<N_k/2^{j-1}} x_{H_k}(m) = 0,$$

because $\lim_{k\to\infty} H_k/N_k/2^j = 0$ for all $j \in \{1, \dots, \ell\}$. Hence,

$$\limsup_{k\to\infty} \frac{1}{N_k}\left|\sum_{n=1}^{N_k} x_{H_k}(n)\right| \leqslant \frac{1}{2^\ell}.$$

Since $\ell \in \mathbb{N}$ was arbitrary, this finishes the proof. □

Proof that Theorem 5.5 implies Theorem 5.1. Define

$$b_m(H) := \frac{1}{H}\left|\sum_{m\leqslant h<m+H} f(S^h z)\boldsymbol{\mu}(h)\right|.$$

First, we observe that, according to Lemma 5.6,

$$\frac{1}{M}\sum_{M\leqslant m<2M} b_m(H) \xrightarrow{H\to\infty, H/M\to 0} 0$$

implies

$$\frac{1}{N}\sum_{n=1}^{N} b_n(H) \xrightarrow{H\to\infty, H/N\to 0} 0.$$

Let $\epsilon > 0$ be arbitrary, let H_k and N_k be two sequences such that $\lim_{k\to\infty} H_k = \infty$ and $\lim_{k\to\infty}(H_k/N_k) = 0$ and take $J_k := \{1 \leqslant n \leqslant N_k : b_n(H_k) \leqslant \epsilon^2\}$. It follows from $\lim_{k\to\infty}(1/N_k)\sum_{n=1}^{N_k} b_n(H_k) = 0$ that for sufficiently large k we have $|J_k|/N_k \geqslant 1-\epsilon$. For $t \in \{0, 1, \dots, H_k - 1\}$ define $J_{k,t} := J_k \cap (H_k\mathbb{Z} + t)$. Then for some $r \in \{0, 1, \dots, H_k - 1\}$ we must have

$$\frac{|J_{k,r}|}{|(H_k\mathbb{Z}+r)\cap\{1,\dots,N_k\}|} \geqslant 1-\epsilon.$$

We get

$$\begin{aligned}\left|\frac{1}{N_k}\sum_{n=1}^{N_k} f(S^h z)\boldsymbol{\mu}(h)\right| &\leqslant \left|\frac{H_k}{N_k}\sum_{n\in(H_k\mathbb{Z}+r)\cap\{1,\dots,N_k\}} \frac{1}{H_k}\sum_{h=n}^{n+H_k-1} f(S^h z)\boldsymbol{\mu}(h)\right| + \frac{H_k}{N_k}\\ &\leqslant \frac{H_k}{N_k}\sum_{n\in(H_k\mathbb{Z}+r)\cap\{1,\dots,N_k\}} b_n(H_k) + \frac{H_k}{N_k}\\ &\leqslant \frac{H_k}{N_k}\sum_{n\in J_{k,r}} b_n(H_k) + \epsilon + \frac{H_k}{N_k} \leqslant \frac{H_k|J_{k,r}|\epsilon^2}{N_k} + \epsilon + \frac{H_k}{N_k}.\end{aligned}$$

As $k\to\infty$ the expression $H_k|J_{k,r}|\epsilon^2/N_k + \epsilon + H_k/N_k$ converges to $\epsilon^2 + \epsilon$. Since ϵ is arbitrary, this finishes the proof. □

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Proof that Theorem 5.5 implies Theorem 5.4. We present a proof by contradiction. Assume there exists some $\delta > 0$ such that one can find an increasing sequence $(H_k)_{k\geqslant 1}$ with the property that for every k the set

$$D_k := \left\{ m \in \mathbb{N} : \left| \frac{1}{H_k} \sum_{m \leqslant h < m+H_k} f(S^h z)\boldsymbol{\mu}(h) \right| \geqslant \delta \right\}$$

satisfies $\overline{d}(D_k) \geqslant \delta$. Since $\overline{d}(D_k) \geqslant \delta$, we can find $N_k \in \mathbb{N}$ such that $H_k \leqslant N_k/k$ and such that

$$\frac{1}{N_k} \sum_{m=1}^{N_k} \left| \frac{1}{H_k} \sum_{m \leqslant h < m+H_k} f(S^h z)\boldsymbol{\mu}(h) \right| \geqslant \delta^2.$$

This contradicts the fact that according to Theorem 5.5 and Lemma 5.6,

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{m=1}^{N_k} \left| \frac{1}{H_k} \sum_{m \leqslant h < m+H_k} f(S^h z)\boldsymbol{\mu}(h) \right| = 0. \qquad \square$$

Remark 5.7. From Theorem 5.4 it follows that for all $x \in \mathcal{A}^{\mathbb{N}}$ that are WRAP, $f \in C(X_x)$ and $z \in X_x$, we have

$$\lim_{H\to\infty} \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left| \frac{1}{H} \sum_{h=1}^{H} f(S^{n+h} z)\boldsymbol{\mu}(n+h) \right| = 0.$$

The remainder of this section is dedicated to proving Theorem 5.5. The following lemma (which is a variant of Lemma 5.3) will be useful for this purpose.

LEMMA 5.8. *Suppose $x \in \mathcal{A}^{\mathbb{N}}$ is WRAP and $f \in C(X_x)$. Then*

$$\lim_{\substack{H\to\infty \\ H/M\to 0}} \frac{1}{M} \sum_{M \leqslant m < 2M} \left| \frac{1}{H} \sum_{m \leqslant h < m+H} f(S^h \tilde{x})\boldsymbol{\mu}(h) \right| = 0. \tag{5.9}$$

where $\tilde{x} \in \mathcal{A}^{\mathbb{Z}}$ is any two-sided sequence extending $x \in \mathcal{A}^{\mathbb{N}}$.

Proof. Since any continuous function $f \in C(X_x)$ can be approximated uniformly by cylinder sets $C = C^{\alpha_1,\dots,\alpha_\ell}_{n_1,\dots,n_\ell}$, it suffices to show (5.9) for indicator functions of cylinder sets.

Let $\ell \in \mathbb{N}$, $n_1, \dots, n_\ell \in \mathbb{Z}$ and $\alpha_1, \dots, \alpha_\ell \in \mathcal{A}$ be arbitrary. Fix $\epsilon > 0$ and let $y \in \mathcal{A}^{\mathbb{N}}$ be a periodic sequence such that $d_W(x, y) \leqslant \epsilon/\ell$. Let $\tilde{y} \in \mathcal{A}^{\mathbb{Z}}$ be a two-sided periodic sequence that extends y. From Lemma 5.2 it follows that

$$\limsup_{H\to\infty} \sup_{m\in\mathbb{N}} \left| \frac{1}{H} \sum_{m \leqslant h < m+H} (\mathbb{1}_C(S^n \tilde{x}) - \mathbb{1}_C(S^n \tilde{y}))\boldsymbol{\mu}(n) \right| \leqslant \ell d_W(x, y) \leqslant \epsilon. \tag{5.10}$$

By a recent result of Matomäki, Radziwiłł and Tao [**49**], for each periodic sequence $a(n)$,

$$\frac{1}{M} \sum_{M \leqslant m < 2M} \left| \frac{1}{H} \sum_{m \leqslant h < m+H} a(h)\boldsymbol{\mu}(h) \right| \to 0 \quad \text{as } H \to \infty,\ H/M \to 0.$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Choosing $a(n) = \mathbb{1}_C(S^n\tilde{y})$ we thus get

$$\frac{1}{M}\sum_{M\leqslant m<2M}\left|\frac{1}{H}\sum_{m\leqslant h<m+H}\mathbb{1}_C(S^h\tilde{y})\boldsymbol{\mu}(h)\right|\to 0 \quad \text{as } H\to\infty,\ H/M\to 0. \tag{5.11}$$

Combining (5.10) and (5.11), we obtain

$$\begin{aligned}&\lim_{\substack{H\to\infty\\(H/M)\to 0}}\frac{1}{M}\sum_{M\leqslant m<2M}\left|\frac{1}{H}\sum_{m\leqslant h<m+H}\mathbb{1}_C(S^h\tilde{x})\boldsymbol{\mu}(h)\right|\\ &\quad\leqslant \lim_{\substack{H\to\infty\\(H/M)\to 0}}\frac{1}{M}\sum_{M\leqslant m<2M}\left|\frac{1}{H}\sum_{m\leqslant h<m+H}\mathbb{1}_C(S^h\tilde{y})\boldsymbol{\mu}(h)\right|+\epsilon=\epsilon.\end{aligned}$$

Since ϵ is arbitrarily, the proof of (5.9) is complete. □

Proof of Theorem 5.5. The following argument is analogous to the one used in the proof of Theorem 5.1. Let $x\in\mathcal{A}^{\mathbb{N}}$ be WRAP and let $f\in C(X_x)$ and $z\in X_x$ be arbitrary. It follows from Proposition 3.6 that $z|_{\mathbb{N}}$ is WRAP. Therefore, equation (5.8) follows from (5.9). □

Acknowledgements. We thank the anonymous referee for many helpful comments. The first author gratefully acknowledges the support of the NSF under grant DMS-1500575. The second author's research is supported by Narodowe Centrum Nauki UMO-2014/15/B/ST1/03736 and by the Foundation for Polish Science (FNP). The third author's research was supported by Narodowe Centrum Nauki UMO-2014/15/B/ST1/03736 and the EU grant 'AOS', FP7-PEOPLE-2012-IRSES, No. 318910.

A. *Appendix. Uniformity of polynomial multiple recurrence*

In this appendix we derive a uniform version of the following polynomial multiple recurrence theorem obtained in [**12**].

THEOREM A.1. (See [**12**, Theorem A]) *Let $\ell, u\in\mathbb{N}$ and let $p_{i,j}\in\mathbb{Q}[t]$ be polynomials satisfying $p_{i,j}(\mathbb{Z})\subset\mathbb{Z}$ and $p_{i,j}(0)=0$, $i=1,\dots,\ell$, $j=1,\dots,u$. Then for any probability space $(X,\mathcal{B},\mu)$, any u-tuple of commuting invertible measure-preserving transformations $T_1,\dots,T_u$ on $(X,\mathcal{B},\mu)$ and any $A\in\mathcal{B}$ with $\mu(A)>0$,*

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu\left(A\cap\prod_{j=1}^{u}T_j^{-p_{1,j}(n)}A\cap\prod_{j=1}^{u}T_j^{-p_{2,j}(n)}A\cap\cdots\cap\prod_{j=1}^{u}T_j^{-p_{\ell,j}(n)}A\right)>0.$$

The uniform version in question is given by the following theorem (a special case of which was used in the proofs of Theorems 2.6 and 3.22).

THEOREM A.2. *For all $\ell, d\in\mathbb{N}$ and all $\varepsilon>0$ there exists $\delta>0$ such that the following holds. For any $u\in\mathbb{N}$, for any polynomials $p_{i,j}\in\mathbb{Q}[t]$, $i=1,\dots,\ell$, $j=1,\dots,u$, satisfying $\deg(p_{i,j})\leqslant d$, $p_{i,j}(\mathbb{Z})\subset\mathbb{Z}$, $p_{i,j}(0)=0$, for any probability space $(X,\mathcal{B},\mu)$,*

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

for any u-tuple of commuting invertible measure-preserving transformations $T_1, \dots, T_u$ on $(X, \mathcal{B}, \mu)$, for any $A \in \mathcal{B}$ with $\mu(A) \geqslant \varepsilon$ and for any $s \in \mathbb{N}$, has

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \mu\Bigg(A \cap \prod_{j=1}^{u} T_j^{-p_{1,j}(sn)} A \cap \prod_{j=1}^{u} T_j^{-p_{2,j}(sn)} A \cap \cdots \cap \prod_{j=1}^{u} T_j^{-p_{\ell,j}(sn)} A\Bigg) > \delta. \tag{A.1}$$

We remark that a slightly less general version of Theorem A.2 is stated in [**30**, Theorem 4.1] without a proof.

In the course of proving Theorem A.2 we will make use of the following equivalent combinatorial form of Theorem A.1.

THEOREM A.3. (See [**11**, Theorem 3.2]) *Let $\ell, u \in \mathbb{N}$, let $\varepsilon > 0$ and let $p_{i,j} \in \mathbb{Z}[t]$ be polynomials satisfying $p_{i,j}(0) = 0$, $i = 1, \dots, \ell$, $j = 1, \dots, u$. Then there exists a positive integer $N = N(\ell, u, \varepsilon, p_{i,j})$ such that for all sets $A \subset \mathbb{Z}^d$ with*

$$\frac{|A \cap [1, N]^u|}{N^u} > \varepsilon$$

there exist $n \in \mathbb{N}$ and $a \in A$ such that $a + (p_{i,1}(n), \dots, p_{i,u}(n)) \in A$ for all $i \in \{1, 2, \dots, \ell\}$.

We will need the following theorem, which is of independent interest and can be interpreted as a polynomial extension of [**11**, Theorem F2].

THEOREM A.4. *For every $\ell, d \in \mathbb{N}$ and every $\varepsilon > 0$ there exist $K \in \mathbb{N}$ and $\beta > 0$ such that for any probability space $(X, \mathcal{B}, \mu)$, any commuting invertible measure-preserving transformations $T_{i,j}$, $1 \leqslant i \leqslant \ell$ and $1 \leqslant j \leqslant d$, and any $A \in \mathcal{B}$ with $\mu(A) \geqslant \varepsilon$, there exists $n \in \{1, \dots, K\}$ such that*

$$\mu\Bigg(A \cap \prod_{j=1}^{d} T_{1,j}^{-n^j} A \cap \prod_{j=1}^{d} T_{2,j}^{-n^j} A \cap \cdots \cap \prod_{j=1}^{d} T_{\ell,j}^{-n^j} A\Bigg) > \beta. \tag{A.2}$$

Moreover,

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \mu\Bigg(A \cap \prod_{j=1}^{d} T_{1,j}^{-n^j} A \cap \prod_{j=1}^{d} T_{2,j}^{-n^j} A \cap \cdots \cap \prod_{j=1}^{d} T_{\ell,j}^{-n^j} A\Bigg) \geqslant \frac{\beta}{K^2}. \tag{A.3}$$

Proof. Let $u := d\ell$ and, for $1 \leqslant i \leqslant \ell$ and $1 \leqslant t \leqslant u$, define

$$p_{i,t}(n) = \begin{cases} n^j & \text{if } t = (i-1)d + j \text{ with } 1 \leqslant i \leqslant \ell \text{ and } 1 \leqslant j \leqslant d, \\ 0 & \text{otherwise.} \end{cases} \tag{A.4}$$

Let $K = N(\ell, u, \varepsilon/2, p_{i,t})$ as guaranteed by Theorem A.3. For the remainder of this proof let us call a set of the form $\{a\} \cup \{a + (p_{i,1}(n), \dots, p_{i,u}(n)) : 1 \leqslant i \leqslant \ell\}$ for some $a = (a_1, \dots, a_u) \in \mathbb{N}^u$ and $n \in \mathbb{N}$ a *basic arrangement*. Let J denote the collection of all basic

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

arrangements contained in $\{1, \dots, K\}^u$. Set $\beta := \varepsilon/4|J|$. We claim that (A.2) and (A.3) are satisfied with this choice of K and β.

Let $(X, \mathcal{B}, \mu)$ be an arbitrary probability space, let $T_{i,j}$, $1 \leqslant i \leqslant \ell$ and $1 \leqslant j \leqslant d$, be commuting invertible measure-preserving transformations on X and let $A \in \mathcal{B}$ with $\mu(A) \geqslant \varepsilon$. For $1 \leqslant t \leqslant u$, let $S_t := T_{i,j}$ where $(i, j) \in \{1, \dots, \ell\} \times \{1, \dots, d\}$ is such that $t = (i-1)d + j$. It thus follows from (A.4) that

$$\prod_{j=1}^{d} T_{i,j}^{n^j} = \prod_{t=1}^{u} S_t^{p_{i,t}(n)}. \tag{A.5}$$

Define

$$f(x) := \frac{1}{K^u} \sum_{(n_1,\dots,n_u)\in[1,K]^u} \mathbb{1}_A\Bigg(\prod_{t=1}^{u} S_t^{n_t}(x)\Bigg).$$

Clearly, f is a non-negative function and $\int_X f \, d\mu \geqslant \varepsilon$. Therefore the set $B := \{x \in X : f(x) \geqslant \varepsilon/2\}$ satisfies $\mu(B) \geqslant \varepsilon/2$. Also, for every $x \in B$ the set

$$E_x := \Bigg\{(n_1, \dots, n_u) \in [1, K]^u : \prod_{t=1}^{u} S_t^{n_t}(x) \in A\Bigg\}$$

has density at least $\varepsilon/2$ in $[1, K]^u$, i.e. $|E_x| \geqslant (\varepsilon/2)K^u$. By our choice of K, we are guaranteed to find at least one basic arrangement contained in E_x.

We have shown that for every $x \in B$ there exists a basic arrangement contained in $E_x \subset [1, K]^u$. Since there are $|J|$ basic arrangements in $[1, K]^u$, by the pigeonhole principle there exists a set $C \subset B$ with $\mu(C) \geqslant \varepsilon/2|J|$ such that E_x contains the same basic arrangement for every $x \in C$. Suppose this basic arrangement is given by $\{(a_1, \dots, a_u)\} \cup \{(a_1, \dots, a_u) + (p_{i,1}(n), \dots, p_{i,u}(n)) : 1 \leqslant i \leqslant \ell\}$. Let $C' := \prod_{t=1}^{u} S_t^{a_t} C$. Then for any $x' \in C'$ and any $i \in \{1, \dots, \ell\}$, if $x := \prod_{t=1}^{u} S_t^{-a_t}(x')$ then by (A.5) and the definition of E_x,

$$\prod_{j=1}^{d} T_{i,j}^{n^j}(x') = \prod_{t=1}^{u} S_t^{p_{i,t}(n)}(x') = \prod_{t=1}^{u} S_t^{a_t + p_{i,t}(n)}(x) \in A.$$

This shows that C' is contained in the intersection $A \cap \prod_{j=1}^{d} T_{1,j}^{-n^j} A \cap \prod_{j=1}^{d} T_{2,j}^{-n^j} A \cap \cdots \cap \prod_{j=1}^{d} T_{\ell,j}^{-n^j} A$. Since $\mu(C') = \mu(C) > \beta$, this finishes the proof of (A.2).

Next, we give a proof of (A.3). Let $M \geqslant 1$ be arbitrary. Note that for all m with $M(K-1) < m \leqslant MK$ and all k with $1 \leqslant k \leqslant K$ the products mk are pairwise distinct. For $1 \leqslant j \leqslant d$, $1 \leqslant m \leqslant M$ and $1 \leqslant i \leqslant \ell$, define $R_{i,j,m} := T_{i,j}^{m^j}$. It follows that

$$\begin{aligned}
&\frac{1}{MK^2} \sum_{n=1}^{MK^2} \mu\Bigg(A \cap \prod_{j=1}^{d} T_{1,j}^{-n^j} A \cap \prod_{j=1}^{d} T_{2,j}^{-n^j} A \cap \cdots \cap \prod_{j=1}^{d} T_{\ell,j}^{-n^j} A\Bigg) \\
&\quad \geqslant \frac{1}{MK^2} \sum_{m=M(K-1)+1}^{MK} \sum_{k=1}^{K} \mu\Bigg(A \cap \prod_{j=1}^{d} T_{1,j}^{-(mk)^j} A \cap \prod_{j=1}^{d} T_{2,j}^{-(mk)^j} A \cap \cdots \cap \prod_{j=1}^{d} T_{\ell,j}^{-(mk)^j} A\Bigg) \\
&\quad = \frac{1}{MK^2} \sum_{m=M(K-1)+1}^{MK} \sum_{k=1}^{K} \mu\Bigg(A \cap \prod_{j=1}^{d} R_{1,j,m}^{-k^j} A \cap \prod_{j=1}^{d} R_{2,j,m}^{-k^j} A \cap \cdots \cap \prod_{j=1}^{d} R_{\ell,j,m}^{-k^j} A\Bigg).
\end{aligned}$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

In light of (A.2) we have

$$\sum_{k=1}^{K} \mu\Big(A \cap \prod_{j=1}^{d} R_{1,j,m}^{-k^j} A \cap \prod_{j=1}^{d} R_{2,j,m}^{-k^j} A \cap \cdots \cap \prod_{j=1}^{d} R_{\ell,j,m}^{-k^j} A\Big) > \beta$$

for all $1 \leqslant m \leqslant M$. Therefore,

$$\begin{aligned}\frac{1}{MK^2} \sum_{m=M(K-1)+1}^{MK} \sum_{k=1}^{K} \mu\Big(A \cap \prod_{j=1}^{d} R_{1,j,m}^{-k^j} A \cap \prod_{j=1}^{d} R_{2,j,m}^{-k^j} A \cap \cdots \cap \prod_{j=1}^{d} R_{\ell,j,m}^{-k^j} A\Big) \\ > \frac{1}{MK^2} \sum_{m=M(K-1)+1}^{MK} \beta \\ = \frac{\beta}{K^2}.\end{aligned}$$

This proves that

$$\liminf_{M\to\infty} \frac{1}{MK^2} \sum_{n=1}^{MK^2} \mu\Big(A \cap \prod_{j=1}^{d} T_{1,j}^{-n^j} A \cap \prod_{j=1}^{d} T_{2,j}^{-n^j} A \cap \cdots \cap \prod_{j=1}^{d} T_{\ell,j}^{-n^j} A\Big) \geqslant \frac{\beta}{K^2}. \tag{A.6}$$

Finally, it follows from the results in [**61**] that the limits on the left-hand side of (A.6) and on the left-hand side of (A.3) exist and are equal. This finishes the proof of (A.3). □

Proof of Theorem A.2. Depending only on $\ell, d \in \mathbb{N}$ and $\epsilon > 0$, choose $\beta > 0$ and $K \geqslant 1$ as guaranteed by Theorem A.4. Note that coefficients of integer polynomials of degree d can be written as fractions with denominator $q := d!$. Define $b := q!$ and pick any $\delta > 0$ such that $\delta < \beta/bK^2$. We claim that (A.1) holds with this choice of δ.

Let $u, s \in \mathbb{N}$ and let $p_{i,j} \in \mathbb{Q}[t]$, $i = 1, \ldots, \ell$, $j = 1, \ldots, u$, with $\deg(p_{i,j}) \leqslant d$, $p_{i,j}(\mathbb{Z}) \subset \mathbb{Z}$, $p_{i,j}(0) = 0$ and such that the denominators of the coefficients of $p_{i,j}$ (when written as reduced fractions) are at most q. Furthermore, let $T_1, \ldots, T_u$ be commuting invertible measure-preserving transformations on a probability space $(X, \mathcal{B}, \mu)$ and let $A \in \mathcal{B}$ with $\mu(A) \geqslant \varepsilon$. It follows from [**61**] that the limit on the left-hand side of (A.1) exists and is equal to

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu\Big(A \cap \prod_{j=1}^{u} T_j^{-p_{1,j}(sn)} A \cap \prod_{j=1}^{u} T_j^{-p_{2,j}(sn)} A \cap \cdots \cap \prod_{j=1}^{u} T_j^{-p_{\ell,j}(sn)} A\Big). \tag{A.7}$$

It thus suffices to show that (A.7) is bigger than δ.

For $1 \leqslant i \leqslant \ell$ and $1 \leqslant j \leqslant u$, find $a_{i,j}^{(1)}, \ldots, a_{i,j}^{(d)}$ such that

$$p_{i,j}(n) = a_{i,j}^{(1)} n + a_{i,j}^{(2)} n^2 + \cdots + a_{i,j}^{(d)} n^d.$$

By assumption, $b a_{i,j}^{(k)} \in \mathbb{Z}$ and hence $s^k b^k a_{i,j}^{(k)} \in \mathbb{Z}$ for all $s \in \mathbb{N}$. Define

$$R_{i,k} := \prod_{j=1}^{u} T_j^{s^k b^k a_{i,j}^{(k)}}, \quad 1 \leqslant k \leqslant d.$$

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

Clearly,

$$\prod_{j=1}^{u} T_j^{p_{i,j}(bsn)} = \prod_{j=1}^{d} R_{i,j}^{n^j} \quad \text{for all } n \in \mathbb{N}.$$

We thus have

$$\begin{aligned}
&\frac{1}{N}\sum_{n=1}^{N} \mu\Bigg(A \cap \prod_{j=1}^{u} T_j^{-p_{1,j}(sn)} A \cap \prod_{j=1}^{u} T_j^{-p_{2,j}(sn)} A \cap \cdots \cap \prod_{j=1}^{u} T_j^{-p_{\ell,j}(sn)} A\Bigg) \\
&\quad \geqslant \frac{1}{N}\sum_{n=1}^{\lfloor N/b \rfloor} \mu\Bigg(A \cap \prod_{j=1}^{u} T_j^{-p_{1,j}(bsn)} A \cap \prod_{j=1}^{u} T_j^{-p_{2,j}(bsn)} A \cap \cdots \cap \prod_{j=1}^{u} T_j^{-p_{\ell,j}(bsn)} A\Bigg) \\
&\quad = \frac{1}{N}\sum_{n=1}^{\lfloor N/b \rfloor} \mu\Bigg(A \cap \prod_{j=1}^{d} R_{1,j}^{-n^j} A \cap \prod_{j=1}^{d} R_{2,j}^{-n^j} A \cap \cdots \cap \prod_{j=1}^{d} R_{\ell,j}^{-n^j} A\Bigg).
\end{aligned}$$

From (A.3) it follows that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{\lfloor N/b \rfloor} \mu\Bigg(A \cap \prod_{j=1}^{d} R_{1,j}^{-n^j} A \cap \prod_{j=1}^{d} R_{2,j}^{-n^j} A \cap \cdots \cap \prod_{j=1}^{d} R_{\ell,j}^{-n^j} A\Bigg) \geqslant \frac{\beta}{bK^2} > \delta.$$

Therefore,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu\Bigg(A \cap \prod_{j=1}^{u} T_j^{-p_{1,j}(sn)} A \cap \prod_{j=1}^{u} T_j^{-p_{2,j}(sn)} A \cap \cdots \cap \prod_{j=1}^{u} T_j^{-p_{\ell,j}(sn)} A\Bigg) > \delta.$$

□

References

[1] J.-P. Allouche. The number of factors in a paperfolding sequence. *Bull. Aust. Math. Soc.* **46** (1992), 23–32.

[2] J.-P. Allouche and J. Shallit. *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, Cambridge, 2003.

[3] L. Auslander, L. Green and F. Hahn. *Flows on Homogeneous Spaces (Annals of Mathematics Studies, 53)*. Princeton University Press, Princeton, NJ, 1963, with the assistance of L. Markus and W. Massey, and an appendix by L. Greenberg.

[4] A. Bartnicka, S. Kasjan, J. Kułaga-Przymus and M. Lemańczyk. $\mathscr{B}$-free sets and dynamics. *Trans. Amer. Math. Soc.*, to appear, doi:10.1090/tran/7132.

[5] M. Beiglböck, V. Bergelson, N. Hindman and D. Strauss. Multiplicative structures in additively large sets. *J. Combin. Theory Ser. A* **113** (2006), 1219–1242.

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

[7] V. Bergelson. Sets of recurrence of $\mathbf{Z}^m$-actions and properties of sets of differences in $\mathbf{Z}^m$. *J. Lond. Math. Soc. (2)* **31** (1985), 295–304.

[8] V. Bergelson. Ergodic Ramsey theory. *Logic and Combinatorics (Arcata, CA, 1985) (Contemporary Mathematics, 65)*. American Mathematical Society, Providence, RI, 1987, pp. 63–87.

[9] V. Bergelson. Ergodic Ramsey theory – an update. *Ergodic Theory of* $\mathbf{Z}^d$ *Actions (Warwick, 1993–1994) (London Mathematical Society Lecture Note Series, 228)*. Cambridge University Press, Cambridge, 1996, pp. 1–61.

[10] V. Bergelson and I. J. Håland. Sets of recurrence and generalized polynomials. *Convergence in Ergodic Theory and Probability (Columbus, OH, 1993) (Ohio State University Mathematics Research Institute Publications, 5)*. De Gruyter, Berlin, 1996, pp. 91–110.

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

[**11**] V. Bergelson, B. Host, R. McCutcheon and F. Parreau. Aspects of uniformity in recurrence. *Colloq. Math.* **84/85** (2000), 549–576.
[**12**] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.* **9** (1996), 725–753.
[**13**] V. Bergelson and I. Ruzsa. Squarefree numbers, IP sets and ergodic theory. *Paul Erdős and His Mathematics, I (Budapest, 1999) (Bolyai Society Mathematical Studies, 11)*. János Bolyai Mathematical Socety, Budapest, 2002, pp. 147–160.
[**14**] J. Berstel, D. Perrin and C. Reutenauer. *Codes and Automata (Encyclopedia of Mathematics and its Applications, 129)*. Cambridge University Press, Cambridge, 2010.
[**15**] A. S. Besicovitch. On the density of certain sequences of integers. *Math. Ann.* **110** (1935), 336–341.
[**16**] A. S. Besicovitch. *Almost Periodic Functions*. Dover Publications, New York, 1955.
[**17**] F. Cellarosi and Y. G. Sinai. Ergodic properties of square-free numbers. *J. Eur. Math. Soc. (JEMS)* **15** (2013), 1343–1374.
[**18**] S. Chowla. *The Riemann Hypothesis and Hilbert's Tenth Problem (Mathematics and Its Applications, 4)*. Gordon and Breach Science Publishers, New York, 1965.
[**19**] H. Davenport. Über numeri abundantes. *Sitzungsber. Preuss. Akad. Wiss.,Phys.-Math. Kl.* **6** (1933), 830–837.
[**20**] H. Davenport and P. Erdős. On sequences of positive integers. *Acta Arith.* **2** (1936), 147–151.
[**21**] H. Davenport and P. Erdős. On sequences of positive integers. *J. Indian Math. Soc. (N.S.)* **15** (1951), 19–24.
[**22**] M. Dekking, M. Mendès France and A. van der Poorten. Folds. *Math. Intelligencer* **4** (1982), 130–138.
[**23**] J.-M. Deshouillers, M. Drmota and C. Müllner. Automatic sequences generated by synchronizing automata fulfill the Sarnak conjecture. *Studia Math.* **231** (2015), 83–95.
[**24**] T. Downarowicz. Survey of odometers and Toeplitz flows. *Algebraic and Topological Dynamics (Contemporary Mathematics, 385)*. American Mathematical Society, Providence, RI, 2005, pp. 7–37.
[**25**] T. Downarowicz. *Entropy in Dynamical Systems (New Mathematical Monographs, 18)*. Cambridge University Press, Cambridge, 2011.
[**26**] T. Downarowicz and A. Iwanik. Quasi-uniform convergence in compact dynamical systems. *Studia Math.* **89** (1988), 11–25.
[**27**] E. El Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk and T. De La Rue. The Chowla and the Sarnak conjectures from ergodic theory point of view. *Discrete Contin. Dyn. Syst.* **37**(6) (2017), 2899–2944.
[**28**] E. El Abdalaoui, M. Lemańczyk and T. de la Rue. A dynamical point of view on the set of $\mathscr{B}$-free integers. *Int. Math. Res. Not. IMRN* **2015** (2015), 7258–7286.
[**29**] N. Frantzikinakis. Multiple ergodic averages for three polynomials and applications. *Trans. Amer. Math. Soc.* **360** (2008), 5435–5475.
[**30**] N. Frantzikinakis, B. Host and B. Kra. The polynomial multidimensional Szemerédi theorem along shifted primes. *Israel J. Math.* **194** (2013), 331–348.
[**31**] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory* **1** (1967), 1–49.
[**32**] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.* **31** (1977), 204–256.
[**33**] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory (M. B. Porter Lectures)*. Princeton University Press, Princeton, NJ, 1981.
[**34**] F. Hahn and W. Parry. Some characteristic properties of dynamical systems with quasi-discrete spectra. *Math. Systems Theory* **2** (1968), 179–190.
[**35**] R. R. Hall. *Sets of Multiples (Cambridge Tracts in Mathematics, 118)*. Cambridge University Press, Cambridge, 1996.
[**36**] N. Hindman. Ultrafilters and combinatorial number theory. *Number Rheory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, IL, 1979) (Lecture Notes in Mathematics, 751)*. Springer, Berlin, 1979, pp. 119–184.
[**37**] B. Hornfeck. Zur Dichte der Menge der vollkommenen Zahlen. *Arch. Math. (Basel)* **6** (1955), 442–443.
[**38**] B. Host and B. Kra. Convergence of polynomial ergodic averages. *Israel J. Math.* **149** (2005), 1–19.
[**39**] A. Iwanik. Weyl almost periodic points in topological dynamics. *Colloq. Math.* **56** (1988), 107–119.
[**40**] K. Jacobs and M. Keane. 0–1-sequences of Toeplitz type. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **13** (1969), 123–131.
[**41**] M. Keane. Generalized Morse sequences. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **10** (1968), 335–353.
[**42**] N. Kryloff and N. Bogoliouboff. La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire. *Ann. of Math. (2)* **38** (1937), 65–113.
[**43**] J. Kułaga-Przymus, M. Lemańczyk and B. Weiss. On invariant measures for $\mathscr{B}$-free systems. *Proc. Lond. Math. Soc. (3)* **110** (2015), 1435–1474.
[**44**] E. Lehrer. Topological mixing and uniquely ergodic systems. *Israel J. Math.* **57** (1987), 239–255.

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press

[**45**] A. Leibman. Convergence of multiple ergodic averages along polynomials of several variables. *Israel J. Math.* **146** (2005), 303–315.
[**46**] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold. *Ergod. Th. & Dynam. Syst.* **25** (2005), 201–213.
[**47**] A. Leibman. Multiple polynomial correlation sequences and nilsequences. *Ergod. Th. & Dynam. Syst.* **30** (2010), 841–854.
[**48**] K. Matomäki and M. Radziwiłł. Multiplicative functions in short intervals. *Ann. of Math. (2)* **183** (2016), 1015–1056.
[**49**] K. Matomäki, M. Radziwiłł and T. Tao. An averaged form of Chowla's conjecture. *Algebra Number Theory* **9** (2015), 2167–2196.
[**50**] M. Morse and G. A. Hedlund. Symbolic dynamics. *Amer. J. Math.* **60** (1938), 815–866.
[**51**] M. Morse and G. A. Hedlund. Symbolic dynamics II. Sturmian trajectories. *Amer. J. Math.* **62** (1940), 1–42.
[**52**] W. Parry. Ergodic properties of affine transformations and flows on nilmanifolds. *Amer. J. Math.* **91** (1969), 757–771.
[**53**] R. Peckner. Uniqueness of the measure of maximal entropy for the squarefree flow. *Israel J. Math.* **210** (2015), 335–357.
[**54**] K. Petersen. *Ergodic Theory (Cambridge Studies in Advanced Mathematics, 2)*. Cambridge University Press, Cambridge, 1983.
[**55**] M. Queffélec. *Substitution Dynamical Systems—Spectral Analysis (Lecture Notes in Mathematics, 1294)*. 2nd edn. Springer, Berlin, 2010.
[**56**] P. Sarnak. Three lectures on the Möbius function, randomness and dynamics. http://publications.ias.edu/sarnak/.
[**57**] I. Schoenberg. Über die asymptotische Verteilung reeller Zahlen mod 1. *Math. Z.* **28** (1928), 171–199.
[**58**] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.* **27** (1975), 199–245.
[**59**] A. Thue. Über unendliche Zeichenreihen. *Nor. Vid. Selsk. Skr. I Math-Nat. Kl.* **7** (1906).
[**60**] A. Thue. Die Lösung eines Spezialfalles eines generellen logischen Problems. *Nor. Vid. Selsk. Skr. I Math-Nat. Kl. Chris.* **8** (1910).
[**61**] M. N. Walsh. Norm convergence of nilpotent ergodic averages. *Ann. of Math. (2)* **175** (2012), 1667–1688.
[**62**] Z. Wang. Möbius disjointness for analytic skew products. *Invent. Math.* **209** (2017), 175–196.
[**63**] B. Weiss. *Single Orbit Dynamics (CBMS Regional Conference Series in Mathematics, 95)*. American Mathematical Society, Providence, RI, 2000.

https://doi.org/10.1017/etds.2017.130 Published online by Cambridge University Press