



HAL
open science

Sarnak's Conjecture -what's new

Sébastien Ferenczi, J Kulaga-Przymus, M Lemańczyk

► **To cite this version:**

Sébastien Ferenczi, J Kulaga-Przymus, M Lemańczyk. Sarnak's Conjecture -what's new. Springer. Ergodic Theory and Dynamical Systems in their Interactions with Arithmetics and Combinatorics CIRM Jean-Morlet Chair, Fall 2016, 2213, 2018, Lecture Notes in Mathematics, 10.1007/978-3-319-74908-2_11 . hal-03563204

HAL Id: hal-03563204

<https://hal.science/hal-03563204>

Submitted on 9 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sarnak's Conjecture – what's new

S. Ferenczi J. Kułaga-Przymus M. Lemańczyk

Abstract

An overview of last seven years results concerning Sarnak's conjecture on Möbius disjointness is presented, focusing on ergodic theory aspects of the conjecture.

Contents

| | | |
|----------|--|-----------|
| 1 | From a PNT in dynamics to Sarnak's conjecture | 6 |
| 2 | Multiplicative functions | 9 |
| 2.1 | Definition and examples | 9 |
| 2.2 | Dirichlet convolution, Euler's product | 10 |
| 2.3 | Distance between multiplicative functions | 11 |
| 2.4 | Mean of a multiplicative function. The Prime Number Theorem (PNT) | 12 |
| 2.5 | Aperiodic multiplicative functions | 13 |
| 2.6 | Davenport type estimates on short intervals | 14 |
| 2.7 | The KBSZ criterion | 15 |
| 3 | Chowla conjecture | 16 |
| 3.1 | Formulation and ergodic interpretation | 16 |
| 3.2 | The Chowla conjecture implies Sarnak's conjecture | 19 |
| 3.3 | The logarithmic versions of Chowla and Sarnak's conjectures | 21 |
| 3.4 | Frantzikinakis' theorem | 24 |
| 3.4.1 | Ergodicity of measures for which μ is quasi-generic | 24 |
| 3.4.2 | Frantzikinakis' results | 27 |
| 3.5 | Dynamical properties of Furstenberg systems associated to the Liouville and Möbius functions | 28 |
| 4 | The MOMO and AOP properties | 30 |
| 4.1 | The MOMO property and its consequences | 30 |
| 4.2 | Möbius disjointness and entropy | 33 |
| 4.3 | The AOP property and its consequences | 34 |
| 5 | Glimpses of results on Sarnak's conjecture | 35 |
| 5.1 | Systems of algebraic origin | 35 |
| 5.1.1 | Horocycle flows | 35 |
| 5.1.2 | Nilrotations, affine automorphisms | 37 |
| 5.1.3 | Other algebraic systems | 38 |
| 5.2 | Systems of measure-theoretic origin. Substitutions and interval exchange transformations | 38 |

| | | |
|----------|--|-----------|
| 5.2.1 | Systems whose powers are disjoint | 38 |
| 5.2.2 | Adic systems and Bourgain’s criterion | 38 |
| 5.2.3 | Substitutions | 41 |
| 5.2.4 | Interval exchanges | 42 |
| 5.2.5 | Systems of rank one | 43 |
| 5.2.6 | Rokhlin extensions | 45 |
| 5.3 | Distal systems | 45 |
| 5.3.1 | Discrete spectrum automorphisms | 47 |
| 5.4 | Sub-polynomial complexity | 47 |
| 5.5 | Systems of number-theoretic origin | 48 |
| 5.6 | Other research around Sarnak’s conjecture | 48 |
| 6 | Related research: \mathcal{B}-free numbers | 49 |
| 6.1 | Introduction | 49 |
| 6.2 | Invariant measures and entropy | 51 |
| 6.3 | Topological results | 52 |
| 6.4 | Ergodic Ramsey theory | 53 |

Introduction

Möbius disjointness Assume that T is a continuous map¹ of a compact metric space X . Following Peter Sarnak [148, 149], we will say that T , or, more precisely, the topological dynamical system (X, T) is *Möbius disjoint* (or *Möbius orthogonal*)² if:

$$(1) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} f(T^n x) \mu(n) = 0 \text{ for each } f \in C(X) \text{ and } x \in X.$$

In 2010, Sarnak [148, 149] formulated the following conjecture:³

- (2) Each zero entropy continuous map T of a compact metric space X is Möbius disjoint.

Note that if f is constant then convergence (1) takes place in an arbitrary topological system (X, T) ; indeed, $\frac{1}{N} \sum_{n \leq N} \mu(n) \rightarrow 0$ is equivalent to the Prime Number Theorem (PNT), e.g. [85, 159]. We can also interpret this statement as the equivalence of the PNT and the Möbius disjointness of the one-point dynamical system. The Prime Number Theorem in arithmetic progressions (Dirichlet’s theorem) can also be viewed similarly: it is equivalent to the Möbius disjointness of the system (X, T) , where $Tx = x + 1$ on $X = \mathbb{Z}/k\mathbb{Z}$ for each $k \geq 1$. Note also that the classical Davenport’s [43] estimate: for each $A > 0$, we have

$$(3) \quad \max_{t \in \mathbb{T}} \left| \sum_{n \leq N} e^{2\pi i n t} \mu(n) \right| \leq C_A \frac{N}{\log^A N} \text{ for some } C_A > 0 \text{ and all } N \geq 2,$$

¹Most often, however not always, T will be a homeomorphism.

² μ stands for the arithmetic Möbius function, see next sections for explanations of notions that appear in Introduction.

³To be compared with Möbius Randomness Law by Iwaniec and Kowalski [95], page 338, that any “reasonable” sequence of complex numbers is orthogonal to μ .

yields the Möbius disjointness of irrational rotations.⁴

The present article is concentrated on an overview of research done during the last seven years⁵ on Sarnak’s conjecture (2) from the ergodic theory point of view. It is also rather aimed at the readers with a good orientation in dynamics, especially in ergodic theory. It means that we assume that the reader is familiar with at least basics of ergodic theory, but often more than that is required, monographs [38, 55, 73, 75, 164] are among best sources to be consulted. In contrast to that, we included in the article a selection of some basics of analytic number theory. Those which appear here, in principle, are not contained in [144] and, as we hope, allow one for a better understanding of dynamical aspects of some number-theoretic results. We should however warn the reader that some number-theoretic results will be presented in their simplified (typically, non-quantitative) forms, sufficient for some ergodic interpretations but not putting across the whole complexity and depth of the results. In particular, this remark applies to recent break-through results of Matomäki and Radziwiłł [124] and some related concerning a behavior of multiplicative functions on short intervals.⁶

Ergodic theory viewpoint on Sarnak’s conjecture Sarnak’s conjecture (2) is formulated as a problem in topological dynamics. However, for each topological system (X, T) the set $M(X, T)$ of (Borel, probability) T -invariant measures is non-empty and we can study dynamical properties of (X, T) by looking at all measure-theoretic dynamical systems (X, \mathcal{B}, μ, T) for $\mu \in M(X, T)$. Via the Variational Principle, Sarnak’s conjecture can be now formulated as Möbius disjointness of the topological systems (X, T) whose measure-theoretic systems (X, \mathcal{B}, μ, T) for all $\mu \in M(X, T)$ have zero Kolmogorov-Sinai entropy. But one of main motivations for (2) in [148] was that this condition is weaker than a certain (open since 1965) pure number-theoretic result, known as the Chowla conjecture (see Section 3.1). Since the Chowla conjecture has its pure ergodic theory interpretation (Section 3.1), the approach through invariant measures allows one to see the implication

$$\text{Chowla conjecture} \Rightarrow \text{Sarnak’s conjecture}^7$$

as a consequence of some disjointness (in the sense of Furstenberg) results in ergodic theory. While the Chowla conjecture remains open, some recent break-through results in number theory find their natural interpretation as particular instances of the validity of Sarnak’s conjecture. Samples of such results are (see Sections 3.4.1 and 3.5):

⁴In order to establish Möbius disjointness, we need to show convergence (1) (for all $x \in X$) only for a set of functions linearly dense in $C(X)$, so, for the rotations on the (additive) circle $\mathbb{T} = [0, 1)$, we only need to consider characters. Note also that if the topological system (X, T) is uniquely ergodic then we need to check (1) (for all $x \in X$) only for a subset of $C(X)$ which is linearly dense in L^1 .

In what follows, for inequalities (as (3)), we will also use notation \ll or $O(\cdot)$, or \ll_A or $O_A(\cdot)$ if we need to emphasize a role of $A > 0$.

⁵For a presentation of a part of it, see [145].

⁶For a detailed account of these results, we refer the reader to [152].

⁷As proved by Tao [155], the logarithmic averages version of the Chowla conjecture is equivalent to the logarithmic version of Sarnak’s conjecture. We will see later in Section 3 that once the logarithmic Chowla conjecture holds for the Liouville function λ , we have that all configurations of ± 1 s appear in λ (infinitely often).

1. The result of Matomäki, Radziwiłł and Tao [125]:

$$\sum_{h \leq H} \left| \sum_{m \leq M} \mu(m) \mu(m+h) \right| = o(HM)$$

(when $H, M \rightarrow \infty$, $H \ll M$) implies that each system (X, T) for which all invariant measures yield measure-theoretic systems with discrete spectrum is Möbius disjoint.⁸

2. The result of Tao [154]:

$$\sum_{n \leq N} \frac{\mu(n) \mu(n+h)}{n} = o(\log N)$$

(when $N \rightarrow \infty$) for each $h \neq 0$ implies that each system (X, T) for which all invariant measures yield measure-theoretic systems with singular spectrum are logarithmically Möbius disjoint.

This is done by:

- interpreting the number theoretic results as ergodic properties of the dynamical systems given by the invariant measures of the subshift X_μ for which μ is quasi-generic,
- using classical disjointness results in ergodic theory.

It is surprising and important that the ergodic theoretical methods of the last decades that led to new non-conventional ergodic theorems and showed a particular role of nil-systems, also appear in the context of Sarnak’s conjecture, and again the role of nil-systems seems to be decisive. Together with some new disjointness results in ergodic theory, it pushes forward significantly our understanding of Möbius disjointness, at least on the level of logarithmic version of Sarnak’s conjecture. The most spectacular achievement here is the recent result of Frantzikinakis and Host [68] (see Section 3.5) who proved that each zero entropy topological system (X, T) with only countably many ergodic measures is logarithmically Möbius disjoint.

The proofs reflect the “local” nature of all the aforementioned results. In other words, regardless the total entropy of the system, to obtain (1) for a FIXED $x \in X$ (and all $f \in C(X)$), we only need to look at ergodic properties of the dynamical systems given by measures “produced” by x itself (the limit points of the empiric measures given by x). So, if all such measures yield zero entropy systems, the Chowla conjecture implies (1) (for the fixed x and all $f \in C(X)$). When all such measures yield systems with discrete spectrum / singular spectrum / countably many ergodic components then the relevant Möbius disjointness holds (at x). Points with one of the listed properties may appear in (X, T) having positive entropy. In fact, a positive entropy system can be Möbius disjoint [52]. To distinguish between zero and positive entropy systems it is natural to expect that in the zero entropy case the behavior of sums in (1) is homogenous in x (for a fixed $f \in C(X)$). Indeed, the uniform

⁸The same argument applied to the Liouville function λ implies that the subshift X_λ generated by λ is uncountable, see Section 3.

convergence (in $x \in X$, under the Chowla conjecture) of sums (1) has been proved in [4] (see Section 4); in fact (2) is equivalent to Sarnak’s conjecture in its uniform form and also in a uniform short interval form. Moreover, for the Liouville function, no positive entropy system satisfies (1) in its uniform short interval form. The problem of uniform convergence turns out to be closely related to the general problem whether Möbius disjointness is stable under our ergodic theory approach. More precisely, suppose that the topological dynamical systems (X, T) and (X', T') are such that the dynamical systems obtained from invariant measures are the same for each of them (up to measure-theoretic isomorphism). Does the Möbius disjointness of (X, T) imply the Möbius disjointness of (X', T') ? Although the answer in general seems unknown, in case of uniquely ergodic models of the same measure-theoretic system a satisfactory (positive) answer can be given [4].

Content of the article We include the following topics:

- Sarnak’s conjecture a.e., Sarnak’s conjecture versus Prime Number Theorem in dynamics – see Introduction and Section 1.
- Brief introduction to multiplicative functions, Prime Number Theorem, Kátai-Bourgain-Sarnak-Ziegler criterion – see Section 2.
- Results of Matomäki, Radziwiłł and Matomäki, Radziwiłł, Tao on multiplicative functions and some of their ergodic interpretations – see Section 3.
- Chowla conjecture, logarithmic Chowla and logarithmic Sarnak conjectures (Tao’s results and Frantzikinakis and Host’s results) – see Section 3.
- Frantzikinakis’ theorem on some consequences of ergodicity of measures for which μ is quasi-generic – see Section 3.
- Ergodic criterion for Sarnak’s conjecture – the AOP and MOMO properties (uniform convergence in (1)), Sarnak’s conjecture in topological models – see Section 4.
- Glimpses of results on Sarnak’s conjecture: systems of algebraic origin (horocycle flows, nilflows); systems of measure-theoretic origin (finite rank systems, distal systems), interval exchange transformations, systems of number-theoretic origin (automatic sequences and related) – see Section 5.
- Related research: \mathcal{B} -free systems, applications to ergodic Ramsey theory – see Section 6.

Sarnak’s conjecture a.e. Before we really get into the subject of Sarnak’s conjecture, let us emphasize that this is the requirement “for each $f \in C(X)$ and $x \in X$ ” in (1) that makes Sarnak’s conjecture deep and difficult to establish. As it has been already noticed in [148], the a.e. version of (2) is always true regardless of the entropy assumption:

Proposition 0.1 ([148]). *Let T be an automorphism of a standard Borel probability space (X, \mathcal{B}, μ) and let $f \in L^1(X, \mathcal{B}, \mu)$. Then, for a.e. $x \in X$, we have*

$$\frac{1}{N} \sum_{n \leq N} f(T^n x) \mu(n) \xrightarrow{N \rightarrow \infty} 0.$$

For a complete proof, see [3]. The main ingredient is the Spectral Theorem which replaces $\left\| \frac{1}{N} \sum_{n \leq N} f(T^n x) \boldsymbol{\mu}(n) \right\|_2$ by $\left\| \frac{1}{N} \sum_{n \leq N} z^n \boldsymbol{\mu}(n) \right\|_{L^2(\sigma_f)}$ ⁹, together with Davenport’s estimate (3) (for $A = 2$) which yields

$$\left\| \frac{1}{N} \sum_{n \leq N} f(T^n x) \boldsymbol{\mu}(n) \right\|_2 \ll \frac{1}{\log^2 N}, \quad N \geq 2.$$

The latter shows that, for $\rho > 1$, the function $\sum_{k \geq 1} \left| \frac{1}{\rho^k} \sum_{n \leq \rho^k} f(T^n \cdot) \boldsymbol{\mu}(n) \right|$ is in $L^2(X, \mu)$ which, letting $\rho \rightarrow 1$ allows one to conclude for $f \in L^\infty(X, \mu)$. The general case $f \in L^1(X, \mu)$ follows from the pointwise ergodic theorem.

As shown in [56], a use of Davenport’s type estimate proved in [80] for the nil-case, yields a polynomial version of Proposition 0.1. See also [39] for the pointwise ergodic theorem for other arithmetic weights.

1 From a PNT in dynamics to Sarnak’s conjecture

The content of this section can be viewed as a kind of motivation for Sarnak’s conjecture (and is written on the base of Tao’s post [156] and Sarnak’s lecture given at CIRM [147]).

We denote by $\mathbb{N} := \{1, 2, \dots\}$ the set of positive integers. Given $N \in \mathbb{N}$, we let $\pi(N) := \{p \leq N : p \in \mathbb{P}\}$. The classical Prime Number Theorem states that

$$(4) \quad \lim_{N \rightarrow \infty} \frac{\pi(N)}{N/\log N} = 1.$$

We will always refer to this theorem as the (classical) PNT.

Assume that T is a continuous map of a compact metric space X . Assume moreover that (X, T) is uniquely ergodic, that is, the set $M(X, T)$ of T -invariant probability Borel measures is reduced to one measure, say μ . By unique ergodicity, the ergodic averages go to zero (even uniformly) for zero mean continuous functions:

$$\frac{1}{N} \sum_{n \leq N} f(T^n x) \xrightarrow{N \rightarrow \infty} 0$$

for each $f \in C(X)$, $\int_X f d\mu = 0$, and $x \in X$. Hence, the statement that a PNT holds in (X, T) “should” mean

$$(5) \quad \lim_{N \rightarrow \infty} \frac{1}{\pi(N)} \sum_{\mathbb{P} \ni p \leq N} f(T^p x) = 0$$

for all zero mean $f \in C(X)$ and $x \in X$ (in what follows, instead of $\sum_{\mathbb{P} \ni p}$, we write simply \sum_p if no confusion arises).¹⁰ Let us see how to arrive at (5) differently.

⁹ σ_f stands for the spectral measure of f .

¹⁰ We recall that Bourgain in [26, 27, 28], proved that for each $\alpha \geq (1 + \sqrt{3})/2$, each automorphism T of a probability standard Borel space (X, \mathcal{B}, μ) and each $f \in L^\alpha(X, \mathcal{B}, \mu)$ the sums in (5) converge for a.e. $x \in X$. The result has been extended by Wierdl in [169] for all $\alpha > 1$.

Recall that the von Mangoldt function Λ is defined by $\Lambda(n) = \log p$ if $n = p^k$ for a prime number p (and $k \geq 1$) and $\Lambda(n) = 0$ otherwise. Contrary to most of arithmetic functions considered in this article, Λ is not multiplicative. It is not bounded either and its support has zero density. The (classical) PNT is equivalent to

$$\frac{1}{N} \sum_{n \leq N} \Lambda(n) \xrightarrow{N \rightarrow \infty} 1.$$

A given sequence $(a_n) \subset \mathbb{C}$ can be said to *satisfy \underline{a} PNT* whenever we can give an asymptotic estimate on

$$\sum_{n \leq N} a_n \Lambda(n)$$

when N tends to infinity; thus the classical PNT is a PNT for the sequence $a_n = 1$. In particular, a sequence (a_n) also satisfies a PNT if

$$(6) \quad \sum_{n \leq N} a_n \Lambda(n) = \sum_{n \leq N} a_n + o(N),$$

and, if additionally (a_n) has zero mean, i.e. if $\frac{1}{N} \sum_{n \leq N} a_n \xrightarrow{N \rightarrow \infty} 0$, then (a_n) satisfies a PNT if

$$(7) \quad \frac{1}{N} \sum_{n \leq N} a_n \Lambda(n) \xrightarrow{N \rightarrow \infty} 0.$$

An interesting special case is $a_n = (-1)^n$, which has zero mean. Here, we do have estimates of the sums of $\Lambda(n)$ over the odd numbers smaller than N , but they are of the order of N , thus (7) is not satisfied. Beyond this point, we will not study such particular cases and we shall always write that *the sequence (a_n) satisfies \underline{a} PNT* whenever (6) holds.

Zero mean sequences are easily “produced” in uniquely ergodic systems. We will say that a *uniquely ergodic topological dynamical system (X, T) satisfies \underline{a} PNT* if

$$(8) \quad \frac{1}{N} \sum_{n \leq N} f(T^n x) \Lambda(n) \xrightarrow{N \rightarrow \infty} 0$$

for all zero mean $f \in C(X)$ and $x \in X$. We have

$$\frac{1}{N} \sum_{n \leq N} f(T^n x) \Lambda(n) = \frac{1}{N} \sum_{p \leq N} f(T^p x) \log p + \frac{1}{N} \sum_{p^k \leq N, k \geq 2} f(T^{p^k} x) \log p.$$

Now, in the second sum if $p^k \leq N$ then $p \in [1, \sqrt{N}]$; the largest value of $\log p$ is bounded by $\frac{1}{2} \log N$, therefore, the second sum is of order $O(\sqrt{N} \cdot \log N/N)$, hence of order $N^{-\frac{1}{2} + \varepsilon}$ for each $\varepsilon > 0$. Thus, a PNT in (X, T) means that

$$(9) \quad \frac{1}{N} \sum_{p \leq N} f(T^p x) \log p \xrightarrow{N \rightarrow \infty} 0$$

for all zero mean $f \in C(X)$ and $x \in X$. Note that by the classical PNT to prove (9), we need to show it for a linearly dense set of functions.¹¹

Let us now write

$$\frac{1}{N} \sum_{p \leq N} f(T^p x) \log p = \frac{1}{N} \sum_{p \leq N/\log N} f(T^p x) \log p + \frac{1}{N} \sum_{N/\log N \leq p \leq N} f(T^p x) \log p.$$

We have $\frac{1}{N} \sum_{p \leq N/\log N} f(T^p x) \log p = O(1/\log N)$ (by $\frac{1}{M} \sum_{p \leq M} \log p \rightarrow 1$ when $M \rightarrow \infty$). Moreover, write $f = f_+ - f_-$ and then we have

$$\begin{aligned} & \frac{\log N - \log \log N}{N} \sum_{N/\log N \leq p \leq N} f_+(T^p x) \\ & \leq \frac{1}{N} \sum_{N/\log N \leq p \leq N} f_+(T^p x) \log p \leq \frac{\log N}{N} \sum_{N/\log N \leq p \leq N} f_+(T^p x) \end{aligned}$$

as $\log N - \log \log N \leq \log p \leq \log N$ for the p in the considered interval. Now, $\pi(N)/(N/(\log N - \log \log N)) \xrightarrow{N \rightarrow \infty} 1$ and $\pi(N)/(N/\log N) \xrightarrow{N \rightarrow \infty} 1$, whence

$$\left| \frac{1}{N} \sum_{p \leq N} f_+(T^p x) \log p - \frac{1}{\pi(N)} \sum_{p \leq N} f_+(T^p x) \right| \xrightarrow{N \rightarrow \infty} 0.$$

Repeating the same reasoning with f_+ replaced by f_- and by (9), we obtain that the statement *a PNT holds in (X, T)* is equivalent to (5) for all zero mean $f \in C(X)$ and $x \in X$.

Remark 1.1. By replacing $\mathbf{\Lambda}$ in (8) by $\boldsymbol{\mu}$, we come back to Sarnak's conjecture. The identity $\mathbf{\Lambda} = \boldsymbol{\mu} * \log$ (see (10) below), i.e. $\mathbf{\Lambda}(n) = \sum_{d|n} \boldsymbol{\mu}(d) \log(n/d) = -\sum_{d|n} \boldsymbol{\mu}(d) \log d$ suggests some other connections between the simultaneous validity of a PNT and Möbius disjointness in (X, T) but no rigorous theorem toward a formal equivalence of the two conditions has been proved. Actually, such an equivalence taken literally does not hold. Indeed, the fact that the support of $\mathbf{\Lambda}$ is of zero upper Banach density makes a PNT vulnerable under zero density replacements of the observable ($f(T^n x)$). On the other hand, Möbius orthogonality is stable under such replacements. We illustrate this using the following simple example.

Consider the classical case $a_n = 1$ for all $n \in \mathbb{N}$. This is the same as to consider a PNT in a uniquely ergodic model¹² of the one-point system. One can now ask if we have a PNT in all uniquely ergodic models of the one-point system

¹¹Indeed, we have

$$\begin{aligned} & \left| \frac{1}{N} \sum_{p \leq N} f(T^p x) \log p - \frac{1}{N} \sum_{p \leq N} g(T^p x) \log p \right| \\ & \leq \frac{1}{N} \sum_{p \leq N} |f(T^p x) - g(T^p x)| \log p \leq \|f - g\| \frac{1}{N} \sum_{p \leq N} \log p = O(\|f - g\|), \end{aligned}$$

as condition $\frac{1}{N} \sum_{n \leq N} \mathbf{\Lambda}(n) \xrightarrow{N \rightarrow \infty} 1$ is equivalent to $\frac{1}{N} \sum_{p \leq N} \log p \xrightarrow{N \rightarrow \infty} 1$.

¹²We recall that if $(Z, \mathcal{D}, \kappa, R)$ is a measure-preserving system then by its *uniquely ergodic model* we mean a uniquely ergodic system (X, T) with the unique (Borel) T -invariant measure μ such that $(Z, \mathcal{D}, \kappa, R)$ is measure-theoretically isomorphic to $(X, \mathcal{B}(X), \mu, T)$.

(it is an exercise to prove that all such models are Möbius disjoint). Take any sequence $(c_{p^k})_{p^k} \in \{-1, 1\}^{\mathbb{N}}$ and define b_n as a_n when $n \neq p^k$ and $b_{p^k} = c_{p^k}$. We can see that

$$\frac{1}{N} \sum_{n \leq N} b_n \Lambda(n) = \frac{1}{N} \sum_{p^k \leq N} c_{p^k} \log p.$$

Now, the subshift $X_b \subset \{-1, 1\}^{\mathbb{N}}$ generated by b (cf. (27)) has only one invariant measure $\delta_{11\dots}$, so it is a uniquely ergodic model of the one-point system and if we take $f(z) = 1 - z(1)$ ($z \in X_b$) as our continuous function, we can see that f has zero mean but neither (8) nor (5) are satisfied if the sequence c is badly behaving. It follows that we can expect a PNT to hold only in some classes of “natural” dynamical systems, samples of which we will see in Section 5.

Returning to our discussion on a PNT, in any such situation, given a bounded sequence $(f(n)) \subset \mathbb{C}$, we can write

$$\sum_{n \leq N} f(n) \Lambda(n) = - \sum_{n \leq N} f(n) \sum_{d|n} \mu(d) \log(d) = - \sum_{d \leq N} \mu(d) \log d \sum_{e \leq N/d} f(ed).$$

Then a further decomposition of the second sum into a structured part and a remainder leads to two sums and allows one for an application of Möbius Randomness Law to the second sum in order to predict the correct main term value of $\sum_{n \leq N} f(n) \Lambda(n)$, see [147].

2 Multiplicative functions

2.1 Definition and examples

An arithmetic function $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$ is called *multiplicative* if $\mathbf{u}(1) = 1$ and $\mathbf{u}(mn) = \mathbf{u}(m)\mathbf{u}(n)$ whenever $(m, n) = 1$. If $\mathbf{u}(mn) = \mathbf{u}(m)\mathbf{u}(n)$ without the coprimeness restriction on m, n , then \mathbf{u} is called *completely multiplicative*. Clearly, each multiplicative function is entirely determined by its values at p^α , where $p \in \mathbb{P}$ is a prime number and $\alpha \in \mathbb{N}$ (for completely multiplicative functions $\alpha = 1$). A prominent example of a multiplicative function is the Möbius function μ determined by $\mu(p) = -1$ and $\mu(p^\alpha) = 0$ for $\alpha \geq 2$. Note that μ^2 (which is obviously also multiplicative) is the characteristic function of the set of square-free numbers. The Liouville function $\lambda: \mathbb{N} \rightarrow \mathbb{C}$ is completely multiplicative and is given by $\lambda(p) = -1$. Clearly, $\mu = \lambda \cdot \mu^2$ and we will see soon some more relations between μ and λ . Many other classical arithmetic functions are multiplicative, for example: the Euler function ϕ ; the function $n \mapsto (-1)^{n+1}$ is a periodic multiplicative function which is not completely multiplicative; $\mathbf{d}(n) :=$ number of divisors of n , $n \mapsto 2^{\omega(n)}$, where $\omega(n)$ stands for the number of different prime divisors of n ; $\sigma(n) = \sum_{d|n} d$. Recall that given $q \geq 1$, a function $\chi: \mathbb{N} \rightarrow \mathbb{C}$ is called a *Dirichlet character of modulus q* if:

- (i) χ is q -periodic and completely multiplicative,
- (ii) $\chi(n) \neq 0$ if and only if $(n, q) = 1$.

It is not hard to see that Dirichlet characters are determined by the ordinary characters of the multiplicative group (of order $\phi(q)$) $(\mathbb{Z}/q\mathbb{Z})^*$ of invertible (under multiplication) elements in $\mathbb{Z}/q\mathbb{Z}$. The Dirichlet character $\chi_1(n) := 1$ iff

$(n, q) = 1$ is called the *principal character* of modulus q . Moreover, each periodic, completely multiplicative function is a Dirichlet character (of a certain modulus). Another class of important (completely) multiplicative functions is given by Archimedean characters $n \mapsto n^{it} = e^{it \log n}$ which are indexed by $t \in \mathbb{R}$.

2.2 Dirichlet convolution, Euler's product

Recall that given two arithmetic functions $\mathbf{u}, \mathbf{v}: \mathbb{N} \rightarrow \mathbb{C}$, by their *Dirichlet convolution* $\mathbf{u} * \mathbf{v}$ we mean the arithmetic function

$$(10) \quad \mathbf{u} * \mathbf{v}(n) := \sum_{d|n} \mathbf{u}(d)\mathbf{v}(n/d), \quad n \in \mathbb{N}.$$

If by A we denote the set of arithmetic functions then $(A, +, *)$ is a ring which is an integral domain and the unit $\mathbf{e} \in A$ is given by $\mathbb{1}_{\{1\}}$.¹³ There is a natural ring isomorphism between A and the ring D of (formal)¹⁴ Dirichlet series

$$A \ni \mathbf{u} \mapsto U(s) := \sum_{n=1}^{\infty} \frac{\mathbf{u}(n)}{n^s} \in D, \quad s \in \mathbb{C},$$

under which

$$U(s)V(s) = \sum_{n=1}^{\infty} \frac{\mathbf{u} * \mathbf{v}(n)}{n^s}.$$

When $\mathbf{u} = \mathbb{1}_{\mathbb{N}}$ then the Dirichlet series defines the Riemann ζ function:¹⁵

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \text{ for } \operatorname{Re} s > 1.$$

It is classical that if \mathbf{u} and \mathbf{v} are multiplicative then so is their Dirichlet convolution. The importance of multiplicativity can be seen in the representation of the Dirichlet series of a multiplicative function \mathbf{u} as an Euler's product. Indeed, a general term of $\prod_{p \in \mathbb{P}} (1 + \mathbf{u}(p)p^{-s} + \mathbf{u}(p^2)p^{-2s} + \dots)$ has the form $\frac{\mathbf{u}(p_1^{\alpha_1}) \dots \mathbf{u}(p_{i_r}^{\alpha_{i_r}})}{(p_1^{\alpha_1} \dots p_{i_r}^{\alpha_{i_r}})^s} = \frac{\mathbf{u}(p_1^{\alpha_1} \dots p_{i_r}^{\alpha_{i_r}})}{(p_1^{\alpha_1} \dots p_{i_r}^{\alpha_{i_r}})^s}$, i.e. equals $\frac{\mathbf{u}(n)}{n^s}$ for some n . It easily follows that

$$\sum_{n \geq 1} \frac{\mathbf{u}(n)}{n^s} = \prod_{p \in \mathbb{P}} (1 + \mathbf{u}(p)p^{-s} + \mathbf{u}(p^2)p^{-2s} + \dots).$$

If additionally \mathbf{u} is completely multiplicative (and $|\mathbf{u}| \leq 1$), then $\mathbf{u}(p^k) = \mathbf{u}(p)^k$ and

$$\sum_{n=1}^{\infty} \frac{\mathbf{u}(n)}{n^s} = \prod_{p \in \mathbb{P}} (1 - \mathbf{u}(p)p^{-s})^{-1}.$$

¹³The Möbius Inversion Formula is given by $\mu * \mathbb{1}_{\mathbb{N}} = \mathbf{e}$.

¹⁴We will not discuss here the problem of convergence of Dirichlet series, see [144].

¹⁵An analytic continuation of ζ yields a meromorphic function on \mathbb{C} (with one pole at $s = 1$) satisfying the functional equation

$$(11) \quad \zeta(s) = 2^s \pi^{s-1} \sin\left(\frac{\pi s}{2}\right) \Gamma(1-s) \zeta(1-s).$$

Because of the sine, $\zeta(-2k) = 0$ for all integers $k \geq 1$ – these are so called trivial zeros of ζ ($\zeta(2k) \neq 0$ since Γ has simple poles at $0, -1, -2, \dots$). In $\operatorname{Re} s > 1$ there are no zeros of ζ (ζ is represented by a convergent infinite product), so except of $-2k, k \geq 1$, there are no zeros for $s \in \mathbb{C}, \operatorname{Re} s < 0$ (as $\operatorname{Re}(1-s) > 1$). The Riemann Hypothesis asserts that all nontrivial zeros of ζ are on the line $x = \frac{1}{2}$. See [144].

Note that if $\mathbf{u} = \boldsymbol{\mu}$, we obtain

$$\sum_{n \geq 1} \frac{\boldsymbol{\mu}(n)}{n^s} = \prod_{p \in \mathbb{P}} (1 - p^{-s})$$

since $\boldsymbol{\mu}(p) = -1$ and $\boldsymbol{\mu}(p^r) = 0$ whenever $r \geq 2$. Since for the Riemann ζ function, we have $\zeta(s) = \prod_{p \in \mathbb{P}} (1 - p^{-s})^{-1}$ for $\operatorname{Re} s > 1$, we obtain the following.

Corollary 2.1. We have $\frac{1}{\zeta(s)} = \sum_{n \geq 1} \frac{\boldsymbol{\mu}(n)}{n^s}$ whenever $\operatorname{Re} s > 1$.

We could have derived the above assertion in a different way. Indeed, $\boldsymbol{\mu} * \mathbb{1}_{\mathbb{N}} = \mathbf{e}$. If $G(s) := \sum_{n=1}^{\infty} \frac{\boldsymbol{\mu}(n)}{n^s}$ stands for the Dirichlet series of the Möbius function, then

$$G(s) \cdot \zeta(s) = \sum_{n=1}^{\infty} \frac{(\boldsymbol{\mu} * \mathbb{1}_{\mathbb{N}})(n)}{n^s} = \sum_{n=1}^{\infty} \frac{\mathbf{e}(n)}{n^s} = 1.$$

2.3 Distance between multiplicative functions

Denote by

$$(12) \quad \mathcal{M} := \{\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C} : \mathbf{u} \text{ is multiplicative and } |\mathbf{u}| \leq 1\}.$$

Let $\mathbf{u}, \mathbf{v} \in \mathcal{M}$. Define the “distance” function D on \mathcal{M} by setting

$$(13) \quad D(\mathbf{u}, \mathbf{v}) := \left(\sum_{p \in \mathbb{P}} \frac{1}{p} \left(1 - \operatorname{Re} \left(\mathbf{u}(p) \overline{\mathbf{v}(p)} \right) \right) \right)^{1/2}.$$

For each $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{M}$, we have:

- $D(\mathbf{u}, \mathbf{u}) \geq 0$; $D(\mathbf{u}, \mathbf{u}) = 0$ iff $\sum_{p \in \mathbb{P}} \frac{1}{p} (1 - |\mathbf{u}(p)|^2) = 0$ iff $|\mathbf{u}(p)| = 1$ for all $p \in \mathbb{P}$, so $D(n^{it}, n^{it}) = 0$ for each $t \in \mathbb{R}$, $D(\boldsymbol{\lambda}, \boldsymbol{\lambda}) = D(\boldsymbol{\mu}, \boldsymbol{\mu}) = 0$. Of course, if $\mathbf{u}(p) = 0$ for each $p \in \mathbb{P}$ then $D(\mathbf{u}, \mathbf{u}) = +\infty$. Moreover, $\boldsymbol{\phi}(n)/n \in \mathcal{M}$ and $D(\boldsymbol{\phi}(n)/n, \boldsymbol{\phi}(n)/n) = \sum_{p \in \mathbb{P}} \frac{1}{p} (1 - \frac{(1-p)^2}{p^2})$ is positive and finite.
- $D(\mathbf{u}, \mathbf{v}) = D(\mathbf{v}, \mathbf{u})$.
- $D(\mathbf{u}, \mathbf{v}) \leq D(\mathbf{u}, \mathbf{w}) + D(\mathbf{w}, \mathbf{v})$, see [78].

When $D(\mathbf{u}, \mathbf{v}) < +\infty$ then one says that \mathbf{u} *pretends to be* \mathbf{v} . For example, $\boldsymbol{\mu}^2$ and $\boldsymbol{\phi}(n)/n$ pretend to be $\mathbb{1}$ (as $\sum_{p \in \mathbb{P}} \frac{1}{p} (1 - \frac{p-1}{p}) = \sum_{p \in \mathbb{P}} \frac{1}{p^2} < +\infty$).

Lemma 2.2 ([78]). *For each $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{w}' \in \mathcal{M}$, we have*

$$(i) \quad D(\mathbf{uw}, \mathbf{vw}') \leq D(\mathbf{u}, \mathbf{v}) + D(\mathbf{w}, \mathbf{w}').$$

Moreover, by (i) and a simple induction,

$$(ii) \quad mD(\mathbf{u}, \mathbf{v}) \geq D(\mathbf{u}^m, \mathbf{v}^m) \text{ for all } m \in \mathbb{N}.$$

If we fix $t \neq 0$ and $k \geq k_0$ then the number of $p \in \mathbb{P}$ satisfying

$$\exp\left(\frac{2\pi}{t}\left(k + \frac{1}{3}\right)\right) \leq p \leq \exp\left(\frac{2\pi}{t}\left(k + \frac{2}{3}\right)\right)$$

is (by the PNT) at least $C \frac{\exp(2\pi k/t)}{k/t}$ (for a constant $C > 0$), whence

$$\left| \left\{ p \in \mathbb{P} : k + \frac{1}{3} \leq \frac{t \log p}{2\pi} \leq k + \frac{2}{3} \right\} \right| \geq C \frac{\exp(2\pi k/t)}{k/t}.$$

It follows that

$$(14) \quad \sum_{\exp(\frac{2\pi}{t}(k+\frac{1}{3})) \leq p \leq \exp(\frac{2\pi}{t}(k+\frac{2}{3}))} \frac{1}{p} (1 - \cos(t \log p)) \geq C' \frac{1}{k}$$

for a constant $C' > 0$. Now, using (13), (14) and summing over k , we obtain the following:¹⁶

$$(15) \quad D(\mathbb{1}, n^{it}) = \infty \text{ for each } t \neq 0.$$

It is not difficult to see that for $t \neq 0$, $D(\chi, n^{it}) = +\infty$ for each Dirichlet character χ , while for $t = 0$, we have $D(\chi, 1) < +\infty$ if and only if χ is principal.

2.4 Mean of a multiplicative function. The Prime Number Theorem (PNT)

The distance D is useful when we want to compute means of multiplicative functions. Given an arithmetic function $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$ its *mean* $M(\mathbf{u})$ is defined as $M(\mathbf{u}) := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \mathbf{u}(n)$ (if the limit exists).

Theorem 2.3 (Halász; e.g. Thm. 6.3 [57]). *Let $\mathbf{u} \in \mathcal{M}$. Then $M(\mathbf{u})$ exists and is non-zero if and only if*

- (i) *there is at least one positive integer k so that $\mathbf{u}(2^k) \neq -1$, and*
- (ii) *the series $\sum_{p \in \mathbb{P}} \frac{1}{p} (1 - \mathbf{u}(p))$ converges.*

When these conditions are satisfied, we have

$$M(\mathbf{u}) = \prod_{p \in \mathbb{P}} \left(1 - \frac{1}{p} \right) \left(1 + \sum_{m=1}^{\infty} p^{-m} \mathbf{u}(p^m) \right).$$

The mean value $M(\mathbf{u})$ exists and is zero if and only if either

- (iii) *there is a real number τ , so that for each positive integer k , $\mathbf{u}(2^k) = -2^{ki\tau}$, moreover $D(\mathbf{u}, n^{i\tau}) < +\infty$; or*
- (iv) *$D(\mathbf{u}, n^{it}) = \infty$ for each $t \in \mathbb{R}$.*

Corollary 2.4 (Wirsing's theorem). *If $\mathbf{u} \in \mathcal{M}$ is real-valued then $M(\mathbf{u})$ exists.*

Proof. Since $\operatorname{Re}(p^{it}) = \operatorname{Re}(p^{-it})$, and $\mathbf{u}(p) \in \mathbb{R}$, we have

$$D(\mathbb{1}, n^{2it}) = D(n^{-it}, n^{it}) \leq 2D(\mathbf{u}, n^{it})$$

by the triangle inequality. By (15), it follows that $D(\mathbf{u}, n^{it}) = +\infty$ for each $0 \neq t \in \mathbb{R}$. Hence, if $D(\mathbf{u}, \mathbb{1}) = +\infty$, then $D(\mathbf{u}, n^{it}) = +\infty$ for each $t \in \mathbb{R}$ and then $M(\mathbf{u}) = 0$ by Halász's theorem (iv).

If not then $D(\mathbf{u}, 1) < +\infty$. Then the series $\sum_{p \in \mathbb{P}} \frac{1}{p} (1 - \mathbf{u}(p))$ converges (so (ii) is satisfied) and we check whether or not $\mathbf{u}(2^k) = -1$ for all $k \in \mathbb{N}$, that is, either (i) holds or (iii) holds. \square

¹⁶This proof of (15) has been shown to us by G. Tenenbaum.

Remark 2.5. It follows from (15) that in Halász’s theorem (iii) and (iv) are two disjoint conditions.

Remark 2.6. Not all functions from \mathcal{M} have mean. Indeed, an Archimedean character n^{it} has mean iff $t = 0$. This can be shown by a direct computation: apply Euler’s summation formula to $f(x) = x^{it}$ with $t \neq 0$, to obtain $\frac{1}{N} \sum_{n \leq N} n^{it} = \frac{N^{it}}{it+1} + O\left(\frac{\log N}{N}\right)$.

Theorem 2.7 (e.g. [78, 85, 159]). *The PNT is equivalent to $M(\boldsymbol{\mu}) = 0$.*

Remark 2.8. The statement above is an elementary equivalence, see the discussion in Section 4 [49]. For a PNT for a more general f (i.e. not for $f = 1$) the relation between such a disjointness and sums over the primes requires more quantitative estimates than simply $o(N)$.

Remark 2.9. By Halász’s theorem, condition $M(\boldsymbol{\mu}) = 0$ is equivalent to $D(\boldsymbol{\mu}, n^{it}) = \infty$ for each $t \in \mathbb{R}$ ($\boldsymbol{\mu}$ does not pretend to be n^{it}), and this can be established similarly to the proof of (15).

The PNT tells us about cancelations of $+1$ and -1 for $\boldsymbol{\mu}$. When one requires a behavior similar to random sequences, say “square-root type cancelation”, the result is much stronger:

Theorem 2.10 (Littlewood, see [36]). *The Riemann Hypothesis holds if and only if for every $\varepsilon > 0$, we have $\sum_{n \leq N} \boldsymbol{\mu}(n) = O_\varepsilon(N^{\frac{1}{2} + \varepsilon})$.*

This result is not hard to establish and we show the sufficiency: By Corollary 2.1, we have

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\boldsymbol{\mu}(n)}{n^s} = - \sum_{n=1}^{\infty} \boldsymbol{\mu}(n) \int_n^{\infty} dx^{-s} = s \sum_{n=1}^{\infty} \boldsymbol{\mu}(n) \int_n^{\infty} \frac{dx}{x^{s+1}}.$$

Setting $M(x) = \sum_{n \leq x} \boldsymbol{\mu}(n)$, we obtain

$$(16) \quad \frac{1}{\zeta(s)} = s \int_1^{\infty} \frac{M(x)}{x^{s+1}} dx, \quad \operatorname{Re} s > 1$$

and, by the assumption on $M(\cdot)$,

$$\int_1^{\infty} \left| \frac{M(x)}{x^{s+1}} \right| dx = \int_1^{\infty} \frac{|M(x)|}{x^{\operatorname{Re} s + 1}} dx \ll \int_1^{\infty} x^{\frac{1}{2} + \varepsilon - (\operatorname{Re} s + 1)} dx = \int_1^{\infty} x^{-\operatorname{Re} s - \frac{1}{2} + \varepsilon} dx.$$

It follows that the integral on the RHS of (16) is absolutely convergent for $\operatorname{Re} s > \frac{1}{2} + \varepsilon$. Hence, (16) yields an analytic extension of $\frac{1}{\zeta(\cdot)}$ to $\{s \in \mathbb{C} : \operatorname{Re} s > \frac{1}{2} + \varepsilon\}$. In this domain there are no zeros of ζ and by the functional equation (see (11)) on ζ , we obtain the Riemann Hypothesis.

2.5 Aperiodic multiplicative functions

Denote by

$$\mathcal{M}_{\text{conv}} := \{\boldsymbol{u} \in \mathcal{M} : \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \boldsymbol{u}(an + r) \text{ exists for all } a, r \in \mathbb{N}\}.$$

The following is classical.

Lemma 2.11. *Let $\mathbf{u} \in \mathcal{M}$. Then $\mathbf{u} \in \mathcal{M}_{\text{conv}}$ if and only if the mean value $M(\chi \cdot \mathbf{u})$ exists for each Dirichlet character χ .*

An arithmetic function $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$ is called *aperiodic* if, for all $a, r \in \mathbb{N}$, we have $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \mathbf{u}(an + r) = 0$. Similarly to Lemma 2.11, we obtain that $\mathbf{u} \in \mathcal{M}$ is aperiodic if and only if $M(\chi \cdot \mathbf{u}) = 0$ for each Dirichlet character χ . Delange theorem (see, e.g., [78]) gives necessary and sufficient conditions for \mathbf{u} to be aperiodic. In particular, each $\mathbf{u} \in \mathcal{M}$ satisfying $D(\mathbf{u}, \chi \cdot n^{it}) = 0$ for all Dirichlet characters χ and all $t \in \mathbb{R}$, is aperiodic. Classical multiplicative functions as $\boldsymbol{\mu}$ or $\boldsymbol{\lambda}$ are aperiodic.

Frantzikinakis and Host in [67] prove a deep structure theorem for multiplicative functions from \mathcal{M} . One of the consequences of it is the following characterization of aperiodic functions: $\mathbf{u} \in \mathcal{M}$ is aperiodic if and only if it is uniform, that is, all Gowers uniformity seminorms¹⁷ vanish [67]. In [23] (see Theorem 1.3 therein), this result is extended to show that $\mathbf{u} \in \mathcal{M}_{\text{conv}}$ is either uniform or rational.¹⁸ Also, a variation of this result has been proved in [23] (see Theorem A therein):

$$(17) \quad \begin{array}{l} \text{for each positive density level set } E = \{n \in \mathbb{N} : \mathbf{u}(n) = c\} \text{ of } \mathbf{u} \in \mathcal{M} \\ \text{there is a (unique if density is smaller than 1) rational (i.e. coming} \\ \text{from a rational function from } \mathcal{M}) \text{ level set } R \text{ of } \mathbf{v} \in \mathcal{M} \text{ such that} \\ d(R)\mathbb{1}_E - d(E)\mathbb{1}_R \text{ is Gowers uniform.} \end{array}$$

For example, for $E = \{n \in \mathbb{N} : \boldsymbol{\mu}(n) = 1\}$ the unique set R is just the set of square-free numbers.

2.6 Davenport type estimates on short intervals

Given $\mathbf{u} \in \mathcal{M}$, for our purposes we will need additionally the following:¹⁹ for each $(b_n) \subset \mathbb{N}$ with $b_{n+1} - b_n \rightarrow \infty$ and any $c \in \mathbb{C}$, $|c| = 1$, we have

$$(18) \quad \frac{1}{b_{K+1}} \sum_{k \leq K} \left| \sum_{b_k \leq n < b_{k+1}} c^n \mathbf{u}(n) \right| \xrightarrow{K \rightarrow \infty} 0.$$

¹⁷For $N \in \mathbb{N}$ we write $[N]$ for the set $\{1, 2, \dots, N\}$. Given $h, N \in \mathbb{N}$ and $f: \mathbb{N} \rightarrow \mathbb{C}$, we let $S_h f(n) = f(n+h)$ and $f_N = \mathbb{1}_{[N]} \cdot f$. For $s \in \mathbb{N}$, the *Gowers uniformity seminorm* [77] $\|\cdot\|_{U_{[N]}^s}$ is defined in the following way:

$$\|f\|_{U_{[N]}^1} := \left| \frac{1}{N} \sum_{n=1}^N f_N(n) \right|$$

and for $s \geq 1$

$$\|f\|_{U_{[N]}^{2^{s+1}}} := \frac{1}{N} \sum_{h=1}^N \left\| f_N S_h \overline{f_N} \right\|_{U_{[N]}^s}^{2^s}.$$

A bounded function $f: \mathbb{N} \rightarrow \mathbb{C}$ is called *uniform* if $\|f\|_{U_{[N]}^s}$ converges to zero as $N \rightarrow \infty$ for each $s \geq 1$.

¹⁸An arithmetic function \mathbf{u} is *rational* if for each $\varepsilon > 0$ there is a periodic function \mathbf{v} such that $\limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} |\mathbf{u}(n) - \mathbf{v}(n)| < \varepsilon$. Note that since $\boldsymbol{\mu}$ is aperiodic, whence orthogonal to all periodic sequences it will also be orthogonal to each rational \mathbf{u} [23]. An example of rational sequence is given by $\boldsymbol{\mu}^2$. For more examples, see the sets of \mathcal{B} -free numbers in the Erdős case in Section 6.

¹⁹To be compared with the estimates (3), where we drop the sup requirement.

It is not hard to see that if $\mathbf{u} \in \mathcal{M}$ satisfies (18) for each (b_n) and c as above, then it must be aperiodic.

In fact, it follows from a break-through result in [124] and [125] that the class of $\mathbf{u} \in \mathcal{M}$ for which (18) holds contains all \mathbf{u} for which

$$(19) \quad \inf_{|t| \leq M, \chi \bmod q, q \leq Q} D(\mathbf{u}, n \mapsto \chi(n)n^{it}; M)^2 \rightarrow \infty,$$

when $10 \leq H \leq M$, $H \rightarrow \infty$ and $Q = \min(\log^{1/125} M, \log^5 H)$; here χ runs over all Dirichlet characters of modulus $q \leq Q$ and

$$D(\mathbf{u}, \mathbf{v}; M) := \left(\sum_{p \leq M, p \in \mathbb{P}} \frac{1 - \operatorname{Re}(\mathbf{u}(p)\overline{\mathbf{v}(p)})}{p} \right)^{1/2}$$

for each $\mathbf{u}, \mathbf{v} \in \mathcal{M}$. Moreover, classical multiplicative functions like $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ satisfy (19), see [125].

Finally, note that (18) true for all (b_n) as above is equivalent to the following statement:

$$(20) \quad \frac{1}{M} \sum_{M \leq m < 2M} \left| \sum_{m \leq h < m+H} c^h \mathbf{u}(h) \right| \xrightarrow{M, H \rightarrow \infty, H = o(M)} 0$$

(we can also replace the first sum by $\sum_{1 \leq m < M}$), see [7] for details. This statement is much closer to the original formulations of (simplified versions of) theorems from [124, 125].

One more consequence of the main result in [124] is the following:

Theorem 2.12 (Thm. 1.1 in [125] and a corollary for $k = 2$ therein). *For $H \rightarrow \infty$ arbitrarily slowly with $M \rightarrow \infty$ ($H \leq M$), we have*

$$\sum_{h \leq H} \left| \sum_{m \leq M} \boldsymbol{\mu}(m)\boldsymbol{\mu}(m+h) \right| = o(HM).$$

2.7 The KBSZ criterion

Sarnak's conjecture is aimed at showing that deterministic sequences (i.e. those given as observable sequences in the zero entropy systems) are orthogonal to $\boldsymbol{\mu}$. In particular, as $\boldsymbol{\mu}$ is a multiplicative function, the result²⁰ below establishes disjointness with $\boldsymbol{\mu}$.

Theorem 2.13 ([31, 104]). *Assume that (a_n) is a bounded sequence of complex numbers. Assume that for all prime numbers $p \neq q$*

$$(21) \quad \frac{1}{N} \sum_{n \leq N} a_{pn} \overline{a_{qn}} \xrightarrow{N \rightarrow \infty} 0.$$

²⁰The main ideas for this result appeared in [40] and [132]. It was first established in a slightly different form in [104] and then in [31], see also [84] for a proof. The criterion has its origin in the bilinear method of Vinogradov [163] which is a technique to study sums of a over primes in terms of sums over progressions $\sum_{n \leq N} a_{dn}$ and sums $\sum_{n \leq N} a_{d_1 n} a_{d_2 n}$. If $a_n = f(T^n x)$ then these sums are Birkhoff sums for powers of T and their joinings.

In what follows we will refer to Theorem 2.13 as to the KBSZ criterion.

Then, for each multiplicative function $\mathbf{u} \in \mathcal{M}$, we have

$$(22) \quad \frac{1}{N} \sum_{n \leq N} a_n \mathbf{u}(n) \xrightarrow{N \rightarrow \infty} 0.$$

For example, see [104], the criterion applies to the sequences of the form $(e^{iP(n)})$, where $P \in \mathbb{R}[x]$ has at least one irrational coefficient (different from the constant term).

In the context of dynamical systems, we use this criterion for $a_n = f(T^n x)$, $n \geq 1$. Clearly, this leads us to study the behavior of different (prime) powers of a fixed map T . We should warn the reader that when applying Theorem 2.13, we do not expect to have (21) satisfied for all continuous functions, in fact, even in uniquely ergodic systems, in general, it cannot hold for all zero mean functions²¹ but we need a subset of $C(X)$ which is linearly dense, cf. footnote 4.

We will also need the following variation of Theorem 2.13, see [7]:

Proposition 2.14. *Assume that (a_n) is a bounded sequence of complex numbers. Assume, moreover, that*

$$(23) \quad \limsup_{\substack{p, q \rightarrow \infty \\ \text{different primes}}} \left(\limsup_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{n \leq N} a_{pn} \bar{a}_{qn} \right| \right) = 0.$$

Then, for each multiplicative function $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$, $\mathbf{u} \in \mathcal{M}$, we have

$$(24) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} a_n \cdot \mathbf{u}(n) = 0.$$

Remark 2.15. In contrast to the KBSZ criterion given by Theorem 2.13, condition (23) has its ergodic theoretical counterpart – the property called AOP (see Section 4) which is a measure-theoretic invariant.

3 Chowla conjecture

In this section we get into the subject of the Chowla conjecture which is the main motivation for Sarnak’s conjecture.

3.1 Formulation and ergodic interpretation

The Chowla conjecture deals with higher order correlations of the Möbius function,²² that is, the conjecture asserts that

$$(25) \quad \frac{1}{N} \sum_{n \leq N} \boldsymbol{\mu}^{j_0}(n) \boldsymbol{\mu}^{j_1}(n + k_1) \dots \boldsymbol{\mu}^{j_r}(n + k_r) \xrightarrow{N \rightarrow \infty} 0$$

²¹We can easily see that when $Tx = x + \alpha$ is an irrational rotation on $\mathbb{T} = [0, 1)$, then, by the Weyl criterion on uniform distribution, (21) is satisfied for all characters (for all $x \in \mathbb{T}$), but there are continuous zero mean functions for which (21) fails [110].

²²As a matter of fact, in [36], it is formulated for the Liouville function. We follow [148]. For a discussion on an equivalence of the Chowla conjecture with $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$, we invite the reader to [141]. As shown in [125], there are non-pretentious (completely) multiplicative functions for which Chowla conjecture fails. For more information, see the discussion on Elliot’s conjecture in [125].

whenever $1 \leq k_1 < \dots < k_r$, $j_s \in \{1, 2\}$ not all equal to 2, $r \geq 0$.²³

We will now explain an ergodic meaning of the Chowla conjecture. Recall that given a dynamical system (X, T) and $\mu \in M(X, T)$, a point $x \in X$ is called *generic for μ* if

$$\frac{1}{N} \sum_{n \leq N} f(T^n x) \xrightarrow{N \rightarrow \infty} \int_X f d\mu$$

for each $f \in C(X)$. Equivalently, $\frac{1}{N} \sum_{n \leq N} \delta_{T^n x} \xrightarrow{N \rightarrow \infty} \mu$ (we recall that $M(X, T)$ is endowed with the weak* topology which makes it a compact metrizable space). By compactness, each point is *quasi-generic* for a certain measure $\nu \in M(X, T)$, i.e.

$$\frac{1}{N_k} \sum_{n \leq N_k} \delta_{T^n x} \xrightarrow{k \rightarrow \infty} \nu$$

for a certain subsequence $N_k \rightarrow \infty$. Let

$$(26) \quad \text{Q-gen}(x) := \{\nu \in M(X, T) : x \text{ is quasi-generic for } \nu\}.$$
²⁴

Assume now that we have a finite alphabet A . We consider $(A^{\mathbb{Z}}, S)$, so called *full shift*, or more precisely, *two-sided full shift*, where $A^{\mathbb{Z}}$ is endowed with the product topology and $S((x_n)) = (y_n)$ with $y_n = x_{n+1}$ for each $n \in \mathbb{Z}$. Each $X \subset A^{\mathbb{Z}}$ that is closed and S -invariant yields a *subshift*, i.e. the dynamical system (X, S) . One way to obtain a subshift is to choose $x \in A^{\mathbb{Z}}$ and consider the closure X_x of the orbit of x via S . If x is given as a one-sided sequence, $x \in A^{\mathbb{N}}$, we still might consider

$$(27) \quad X_x := \{y \in A^{\mathbb{Z}} : \text{each block appearing in } y \text{ appears in } x\}$$

to obtain a two-sided subshift. In case when each block appearing in x reappears infinitely often, $X_x = \{S^n \bar{x} : n \in \mathbb{Z}\}$, for some \bar{x} for which $\bar{x}(j) = x(j)$ for each $j \geq 1$ but, in general, there is no such a good \bar{x} . Moreover, we will let ourselves speak about a one-sided sequence x to be generic or quasi-generic for a measure $\nu \in M(X_x, S)$.

Now take $A = \{-1, 0, 1\}$. For each subshift $X \subset \{-1, 0, 1\}^{\mathbb{Z}}$ let $\theta \in C(X)$ be defined as

$$(28) \quad \theta(y) = y(0), \quad y \in X.$$

Note that directly from the Stone-Weierstrass theorem we obtain the following.

Lemma 3.1. *The linear subspace generated by the constants and the family*

$$\{\theta^{j_0} \circ S^{k_0} \cdot \theta^{j_1} \circ S^{k_1} \cdot \dots \cdot \theta^{j_r} \circ S^{k_r} : k_i \in \mathbb{Z}, j_i \in \{1, 2\}, i = 0, 1, \dots, r, r \geq 0\}$$

of continuous functions is an algebra of functions separating points, hence it is dense in $C(X)$.

²³The Chowla conjecture is rather “close” in spirit to the Twin Number Conjecture in the sense that the latter is expressed by $(*) \sum_{n \leq x} \mathbf{\Lambda}(n)\mathbf{\Lambda}(n+2) = (2\Pi_2) \cdot x + o(x)$, where $\Pi_2 = \prod_p (1 - \frac{1}{(p-1)^2}) = 0,66016\dots$ which can be compared with $\sum_{n \leq x} \mu(n)\mu(n+2) = o(x)$ which is “close” to the Chowla conjecture, see e.g. [157]. A recent development shows that it is realistic to claim that the Chowla conjecture with an error term of the form $o((\log N)^{-A})$ for some A large enough (A depending on the number of shifts of μ that are considered) implies $(*)$. (Of course, everywhere $\mathbf{\Lambda}$ is a good approximation of $\mathbf{1}_{\mathbb{P}}$).

See also [140] for a (conditional) equivalence of $(*)$ with $\sum_{n \leq N} \mathbf{\Lambda}(n)\mu(n+2) = o(N)$.

²⁴We recall that either x is generic or $\text{Q-gen}(x)$ is a connected uncountable set, see Proposition 3.8 in [47].

The subshift (X_μ, S) is called the *Möbius system* and $X_{\mu^2} \subset \{0, 1\}^{\mathbb{Z}} \subset \{-1, 0, 1\}^{\mathbb{Z}}$ is the *square-free system*.²⁵ Note that $s: (z(n)) \mapsto (z(n)^2)$ will settle a factor map between the Möbius system and the square-free system. The point μ^2 is a generic point for so called *Mirsky measure* ν_{μ^2} [33, 131] (see Section 6.2). In other words, there are frequencies of blocks on μ^2 : for each block $B \in \{0, 1\}^\ell$, the following limit exists:

$$\lim_{N \rightarrow \infty} \frac{1}{N} |\{1 \leq n \leq N - \ell : \mu^2(n, n + \ell - 1) = B\}| =: \nu_{\mu^2}(B).$$

We can now consider the relatively independent extension²⁶ $\widehat{\nu}_{\mu^2}$ of ν_{μ^2} which is the measure on $s^{-1}(X_{\mu^2}) \subset \{-1, 0, 1\}^{\mathbb{Z}}$ given by the following condition: for each block $C \in \{-1, 0, 1\}^\ell$, we have

$$\widehat{\nu}_{\mu^2}(C) := \frac{1}{2^k} \nu_{\mu^2}(C^2),$$

where C^2 is obtained from B by squaring on each coordinate and k is the number of 1 in C^2 . A straightforward computation shows that

$$(29) \quad \int_{\{-1, 0, 1\}^{\mathbb{Z}}} \theta^{j_0} \circ S^{k_0} \cdot \theta^{j_1} \circ S^{k_1} \cdot \dots \cdot \theta^{j_r} \circ S^{k_r} d\widehat{\nu}_{\mu^2} = 0$$

whenever $\{j_0, \dots, j_r\} \neq \{2\}$. On the other hand, in view of Lemma 3.1, the values of integrals

$$\int_{\{-1, 0, 1\}^{\mathbb{Z}}} \theta^2 \circ S^{k_0} \cdot \theta^2 \circ S^{k_1} \cdot \dots \cdot \theta^2 \circ S^{k_r} d\widehat{\nu}_{\mu^2}$$

for all $k_i \in \mathbb{Z}$ and $r \geq 0$ entirely determine the Mirsky measure ν_{μ^2} .

Corollary 3.2. The Chowla conjecture holds if and only if μ is a generic point for $\widehat{\nu}_{\mu^2}$.

Proof. We consider any extension of μ to a two-sided sequence (for example we set $\mu(n) = 0$ for each $n \leq 0$). Suppose that

$$(30) \quad \frac{1}{N_k} \sum_{n \leq N_k} \delta_{S^n \mu} \xrightarrow[k \rightarrow \infty]{} \kappa.$$

In order to get $\kappa = \widehat{\nu}_{\mu^2}$, in view of Lemma 3.1, we need to show that

$$\int_{\{-1, 0, 1\}^{\mathbb{Z}}} \theta^{j_0} \circ S^{k_0} \cdot \theta^{j_1} \circ S^{k_1} \cdot \dots \cdot \theta^{j_r} \circ S^{k_r} d\kappa = 0$$

for any choice of integers $k_0 < k_1 < \dots < k_r$, $\{j_0, j_1, \dots, j_r\} \neq \{2\}$ and $r \geq 0$. Since the measure ν is S -invariant, it is the same as to show that

$$\int_{\{-1, 0, 1\}^{\mathbb{Z}}} \theta^{j_0} \cdot \theta^{j_1} \circ S^{k_1 - k_0} \cdot \dots \cdot \theta^{j_r} \circ S^{k_r - k_0} d\kappa = 0.$$

²⁵The point μ^2 is recurrent, so there is a “completion” of μ^2 to a two-sided sequence generating the same subshift.

²⁶Consider Bernoulli measure $B(1/2, 1/2)$ on $\{-1, 1\}^{\mathbb{Z}}$ and Mirsky measure ν_{μ^2} on $\{0, 1\}^{\mathbb{Z}}$. Measure $\widehat{\nu}_{\mu^2}$ is the image of the product measure $B(1/2, 1/2) \otimes \nu_{\mu^2}$ via the map

$$(x, y) \mapsto ((x(n) \cdot y(n)))_{n \in \mathbb{Z}} \in \{-1, 0, 1\}^{\mathbb{Z}}.$$

Now, we have $1 \leq k_1 - k_0 < \dots < k_r - k_0$ and the result follows from (25) and (30). \square

The Chowla conjecture for $r = 0$ is just the PNT, however, it remains open even for $r = 1$. As in [148], we could consider a weaker version of the Chowla conjecture. Namely, we say that μ satisfies the *topological Chowla conjecture* if $X_\mu = s^{-1}(X_{\mu^2})$.

Remark 3.3. Note that (25) holds if

$$|\{0 \leq t \leq r : j_t = 1\}| = 1.$$

Indeed, it is not hard to see that if t_0 is the only index for which $j_{t_0} = 1$ then the sequence $a(n) := \prod_{t \neq t_0} \mu^2(n + k_t)$ is rational. Hence, μ is orthogonal to $a(\cdot)$, cf. footnote 18.

3.2 The Chowla conjecture implies Sarnak's conjecture

Assume that (X, T) is a topological system. Following [99, 166] a point $x \in X$ is called *completely deterministic* if for each measure $\nu \in \text{Q-gen}(x)$ (see (26)), the measure theoretic dynamical system $(X, \mathcal{B}(X), \nu, T)$ has zero Kolmogorov-Sinai entropy: $h_\nu(T) = 0$. Of course, if the topological entropy of T is zero, then by the Variational Principle, each $x \in X$ is completely deterministic. On the other hand, (X_{μ^2}, S) has positive topological entropy [6, 136, 148] and $\mu^2 \in X_{\mu^2}$ is completely deterministic, see [3, 33].

Let $f \in C(X)$ and $x \in X$ be completely deterministic. We have

$$\frac{1}{N} \sum_{n \leq N} f(T^n x) \mu(n) = \int_{X \times X_\mu} (f \otimes \theta) d \left(\frac{1}{N} \sum_{n \leq N} \delta_{(T \times S)^n(x, \mu)} \right).$$

We can assume that

$$\frac{1}{N_k} \sum_{n \leq N_k} \delta_{(T \times S)^n(x, \mu)} \xrightarrow[k \rightarrow \infty]{} \rho \text{ in the space } M(X \times X_\mu, T \times S).$$

Under the Chowla conjecture, the projection of ρ on X_μ is equal to $\hat{\nu}_{\mu^2}$ (since, by Corollary 3.2, μ is a generic point for $\hat{\nu}_{\mu^2}$), while the projection of ρ on X is some T -invariant measure κ and $h_\kappa(T) = 0$ (since x is completely deterministic). Note that ρ is a joining²⁷ of the (measure-theoretic) dynamical systems (X, κ, T) and $(X_\mu, \hat{\nu}_{\mu^2}, S)$. Moreover, the latter automorphism has the so called relative Kolmogorov property with respect to the factor $(X_{\mu^2}, \nu_{\mu^2}, S)$. We then consider the restriction of the joining $\rho|_{X \times X_{\mu^2}}$ and $\rho|_{X_\mu}$ to obtain two systems that have a common factor (namely X_{μ^2}) relatively to which the first one has zero entropy and the second being relatively Kolmogorov. Since the function θ is orthogonal to $L^2(X_{\mu^2}, \nu_{\mu^2})$, the relative disjointness theorem on zero entropy and Kolmogorov property yields the following (see also Remark 3.6):

²⁷Recall that if R_i is an automorphism of a probability standard Borel space $(Z_i, \mathcal{D}_i, \nu_i)$, $i = 1, 2$, then each $R_1 \times R_2$ -invariant measure λ on $(Z_1 \times Z_2, \mathcal{D}_1 \otimes \mathcal{D}_2)$ having the projections ν_1 and ν_2 , respectively is called a *joining* of R_1 and R_2 : we write $\lambda \in J(R_1, R_2)$. If R_1, R_2 are ergodic then the set $J^e(R_1, R_2)$ of ergodic joinings between R_1 and R_2 is non-empty. A fundamental notion here is the *disjointness* (in sense of Furstenberg) [70]: R_1 and R_2 are disjoint if $J(R_1, R_2) = \{\nu_1 \otimes \nu_2\}$: we write $R_1 \perp R_2$. For example, zero entropy automorphisms are disjoint with automorphisms having completely positive entropy (Kolmogorov automorphisms) and also a relativized version of this assertion holds.

Theorem 3.4 ([3]). *The Chowla conjecture implies*

$$\frac{1}{N} \sum_{n \leq N} f(T^n x) \boldsymbol{\mu}(n) \rightarrow 0$$

for each dynamical system (X, T) , $f \in C(X)$ and $x \in X$ completely deterministic. In particular, the Chowla conjecture implies Sarnak's conjecture.²⁸

Remark 3.5. It is also proved in [3] that this seemingly stronger statement of the validity of Sarnak's conjecture at completely deterministic points is in fact equivalent to the Möbius disjointness of all zero entropy systems.

Remark 3.6. A word for word repetition of the above proof²⁹ yields the same result when we replace $\boldsymbol{\mu}$ by another generic point of $\widehat{\nu}_{\boldsymbol{\mu}^2}$ in which we control the relative Kolmogorov property over the maximal factor with zero entropy, so called Pinsker factor. In particular, we can replace $\boldsymbol{\mu}$ by $\boldsymbol{\lambda}$ (for which the Pinsker factor will be just the one-point dynamical system).

As a matter of fact, it is expected that each aperiodic real-valued multiplicative function satisfies the Chowla type result (and hence satisfies the Sarnak type result), see the conjectures by Frantzikinakis and Host formulated after Theorem 3.30.

Remark 3.7. The original proof of Sarnak of the implication “Chowla conjecture \Rightarrow Sarnak's conjecture” used some combinatorial arguments and probabilistic methods, see [157].

Sarnak's conjecture (2) is formulated for the Möbius function. But of course one can consider other multiplicative functions.³⁰ Below, we show that if we use the Liouville function then nothing changes.

Corollary 3.8. Sarnak's conjecture with respect to $\boldsymbol{\mu}$ is equivalent to Sarnak's conjecture with respect to $\boldsymbol{\lambda}$.

Proof. Let us recall the basic relation between these two functions: $\boldsymbol{\lambda}(n) = \sum_{d^2 | n} \boldsymbol{\mu}(n/d^2)$.

Assume that (X, T) is a dynamical system with $h(T) = 0$. As the zero entropy class is closed under taking powers, we assume Möbius disjointness for all powers of T . Then

$$\begin{aligned} \frac{1}{N} \sum_{n \leq N} f(T^n x) \boldsymbol{\lambda}(n) &= \frac{1}{N} \sum_{n \leq N} f(T^n x) \left(\sum_{d^2 | n} \boldsymbol{\mu}(n/d^2) \right) \\ &= \frac{1}{N} \sum_{n \leq N} \sum_{d^2 | n} \boldsymbol{\mu}(n/d^2) f((T^{d^2})^{n/d^2} x) \\ &= \sum_{d \leq \sqrt{N}} \frac{1}{d^2} \cdot \frac{1}{N/d^2} \sum_{n \leq N/d^2} \boldsymbol{\mu}(n) f((T^{d^2})^n x). \end{aligned}$$

²⁸We will see later that some special cases of validity of convergence in (25) also have their ergodic interpretations and they imply Möbius disjointness for restricted classes of dynamical systems of zero entropy; in particular, see Corollary 3.20 and Corollary 3.25.

²⁹The above proof was already suggested by Sarnak in [148].

³⁰If Möbius disjointness in a dynamical system is shown through the KBSZ criterion then we obtain orthogonality with respect to all multiplicative functions.

Take $\varepsilon > 0$ and select $M \geq 1$ so that $\sum_{d \geq M} \frac{1}{d^2} < \varepsilon$. Consider T, T^2, T^3, \dots, T^M . We have

$$\left| \frac{1}{N} \sum_{n \leq N} f(T^{kn}x) \mu(n) \right| < \varepsilon$$

for all $k = 1, \dots, M$ whenever $N \geq N_0$. It follows that

$$\left| \frac{1}{N/d^2} \sum_{n \leq N/d^2} \mu(n) f(T^{d^2 n}x) \right| < \varepsilon$$

for all $d = 1, \dots, M$ if $N > MN_0$. Otherwise we estimate such a sum by $\|f\|_\infty$.

To obtain the other direction, we first recall that μ^2 is a completely deterministic point. Then use Theorem 3.4 for λ (see Remark 3.6), write $\lambda(n)\mu^2(n) = \mu(n)$ for each $n \geq 1$ and we obtain

$$\frac{1}{N} \sum_{n \leq N} f(T^n x) \mu^2(n) \lambda(n) = \frac{1}{N} \sum_{n \leq N} (f \otimes \theta)((T \times S)^n(x, \mu^2)) \lambda(n) \rightarrow 0$$

as the point (x, μ^2) is completely deterministic. \square

3.3 The logarithmic versions of Chowla and Sarnak's conjectures

An intriguing problem arises whether the Chowla and Sarnak's conjecture are equivalent. An intuition from ergodic theory would say that this is rather not the case as the class of systems that are disjoint (in the Furstenberg sense) from all zero entropy measure-theoretic systems is the class of Kolmogorov automorphisms and not only Bernoulli automorphisms (and a relative version of this result persists).³¹

From that point of view a recent remarkable result of Terence Tao [155] about the equivalence of logarithmic versions of the Chowla and Sarnak's conjectures is quite surprising. We will formulate some versions³² of three (out of five) conjectures from [155].

Conjecture A: We have

$$\frac{1}{\log N} \sum_{n \leq N} \frac{\mu^{j_0}(n) \mu^{j_1}(n+k_1) \dots \mu^{j_r}(n+k_r)}{n} \xrightarrow{N \rightarrow \infty} 0$$

whenever $1 \leq k_1 < \dots < k_r$, $j_s \in \{1, 2\}$ not all equal to 2, $r \geq 0$.

³¹If we consider general sequences $z \in \{-1, 0, 1\}^{\mathbb{N}}$ then we can speak about the Sarnak and Chowla properties on a more abstract level: for example the Chowla property of z means (25) with μ replaced by z . See Example 5.1 and Remark 5.3 in [3] for sequences orthogonal to all deterministic sequences but not satisfying the Chowla property. However, arithmetic functions in these examples are not multiplicative.

However, an analogy between disjointness results in ergodic theory and disjointness of sequences is sometimes accurate. For example, a measure-theoretic dynamical system has zero entropy if and only if it is disjoint with all Bernoulli automorphisms. As pointed out in [3] (Prop. 5.21), a sequence $t \in \{-1, 1\}^{\mathbb{N}}$ is completely deterministic if and only if it is disjoint with any sequence $z \in \{-1, 0, 1\}^{\mathbb{N}}$ satisfying the Chowla property.

³²See Remark 1.9. Also, in [155] the Liouville function λ is considered, see page 2 in [155] how to replace λ by μ .

Remark 3.9. It should be noted that passing to such logarithmic averages moves one away from questions about primes, twin primes and subtleties such as the parity problem. For example, the statement $\sum_{n \leq N} \frac{\mu(n)}{n} = o(\log N)$ is easy to establish (in fact, $\left| \sum_{n \leq N} \frac{\mu(n)}{n} \right| \leq 1$), while the PNT is equivalent to much stronger statement $\sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0$ (as conditionally convergent series).

On the other hand, the logarithmically averaged Chowla conjecture implies that all “admissible” configurations do appear on μ , see Corollary 3.13 below (the topological Chowla conjecture for λ implies that all blocks of ± 1 appear in λ).

Conjecture B: We have

$$\frac{1}{\log N} \sum_{n \leq N} \frac{f(T^n x) \mu(n)}{n} \xrightarrow{N \rightarrow \infty} 0$$

whenever (X, T) is a topological system of zero topological entropy, $f \in C(X)$ and $x \in X$.

To formulate the third conjecture, we need to recall the definition of a nilrotation. Let G be a connected, simply connected Lie group and $\Gamma \subset G$ a lattice (a discrete, cocompact subgroup). For any $g_0 \in G$ we define $T_{g_0}(gH) := g_0 g H$. Then the topological system $(G/\Gamma, T_{g_0})$ is called a *nilrotation*.

Conjecture C: Let $f \in C(G/\Gamma)$ be Lipschitz continuous and $x_0 \in G$. Then (for $H \leq N$)

$$\sum_{n \leq N} \frac{\sup_{g \in G} \left| \sum_{h \leq H} f(T_g^{h+n}(x_0 \Gamma)) \mu(n+h) \right|}{n} = o(H \log N).$$

Theorem 3.10 ([155]). *Conjectures A, B and C are equivalent.*

Remark 3.11. Tao also shows that if instead of logarithmic averages we come back to Cesàro averages, then

$$\text{Conjecture A} \Rightarrow \text{Conjecture B} \Rightarrow \text{Conjecture C}$$

and it is the implication Conjecture C \Rightarrow Conjecture A that requires logarithmic averages.

Remark 3.12. Let us consider the Cesàro version of Conjecture C with $H = o(N)$ and we drop the assumption on the sup (which is inside), i.e.: for each $g \in G$, we have

$$\frac{1}{N} \sum_{n \leq N} \left| \sum_{h \leq H} f(T_g^{h+n}(x_0 \Gamma)) \mu(n+h) \right| \xrightarrow{H, N \rightarrow \infty, H = o(N)} 0.$$

This is a particular case of what we will see in Section 4, where we introduce the strong MOMO notion (hence, the validity of Sarnak’s conjecture on (typical) short interval).

Corollary 3.13 (a letter of W. Veech in June 2016). Sarnak’s conjecture implies topological Chowla conjecture. Equivalently, Sarnak’s conjecture implies that each block $B \in \{-1, 0, 1\}^\ell$ for which B^2 appears in μ^2 appears in μ (and the entropy of (X_μ, S) equals $\frac{6}{\pi^2} \log 3$).

Proof. Indeed, Sarnak’s conjecture implies its logarithmic version which, by Theorem 3.10, implies logarithmic Chowla conjecture, that is, $\frac{1}{\log N} \sum_{n \leq N} \frac{\delta_{S^n \mu}}{n} \rightarrow \widehat{\nu}_{\mu^2}$. However, the logarithmic averages of the Dirac measures are convex combinations of the consecutive Cesàro averages³³ $\frac{1}{n} \sum_{j \leq n} \delta_{S^j \mu}$, so if we take a block $B \in s^{-1}(X_{\mu^2})$, we have $\widehat{\nu}_{\mu^2}(B) > 0$ and therefore there exists n such that $\frac{1}{n} \sum_{j \leq n} \delta_{S^j \mu}(B) > 0$, which means that B appears in μ . \square

Remark 3.14. (added in October 2017) As a matter of fact, as shown in [76], Sarnak’s conjecture implies the existence of a subsequence (N_k) along which $\frac{1}{N_k} \sum_{n \leq N_k} \delta_{S^n \mu} \rightarrow \widehat{\nu}_{\mu^2}$. This follows from a general observation that, given a topological system (X, T) , whenever an ergodic measure ν is a limit of a subsequence (M_k) of logarithmic averages of Dirac measures: $\nu = \lim_{k \rightarrow \infty} \frac{1}{\log M_k} \sum_{m \leq M_k} \frac{\delta_{T^m x}}{m}$, then there exists a subsequence (N_k) for which $\nu = \lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \delta_{T^n x}$. We apply this to the measure $\widehat{\nu}_{\mu^2}$ which is ergodic.

In [154], Tao proves the logarithmic version of Chowla conjecture for the correlations of order 2 (which we formulate for the Liouville function):

Theorem 3.15 ([154]). *For each $0 \neq h \in \mathbb{Z}$, we have*

$$\frac{1}{\log N} \sum_{n \leq N} \frac{\lambda(n)\lambda(n+h)}{n} \xrightarrow{N \rightarrow \infty} 0.$$

See also [124], where it is proved that for each integer $h \geq 1$ there exists $\delta(h) > 0$ such that $\limsup_{N \rightarrow \infty} \frac{1}{N} \left| \sum_{n \leq N} \lambda(n)\lambda(n+h) \right| \leq 1 - \delta(h)$ and [126], where it is proved that for the Liouville function the eight patterns of length 3 of signs occur with positive lower density, and the density result with lower density replaced by upper density persists for $k+5$ patterns (out of total 2^k) for each $k \in \mathbb{N}$.

For a proof of a function field Chowla’s conjecture, see [32].

³³ Assume that (a_n) is a bounded sequence and set $A_n = a_1 + \dots + a_n$. Then, we have by summation by parts

$$\begin{aligned} (31) \quad \frac{1}{\log N} \sum_{n \leq N} \frac{a_n}{n} &= \frac{1}{\log N} \sum_{n \leq N} (A_{n+1} - A_n) \frac{1}{n} \\ &= \frac{1}{\log N} \sum_{n \leq N} A_n \left(\frac{1}{n} - \frac{1}{n+1} \right) + o(1) = \frac{1}{\log N} \sum_{n \leq N} \frac{A_n}{n} \frac{1}{n+1} + o(1). \end{aligned}$$

It follows that:

- If the Cesàro averages of (a_n) converge, so do the logarithmic averages of (a_n) .
- The converse does not hold (see e.g. [24] in \mathcal{B} -free case, Section 6.1).
- If the Cesàro averages converge along a subsubsequence (N_k) then not necessarily the logarithmic averages do the same. Indeed, by (31), $\frac{1}{\log N_k} \sum_{n \leq N_k} \frac{a_n}{n}$ is (up to a small error) a convex combination of the Cesàro averages for all $n \leq N_k$.

Remark 3.16. See also [158], where, given $k_0, \dots, k_\ell \in \mathbb{Z}$ and $\mathbf{u}_0, \dots, \mathbf{u}_\ell \in \mathcal{M}$, one studies sequences of the form

$$n \mapsto \mathbf{u}_0(n + ak_0) \cdot \dots \cdot \mathbf{u}_\ell(n + ak_\ell), \quad a \in \mathbb{Z}.$$

By considering their logarithmic averages, one obtains a sequence $(f(a))$. The main result of [158] is a structure theorem (depending on whether or not the product $\mathbf{u}_0 \cdot \dots \cdot \mathbf{u}_\ell$ weakly pretends to be a Dirichlet character) for the sequences $(f(a))$. As a corollary, the logarithmically averaged Chowla conjecture is proved for any odd number of shifts.

3.4 Frantzikinakis' theorem

Tao's approach from [155] is continued in [66]. Before we formulate Frantzikinakis' results, let us interpret some arithmetic properties, especially the role of a "good behavior" on (typical) short interval of a multiplicative function in the ergodic theory language.

3.4.1 Ergodicity of measures for which μ is quasi-generic

In this subsection we summarize ergodic consequences of some recent, previously mentioned number-theoretic results, cf. [65]. By that we mean that we consider all measures $\kappa \in \text{Q-gen}(\mu)$ and we study ergodic properties of the dynamical systems (X_μ, κ, S) .

Let $\kappa \in \text{Q-gen}(\mu)$, i.e. $\frac{1}{M_k} \sum_{m \leq M_k} \delta_{S^m \mu} \xrightarrow[k \rightarrow \infty]{} \kappa \in M(X_\mu, S)$ for some increasing sequence (M_k) . As usual, $\theta(x) = x(0)$ ($\theta \in C(X_\mu)$). We have

$$(32) \quad \int_{X_\mu} \theta d\kappa = 0,$$

as the integral equals $\lim_{k \rightarrow \infty} \frac{1}{M_k} \sum_{n \leq M_k} \theta(S^n \mu) = 0$ (by the PNT). Denoting by Inv the σ -algebra of S -invariant (modulo the measure κ) subsets of X_μ , we recall that

$$\frac{1}{H} \sum_{h \leq H} \theta \circ S^h \xrightarrow[H \rightarrow \infty]{} \mathbb{E}(\theta | Inv) \text{ in } L^2(X_\mu, \kappa)$$

(by the von Neumann ergodic theorem). We want to show that

$$\theta \perp L^2(X_\mu, Inv, \kappa)$$

(i.e. κ must be "slightly" ergodic). In other words, we want to show that

$$\int_{X_\mu} \left| \frac{1}{H} \sum_{h \leq H} \theta \circ S^h \right|^2 d\kappa \xrightarrow[H \rightarrow \infty]{} 0.$$

But such integrals can be computed:

$$\frac{1}{M_k} \sum_{m \leq M_k} \left| \frac{1}{H} \sum_{h \leq H} \theta \circ S^h(S^m \mu) \right|^2 \xrightarrow[k \rightarrow \infty]{} \int_{X_\mu} \left| \frac{1}{H} \sum_{h \leq H} \theta \circ S^h \right|^2 d\kappa.$$

Putting things together, given $\varepsilon > 0$, for $H \geq 1$ large enough, we want to see

$$\limsup_{k \rightarrow \infty} \frac{1}{M_k} \sum_{m \leq M_k} \left| \frac{1}{H} \sum_{h \leq H} \mu(m+h) \right|^2 \leq \varepsilon.$$

The latter is true because of [124]: for a „typical” m the sum $|\frac{1}{H} \sum_{m \leq h < m+H} \mu(h)|$ is small.

Remark 3.17. As the calculation above shows, the fact that

$$\frac{1}{M} \sum_{m \leq M} \left| \frac{1}{H} \sum_{h \leq H} \mu(m+h) \right|^2 \rightarrow 0$$

when $H \rightarrow \infty$ and $H = o(M)$ is equivalent to $\theta \perp L^2(X_\mu, Inv, \kappa)$ for each $\kappa \in \text{Q-gen}(\mu)$. In particular, the Chowla conjecture implies the above short interval behavior.

However, remembering that $\kappa|_{X_{\mu^2}} = \nu_{\mu^2}$, one can ask now whether θ is measurable with respect to the factor given by the Mirsky measure. As this factor has rational discrete spectrum [33], to show that this is not the case, we need to prove that $\theta \perp L^2(\Sigma_{rat})$, where Σ_{rat} stands for the factor given by the whole rational spectrum of (X_μ, κ, S) . To do it, we need to show that for each $r \geq 1$, we have

$$\frac{1}{N} \sum_{n \leq N} \theta \circ S^{rn} \xrightarrow{N \rightarrow \infty} 0 \text{ in } L^2(X_\mu, \kappa).$$

This convergence can be shown by using the strong MOMO property (which we will consider in Section 4) for the rotation $j \mapsto j+1$ on $\mathbb{Z}/r\mathbb{Z}$. We skip this argument here and show still a stronger consequence.

Assume that $\kappa \in \text{Q-gen}(\mu)$ and that we want to show that the spectral measure of $\theta \in L^2(X_\mu, \kappa)$ is continuous. Hence, we need to show that

$$\frac{1}{H} \sum_{h \leq H} |\widehat{\sigma}_\theta(h)| \xrightarrow{H \rightarrow \infty} 0$$

when $H \rightarrow \infty$. Equivalently, we need to show that

$$\frac{1}{H} \sum_{h \leq H} \left| \int_{X_\mu} \theta \circ S^h \cdot \theta \, d\kappa \right| \xrightarrow{H \rightarrow \infty} 0.$$

If we fix $H \geq 1$ then

$$\begin{aligned} \int_{X_\mu} \theta \circ S^h \cdot \theta \, d\kappa &= \lim_{k \rightarrow \infty} \frac{1}{M_k} \sum_{m \leq M_k} \theta \circ S^h(S^m \mu) \cdot \theta(S^m \mu) \\ &= \frac{1}{M_k} \sum_{m \leq M_k} \mu(m+h) \mu(m). \end{aligned}$$

It follows that we need to show that

$$\frac{1}{H} \sum_{h \leq H} \left| \frac{1}{M_k} \sum_{m \leq M_k} \mu(m+h) \mu(m) \right| \rightarrow 0$$

when $H, M_k \rightarrow \infty$; to be precise, given $\varepsilon > 0$ we want to show that for $H > H_\varepsilon$, we have $\limsup_{k \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \left| \frac{1}{M_k} \sum_{m \leq M_k} \boldsymbol{\mu}(m+h) \boldsymbol{\mu}(m) \right| < \varepsilon$. Hence, directly from Theorem 2.12, we obtain the following.

Corollary 3.18. The spectral measure of θ is continuous for each $\kappa \in \text{Q-gen}(\boldsymbol{\mu})$.

While it is obvious that the subshift $X_\boldsymbol{\mu}$ is uncountable (indeed, it is the subshift $X_{\boldsymbol{\mu}^2}$ which is already uncountable, see Section 6), it is not clear whether X_λ is uncountable. However, if a subshift (Y, S) is countable, all its ergodic measures are given by periodic orbits, hence there are only countably many of them and it easily follows that each $\kappa \in M(Y, S)$ yield a system with discrete spectrum. Hence, immediately from Corollary 3.18, we obtain that:

Corollary 3.19. The subshift X_λ is uncountable.³⁴

From Corollary 3.18 we derive immediately the Möbius disjointness of all dynamical systems with “trivial” invariant measures (see also [91]). This kind of problems will be the main subject of our discussion in Section 4.

Corollary 3.20. Let (X, T) be any topological dynamical system such that, for each measure $\nu \in M(X, T)$, (X, ν, T) has discrete spectrum (not necessarily ergodic, of course). Then (X, T) is Möbius disjoint. In particular, the result holds if $M^e(X, T)$ is countable with each member of $M^e(X, T)$ yielding a discrete spectrum dynamical system.

Proof. Fix $x \in X$ and consider

$$\frac{1}{M_k} \sum_{m \leq M_k} \delta_{(T^m x, S^m \boldsymbol{\mu})} \xrightarrow{k \rightarrow \infty} \rho.$$

We have $\rho|_{X_\boldsymbol{\mu}} =: \kappa \in \text{Q-gen}(\boldsymbol{\mu})$ and $\rho|_X =: \nu$. Now, we fix $f \in C(X)$ and we need to show that $\int f \otimes \theta d\rho = 0$. But

$$(33) \quad \int_{X \times X_\boldsymbol{\mu}} f \otimes \theta d\rho = \int_{X \times X_\boldsymbol{\mu}} (f \otimes 1) \cdot (1 \otimes \theta) d\rho = 0.$$

Indeed, the spectral measure of $f \otimes 1$ with respect to ρ is the same as the spectral measure of f with respect to ν and the spectral measure of $1 \otimes \theta$ with respect to ρ is the same as the spectral measure of θ with respect to κ . Therefore, these spectral measures are mutually singular by assumption and Corollary 3.18. Hence, the functions $f \otimes 1$ and $1 \otimes \theta$ are orthogonal, i.e. (33) holds.³⁵ \square

If we have all ergodic measures giving discrete spectrum but we have too many ergodic measures then the argument above does not go through. Consider

$$(*) \quad (x, y) \mapsto (x, x + y) \text{ on } \mathbb{T}^2. \quad ^{36}$$

Question 1 (Frantzikinakis (2016)). Can we obtain $\kappa \in \text{Q-gen}(\lambda)$, so that (X_λ, κ, S) is isomorphic to $(*)$?

Of course, the answer to Question 1 is expected to be negative.

³⁴The result has been observed in [68], cf. also [90].

³⁵We use here the standard result in the theory of unitary operators that mutual singularity of spectral measures implies orthogonality. Recall also the classical result in ergodic theory that spectral disjointness implies disjointness.

³⁶Consider $X_1 = X_2 = \mathbb{T}^2$ with $\mu_1 = \mu_2 = \text{Leb}_{\mathbb{T}^2}$, the diagonal joining Δ on $X_1 \times X_2$ and $f(x, y) = \theta(x, y)$ with $\theta(x, y) = e^{2\pi i y}$. The spectral measure of θ is Lebesgue, and all ergodic components of the measure μ_1 have discrete spectra.

3.4.2 Frantzikinakis' results

We now follow [66] and formulate results for the Liouville function, although, up to some obvious modifications, they also hold for μ .

Theorem 3.21 ([66]). *Assume that $N_k \rightarrow \infty$ and let $\frac{1}{\log N_k} \sum_{n \leq N_k} \frac{\delta_{S^n \lambda}}{n} \xrightarrow[k \rightarrow \infty]{} \kappa$. If κ is ergodic then the Chowla conjecture (and Sarnak's conjecture) holds along (N_k) for the logarithmic averages.*

Taking into account footnote 33, we cannot deduce a similar statement for ordinary averages along (N_k) but in view of [76], see Remark 3.14, the Chowla conjecture holds along another subsequence. The situation becomes clear when (N_k) is the sequence of all natural numbers and we assume genericity.

Corollary 3.22 ([66]). *If λ is generic for an ergodic measure then the Chowla conjecture holds.*

Let us say a few words on the proof. Recall that given a bounded sequence $(a(n)) \subset \mathbb{C}$ admitting correlations,³⁷ one defines its local uniformity seminorms (see Host and Kra [87]) in the following manner:

$$(34) \quad \|a\|_{U^1(\mathbb{N})}^2 = \mathbb{E}_{h \in \mathbb{N}} \mathbb{E}_{n \in \mathbb{N}} a(n+h) \overline{a(n)},$$

$$(35) \quad \|a\|_{U^{s+1}(\mathbb{N})}^{2^{s+1}} = \mathbb{E}_{h \in \mathbb{N}} \|S_h a \cdot \bar{a}\|_{U^s(\mathbb{N})}^{2^s}, \quad s \geq 2,$$

where, for each bounded sequence $(b(n))$, $(S_h b)(n) := b(h+n)$ and $\mathbb{E}_{n \in \mathbb{N}} b(n) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} b(n)$. (Similar definitions are considered along a subsequence (N_k) .)

The following result has been proved by Tao:

Theorem 3.23 ([155]). *Assume that λ is generic. The Chowla conjecture holds if and only if $\|\lambda\|_{U^s(\mathbb{N})} = 0$ for each $s \geq 1$.*³⁸

Remark 3.24. We have assumed in the statement of Theorem 3.23 that λ is generic but we would like also to note that, without this latter (strong) assumption, Tao obtained the equivalence in Theorem 3.23 for the logarithmic averages, see Conjecture 1.6 and Theorem 1.9 in [155] (however, one has to modify the definition of seminorms [155]).

Hence, under the assumption of Corollary 3.22, we need to prove that all local uniform seminorms of λ vanish. The inverse theorem for seminorms reduces this problem to the statement: for every basic nilsequence $(a(n))$ ³⁹ on an $s-1$ -step nilmanifold G/Γ and every $s-2$ -step manifold H/Λ , we have

$$\lim_{N \rightarrow \infty} \mathbb{E}_{m \in \mathbb{N}} \sup_{b \in \Psi_{H/\Lambda}} \left| \mathbb{E}_{n \in [m, m+N]} \lambda(n) a(n) b(n) \right| = 0,$$

where $\Psi_{H/\Lambda}$ is a special class of basic nil-sequences (coming from Lipschitz functions). The latter is then proved using a deep induction argument.

³⁷I.e., we assume the existence of the limits of sequences $(\frac{1}{N} \sum_{n \leq N} a'(n) a'(n+k_1) \dots a'(n+k_r))_{N \geq 1}$ for every $r \in \mathbb{N}$ and $k_1, \dots, k_r \in \mathbb{N}$ (not necessarily distinct) with $a' = a$ or \bar{a} . It is not hard to see that a admits correlations if and only if it is generic, cf. Section 3.1.

³⁸We have $\|\lambda\|_{U^1(\mathbb{N})} = 0$ by [124], moreover $\|\lambda\|_{U^2(\mathbb{N})} = 0$ is equivalent to $\lim_{N \rightarrow \infty} \mathbb{E}_{m \in \mathbb{N}} \sup_{\alpha \in [0,1]} \left| \mathbb{E}_{n \in [m, m+N]} \lambda(n) e^{2\pi i n \alpha} \right| = 0$ (cf. Conjecture C) and remains open. For a subsequence version of Theorem 3.23 for logarithmic averages, see [155].

³⁹By that we mean $a(n) = f(g^n \Gamma)$ for some continuous $f \in C(G/\Gamma)$ and $g \in G$.

3.5 Dynamical properties of Furstenberg systems associated to the Liouville and Möbius functions

We now continue considerations about logarithmic version of Sarnak's conjecture, cf. Conjecture B, Theorem 3.21. Consider all measures κ for which λ is logarithmically quasi-generic, i.e. $\frac{1}{\log N_k} \sum_{n \leq N_k} \frac{\delta_{S^n \lambda}}{n} \rightarrow \kappa$ for some $N_k \rightarrow \infty$. We denote the set of all such measures by $\log\text{-Q-gen}(\lambda)$. Following [68], for each $\kappa \in \log\text{-Q-gen}(\lambda)$ the corresponding measure-theoretic dynamical system (X_λ, κ, S) will be called a *Furstenberg system of λ* . Before we get closer to the results of [68], let us see first some consequence of Theorem 3.15 for the logarithmic Sarnak's conjecture:

(36) For each Furstenberg system (X_λ, κ, S) , the spectral measure σ_θ of θ is Lebesgue.

Indeed, assuming $\frac{1}{\log N_k} \sum_{n \leq N_k} \frac{\delta_{S^n \lambda}}{n} \xrightarrow[k \rightarrow \infty]{} \kappa$, Theorem 3.15 tells us that for each $h \in \mathbb{Z} \setminus \{0\}$, we have

$$\widehat{\sigma}_\theta(h) = \int_{X_\lambda} \theta \circ S^h \cdot \theta \, d\kappa = \lim_{k \rightarrow \infty} \frac{1}{\log N_k} \sum_{n \leq N_k} \frac{\lambda(n+h)\lambda(n)}{n} = 0.$$

Using (36) and repeating the proof of Corollary 3.20, we obtain the following.

Corollary 3.25. Let (X, T) be a topological system such that each of its Furstenberg's systems has singular spectrum. Then (X, T) is logarithmically Liouville disjoint.

The starting point of the paper [68] is a surprising Tao's identity (implicit in [154]) for general sequences which in its ergodic theory language (cf. Subsection 3.4.1) takes the following form:

Theorem 3.26 (Tao's identity, [68]). *Let $\kappa \in \log\text{-Q-gen}(\lambda)$. Then*

$$\int_{X_\lambda} \left(\prod_{j=1}^{\ell} \theta \circ S^{k_j} \right) d\kappa = (-1)^\ell \lim_{N \rightarrow \infty} \frac{\log N}{N} \sum_{\mathbb{P} \ni p \leq N} \int_{X_\lambda} \left(\prod_{j=1}^{\ell} \theta \circ S^{pk_j} \right) d\kappa$$

for all $\ell \in \mathbb{N}$ and $k_1, \dots, k_\ell \in \mathbb{Z}$.

Now, the condition in Theorem 3.26 is purely abstract (indeed, the function θ generates the Borel σ -algebra), and the strategy to cope with logarithmic Sarnak's conjecture is to describe the class of measure-theoretic dynamical systems satisfying the assertion of Theorem 3.26 and then to obtain Liouville disjointness for all systems which are disjoint (in the Furstenberg sense) from all members of the class. In fact, Frantzikinakis and Host deal with extensions of Furstenberg systems of λ , so called *systems of arithmetic progressions with prime steps*.⁴⁰

⁴⁰Given a measure-theoretic dynamical system $(Z, \mathcal{D}, \rho, R)$, its system of arithmetic progressions with prime steps is of the form $(Z^{\mathbb{Z}}, \mathcal{B}(Z^{\mathbb{Z}}), \tilde{\rho}, S)$, where S is the shift and the (shift invariant) measure $\tilde{\rho}$ is determined by

$$\int_{Z^{\mathbb{Z}}} \prod_{j=-m}^m f_j(z_j) d\tilde{\rho}(z) = \lim_{N \rightarrow \infty} \frac{\log N}{N} \sum_{p \leq N} \int_Z \prod_{j=-m}^m f_j \circ R^{pj} d\rho$$

for all $m \geq 0$, $f_{-m}, \dots, f_m \in L^\infty(Z, \rho)$ (here $z = (z_j)$). It is proved that such shift systems have no irrational spectrum. One of key observations is that each Furstenberg system of the Liouville function is a factor of the associated system of arithmetic progressions with prime steps.

They prove the following result.

Theorem 3.27 ([68]). *For each system of arithmetic progressions with prime steps, its “typical” ergodic component is isomorphic to the direct product of an infinite-step nilsystem and a Bernoulli automorphism.⁴¹ In particular, each Furstenberg system (X_λ, κ, S) of λ is a factor of a system which:*

- (i) *has no irrational spectrum and*
- (ii) *has ergodic components isomorphic to the direct product of an infinite-step nilsystem and a Bernoulli automorphism.*

Remark 3.28. All the above results are also true when we replace λ by μ .

Then, some new disjointness results in ergodic theory are proved (for example, all totally ergodic automorphisms are disjoint from an automorphism satisfying (i) and (ii) in Theorem 3.27) and the following remarkable result is obtained:

Theorem 3.29 ([68]). *Let (X, T) be a topological dynamical system of zero entropy with countably many ergodic invariant measures. Then Conjecture B holds for (X, T) .*

In particular, logarithmic Sarnak’s conjecture holds for all zero entropy uniquely ergodic systems. As a matter of fact, some new⁴² consequences are derived:

Theorem 3.30 ([68]). *Let (X, T) be a topological dynamical system with zero entropy. Assume that $x \in X$ is generic for a measure ν with only countably many ergodic components all of which yield totally ergodic systems. Then, for every $f \in C(X)$, $\int_X f d\nu = 0$, we have*

$$\lim_{N \rightarrow \infty} \frac{1}{\log N} \sum_{n \leq N} \frac{f(T^n x) \prod_{j=1}^{\ell} \mu(n + k_j)}{n} = 0$$

for all $\ell \in \mathbb{N}$ and $k_1, \dots, k_\ell \in \mathbb{Z}$.

New conjectures are proposed in [68]:

1. Every real-valued $\mathbf{u} \in \mathcal{M}$ has a unique Furstenberg system (i.e. \mathbf{u} is generic) which is ergodic and isomorphic to the direct product of a Bernoulli automorphism and an odometer.
2. If, additionally, $\mathbf{u} \in \mathcal{M}$ takes values ± 1 then its Furstenberg system is either Bernoulli or it is an odometer.

Finally, it is noticed in [68] that the complexity of the Liouville function has to be superlinear, that is

$$(37) \quad \lim_{N \rightarrow \infty} \frac{1}{N} |\{B \in \{-1, 1\}^N : B \text{ appears in } \lambda\}| = \infty.$$

The reason is that, as shown in [68], for transitive systems having linear block growth we have only finitely many ergodic measures (and clearly systems with linear block growth have zero topological entropy). Hence, by Theorem 3.29, such systems are Liouville disjoint. As X_λ is not Liouville disjoint, λ cannot have linear block growth, i.e. (37) holds.

⁴¹The product decomposition depends on the component.

⁴²They are new even for irrational rotations. Cf. the notions of (S)-strong and (S₀)-strong and their equivalence to the Chowla type condition in [3].

4 The MOMO and AOP properties

4.1 The MOMO property and its consequences

We will now consider Sarnak’s conjecture from the ergodic theory point of view. We ask whether (already) measure-theoretic properties of a measurable system $(Z, \mathcal{D}, \kappa, R)$ imply the validity of (1) for any (X, T) , $f \in C(X)$ provided that $x \in X$ is a generic point for a measure μ such that the measure-theoretic system $(X, \mathcal{B}(X), \mu, T)$ is measure-theoretically isomorphic to $(Z, \mathcal{D}, \kappa, R)$. More specifically, we can ask whether some measure-theoretic properties of $(Z, \mathcal{D}, \kappa, R)$ can imply Möbius disjointness of all its uniquely ergodic models.⁴³ We recall that the Jewett-Krieger theorem implies the existence of a uniquely ergodic model of each ergodic system.⁴⁴ As a matter of fact, there are plenty of such models and they can have various additional topological properties including topological mixing⁴⁵ [115]. Here is another variation of the approach to view Möbius disjointness as a measure-theoretic invariant:

Question 2. Does the Möbius disjointness in a certain uniquely ergodic model of an ergodic system yield the Möbius disjointness in all its uniquely ergodic models?

To cope with these questions we need a definition. Let $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$ be an arithmetic function.⁴⁶

Definition 4.1 (strong MOMO⁴⁷ property [4]). We say that (X, T) satisfies the *strong MOMO property* (relatively to \mathbf{u}) if, for any increasing sequence of integers $0 = b_0 < b_1 < b_2 < \dots$ with $b_{k+1} - b_k \rightarrow \infty$, for any sequence (x_k) of points in X , and any $f \in C(X)$, we have

$$(38) \quad \frac{1}{b_K} \sum_{k < K} \left| \sum_{b_k \leq n < b_{k+1}} f(T^{n-b_k} x_k) \mathbf{u}(n) \right| \xrightarrow{K \rightarrow \infty} 0.$$

Remark 4.1. The property (38) looks stronger than the condition on Möbius disjointness. The idea behind it is to look at the pieces of orbits (of different points) in one system as a single orbit of a point in a different, larger but “controllable” (from measure-theoretic point of view) system.

⁴³Note that the answer is positive in all uniquely ergodic models of the one-point system: each such a model has a unique fixed point that attracts each orbit on a subset of density 1, cf. the map $e^{2\pi i x} \mapsto e^{2\pi i x^2}$, $x \in [0, 1)$. This argument is however insufficient already for uniquely ergodic models of the exchange of two points: in this case we have a density 1 attracting 2-periodic orbit $\{a, b\}$, but we do not control to which point a or b the orbit returns first. Quite surprisingly, it seems that already in this case we need [124] to obtain Möbius disjointness of all uniquely ergodic models.

⁴⁴If all uniquely ergodic systems were Möbius disjoint, then as noticed by T. Downarowicz, we would get that the Chowla conjecture fails in view of the result of B. Weiss [168] Thm. 4.4’ on approximation of generic points of ergodic measures by uniquely ergodic sequences.

⁴⁵Topological mixing for example excludes the possibility of having eigenfunctions continuous.

⁴⁶Our objective is of course the Möbius function μ , however the whole approach can be developed for an arithmetic function satisfying some additional properties.

⁴⁷The acronym comes from Möbius Orthogonality of Moving Orbits.

Remark 4.2. One can easily show (as in Section 3.4.1) that the strong MOMO property (relative to μ) implies $f \otimes \theta \perp L^2(Inv, \rho)$ for each $\rho \in \text{Q-gen}((x, \mu), T \times S)$.⁴⁸

By taking $f = 1$ in Definition 4.1, we obtain that whenever strong MOMO holds, \mathbf{u} has to satisfy:

$$(39) \quad \frac{1}{b_K} \sum_{k < K} \left| \sum_{b_k \leq n < b_{k+1}} \mathbf{u}(n) \right| \xrightarrow{K \rightarrow \infty} 0$$

for every sequence $0 = b_0 < b_1 < b_2 < \dots$ with $b_{k+1} - b_k \rightarrow \infty$. In particular, $\frac{1}{N} \sum_{n < N} \mathbf{u}(n) \xrightarrow{N \rightarrow \infty} 0$. This is to be compared with (18), (20) and (19) to realize that we require a special behavior of \mathbf{u} on a typical short interval.

Theorem 4.3 ([4]). *Let $(Z, \mathcal{D}, \kappa, R)$ be an ergodic dynamical system. Let $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$ be an arithmetic function. The following conditions are equivalent:*

- (a) *There exist a topological system (Y, S) enjoying the strong MOMO property (relative to \mathbf{u}) and $\nu \in M^e(Y, S)$ such that the measurable systems $(Y, \mathcal{B}(Y), \nu, S)$ and $(Z, \mathcal{D}, \kappa, R)$ are isomorphic.*
- (b) *For any topological dynamical system (X, T) and any $x \in X$, if there exists a finite number of T -invariant measures μ_j , $1 \leq j \leq t$, such that*

- *$(X, \mathcal{B}(X), \mu_j, T)$ is measure-theoretically isomorphic to $(Z, \mathcal{D}, \kappa, R)$ for each j ,*
- *any measure for which x is quasi-generic is a convex combination of the measures μ_j , i.e. $Q\text{-gen}(x) \subset \text{conv}(\mu_1, \dots, \mu_t)$,*

then $\frac{1}{N} \sum_{n \leq N} f(T^n x) \mathbf{u}(n) \xrightarrow{N \rightarrow \infty} 0$ for each $f \in C(X)$.

- (c) *All uniquely ergodic models of $(Z, \mathcal{D}, \kappa, R)$ enjoy the strong MOMO property (relative to \mathbf{u}).*

The proof of implication (a) \Rightarrow (b) borrows some ideas from [91] and the proof of implication (b) \Rightarrow (c) uses some ideas from [7].

Remark 4.4. It can be easily shown that any minimal (hence uniquely ergodic) rotation on a compact Abelian group satisfies the strong MOMO property (say, relatively to μ). It follows from Theorem 4.3 (and the Halmos-von Neumann theorem) that in each uniquely ergodic model of an ergodic automorphism with discrete spectrum, we also have the strong MOMO property (in particular, the Möbius disjointness).

We now list three consequences of Theorem 4.3:

Corollary 4.5 ([3]). (a) If Sarnak's conjecture holds then the strong MOMO property (relative to μ) holds for every zero entropy dynamical system.⁴⁹

⁴⁸ Inv stands here for the σ -algebra of $T \times S$ -invariant sets modulo ρ .

⁴⁹ That is, Sarnak's conjecture and the strong MOMO property (relative to μ) for all deterministic systems are equivalent statements.

- (b) If Sarnak's conjecture holds then it holds uniformly, that is, the convergence in (1) is uniform in x .⁵⁰
- (c) Fix $\delta_{(\dots 0.00\dots)} \neq \kappa \in M^e((\mathbb{D}_L)^\mathbb{Z}, S)$, where $\mathbb{D}_L = \{z \in \mathbb{C} : |z| \leq L\}$. Let (X, T) be any uniquely ergodic model of $((\mathbb{D}_L)^\mathbb{Z}, \kappa, S)$. Then for any $\mathbf{u} \in (\mathbb{D}_L)^\mathbb{Z}$ for which $\text{Q-gen}(\mathbf{u}) \subset \text{conv}(\kappa_1, \dots, \kappa_m)$, where $((\mathbb{D}_L)^\mathbb{Z}, \kappa_j, S)$ for $j = 1, \dots, m$ is measure-theoretically isomorphic to $((\mathbb{D}_L)^\mathbb{Z}, \kappa, S)$, the system (X, T) does not satisfy the strong MOMO property (relative to \mathbf{u}).⁵¹

Remark 4.6. Let us come back to Theorem 3.4 and Remark 3.6, i.e. to the reformulation of Sarnak's conjecture using completely deterministic sequences. We intend to show that a natural generalization of Corollary 4.5 (b) to the completely deterministic case fails. Indeed, consider the square-free system (X_{μ^2}, S) . In Remark 3.3, we have already noticed that whenever $k_j, j = 1, \dots, r$ are different non-negative integers, then

$$(*) \quad \sum_{n \leq N} \mu^2(n + k_1) \dots \mu^2(n + k_{r-1}) \mu(n + k_r) = o(N).$$

It follows that for each $f \in C(X_{\mu^2})$, for each $k \in \mathbb{Z}$, we have

$$(**) \quad \frac{1}{N} \sum_{n \leq N} f(S^{n+k} \mu^2) \mu(n) \rightarrow 0.$$

On the other hand, the convergence in (**) cannot be uniform in $k \in \mathbb{Z}$. Indeed, if it were then the whole square-free system would be Möbius disjoint. This is however impossible since (X_{μ^2}, S) is hereditary, see Remark 6.2. Indeed, we can find $y \in X_{\mu^2}$ such $y(n) = 1$ if and only if $\mu(n) = 1$ and $y(n) = 0$ otherwise (then $y \leq \mu^2$) and if we set $\theta(z) := z(0)$ then $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \theta(S^n y) \mu(n) = \frac{3}{\pi^2}$.

See also [139], where a quantitative version of (*) has been proved.

Note that Theorem 4.3 does not fully answer Question 2. In certain situations the following general (lifting) lemma of Downarowicz and Lemańczyk can be helpful:

Lemma 4.7 ([2, 51]). *Assume that an ergodic automorphism R is coalescent.⁵² Let (\tilde{X}, \tilde{T}) and (X, T) be uniquely ergodic models of R . Assume that T is a topological factor of \tilde{T} , i.e. there exists $\pi: \tilde{X} \rightarrow X$ which is continuous and onto and which satisfies $\pi \circ \tilde{T} = T \circ \pi$. If T is Möbius disjoint then also \tilde{T} is Möbius disjoint.*

⁵⁰It is not hard to see that the MOMO property implies the relevant uniform convergence. As a matter of fact, the strong MOMO property is equivalent to the uniform convergence (in x , for a fixed $f \in C(X)$) on short intervals: $\frac{1}{M} \sum_{1 \leq m < M} \left| \frac{1}{H} \sum_{m \leq h < m+H} f(T^h x) \mu(n) \right| \rightarrow 0$ (when $H, M \rightarrow \infty$ and $H = o(M)$). It follows that we have equivalence of: Sarnak's conjecture (2), Sarnak's conjecture in its uniform form, Sarnak's conjecture in its short interval uniform form and the strong MOMO property. Moreover, each of these conditions is implied by the Chowla conjecture.

⁵¹This result means that there must be an observable sequence in (X, T) which significantly correlates with \mathbf{u} .

⁵²This means that each measure-preserving transformation commuting with R must be invertible. Finite multiplicity of the Koopman operator associated to R guarantees coalescence. In particular, all ergodic rotations are coalescent.

4.2 Möbius disjointness and entropy

Sarnak’s conjecture deals with deterministic systems but Möbius disjointness, a priori, does not exclude the possibility of positive (topological) entropy systems which are Möbius disjoint.⁵³ The first “natural” trial would be to take the square-free system (X_{μ^2}, S) which has positive entropy (see Section 6.2) and clearly μ^2 is orthogonal to μ . However, in spite of the orthogonality of the two sequences, as we have noticed in Remark 4.6, the square-free system is not Möbius disjoint.

Recently, Downarowicz and Serafin [52] constructed Möbius disjoint positive entropy homeomorphisms of arbitrarily large entropy. On the other hand, see [102], in the subshift of finite type case we do not have Möbius disjointness. Using Katok’s horseshoe theorem, it follows that $C^{1+\delta}$ -diffeomorphisms of surfaces are not Möbius disjoint but the following question seems to be open:

Question 3. Is there a positive entropy diffeomorphism of a compact manifold which is Möbius disjoint?

Viewed all this above, another natural question arises:

Question 4. Does there exist an ergodic positive entropy measure-theoretic system all uniquely ergodic models of which are Möbius disjoint?

Using Theorem 4.3, Sinai’s theorem on Bernoulli factors (see e.g. [75]) and B. Weiss’ theorem [167] on strictly ergodic models of some diagrams a partial answer to Question 4 is given by the following result:

Corollary 4.8 ([4]). Assume that $\mathbf{u} \in (\mathbb{D}_L)^{\mathbb{Z}}$ is generic for a Bernoulli measure κ . Let $\mathbf{v} \in (\mathbb{D}_L)^{\mathbb{Z}}$, \mathbf{u} and \mathbf{v} correlate. Then for each dynamical system (X, T) with $h(X, T) > h((\mathbb{D}_L)^{\mathbb{Z}}, \kappa, S)$, we do not have the strong MOMO property relatively to \mathbf{v} .

By substituting $\mathbf{u} = \boldsymbol{\lambda}$, $\mathbf{v} = \boldsymbol{\mu}$ and assuming the Chowla conjecture for $\boldsymbol{\lambda}$, we obtain that no system (X, T) with entropy $> \log 2$ satisfies the strong MOMO relatively to $\boldsymbol{\mu}$. When $\boldsymbol{\mu}$ is replaced by $\boldsymbol{\lambda}$, we still have a stronger result.

Proposition 4.9 ([4]). Assume that the Chowla conjecture holds for $\boldsymbol{\lambda}$. Then no topological system (X, T) with positive entropy satisfies the strong MOMO property relatively to $\boldsymbol{\lambda}$.

Remark 4.10. The proof of Theorem 4.3 tells us that when $(Z, \mathcal{D}, \kappa, R)$ is ergodic and has positive entropy then there exists a system (X, T) , which is not Liouville disjoint, with at most three ergodic measures and all of these measures yield a measurable system isomorphic to R . Therefore, it seems reasonable to conjecture that the answer to Question 4 is negative.

We now have a completely clear picture for the Liouville function: it follows from Theorem 3.4 (for $\boldsymbol{\lambda}$) and Proposition 4.9 that if the Chowla conjecture holds for $\boldsymbol{\lambda}$ then the strong MOMO property (relatively to $\boldsymbol{\lambda}$) holds for (X, T) if and only if $h(X, T) = 0$. Using footnote 50, we immediately obtain Proposition 4.9 in its equivalent form:

⁵³Sarnak in [148] mentions that Bourgain has constructed a positive entropy system which is Möbius disjoint but this construction has never been published.

Corollary 4.11. Assume that the Chowla conjecture holds for λ . Then, the short interval uniform convergence in (1) (with μ replaced by λ) takes place if and only if $h(X, T) = 0$.

4.3 The AOP property and its consequences

We need an ergodic criterion to establish the strong MOMO property in models of an automorphism. This turns out to be a natural ergodic counterpart of the KBSZ criterion (Theorem 2.13). Following [7] an ergodic automorphism R is said to have *asymptotically orthogonal powers* (AOP) if for each $f, g \in L^2_0(Z, \mathcal{D}, \kappa)$, we have

$$(40) \quad \lim_{\mathbb{P} \ni p, q \rightarrow \infty, p \neq q} \sup_{\kappa \in J^e(R^p, R^q)} \left| \int_{X \times X} f \otimes g d\kappa \right| = 0.$$

Rotation $Rx = x + 1$ acting on $\mathbb{Z}/k\mathbb{Z}$ with $k \geq 2$ has no AOP property because of Dirichlet's theorem on primes in arithmetic progressions. Hence, AOP implies total ergodicity (clearly, AOP is closed under taking factors). The AOP property implies zero entropy [7].

Clearly, if the powers of R are pairwise disjoint⁵⁴ then R enjoys the AOP property. In order to see a less trivial example of an AOP automorphism, consider any totally ergodic discrete spectrum automorphism R on (Z, \mathcal{D}, κ) . For f, g take eigenfunctions corresponding to eigenvalues c, d , respectively. Now, take $\rho \in J^e(R^p, R^q)$ and consider

$$\int_{Z \times Z} f \otimes g d\rho = \int_{Z \times Z} (f \otimes \mathbb{1}_Z) \cdot (\mathbb{1}_Z \otimes g) d\rho.$$

Notice that $f \otimes \mathbb{1}_Z$ and $\mathbb{1}_Z \otimes g$ are eigenfunctions of $(Z \times Z, \rho, R^p \times R^q)$ corresponding to c^p and d^q , respectively. If $c^p \neq d^q$ (and this is the case for all but one pair (p, q) because of total ergodicity) then these eigenfunctions are orthogonal and we are done. We will see more examples in Section 5.

Remark 4.12. For an AOP automorphism the powers need not be disjoint. As a matter of fact, we can have an AOP automorphism with all of its non-zero powers isomorphic.⁵⁵

Theorem 4.13 ([3, 7]). *Let $\mathbf{u} \in \mathcal{M}$. Suppose that $(Z, \mathcal{D}, \kappa, R)$ satisfies AOP. Then the following are equivalent:*

- \mathbf{u} satisfies (39);
- The strong MOMO property relatively to \mathbf{u} is satisfied in each uniquely ergodic model (X, T) of R .

In particular, if the above holds, for each $f \in C(X)$, we have

$$\frac{1}{N} \sum_{n \leq N} f(T^n x) \mathbf{u}(n) \xrightarrow{N \rightarrow \infty} 0 \text{ uniformly in } X.$$

⁵⁴This is a "typical" property of an automorphism of a probability standard Borel space [98]. Möbius disjointness for uniquely ergodic models for this case is already noticed in [31].

⁵⁵Take an ergodic rotation with the group of eigenvalues $\{e^{2\pi i \alpha m/n} : m, n \in \mathbb{Z}, n \neq 0, \alpha \notin \mathbb{Q}\}$.

Corollary 4.14. Assume that $(Z, \mathcal{D}, \kappa, R)$ enjoys the AOP property. Then, in each uniquely ergodic model (X, T) of R , we have

$$(41) \quad \frac{1}{M} \sum_{M \leq m < 2M} \left| \frac{1}{H} \sum_{m \leq h < m+H} f(T^n x) \mu(n) \right| \xrightarrow{H, M \rightarrow \infty, H = o(M)} 0$$

for all $f \in C(X)$, $x \in X$.

The AOP property can be defined for actions of locally compact (second countable) groups. Then, for induced actions this property lifts [64], and in particular (by taking the induced \mathbb{R} -action), if we have an automorphism then the corresponding suspension flow⁵⁶ has this lifted property. In particular, using induced \mathbb{Z} -actions (for $a\mathbb{Z} \subset \mathbb{Z}$), one can derive easily that for uniquely ergodic systems (X, T) with the measure-theoretic AOP property we not only have Möbius disjointness but also

$$(42) \quad \frac{1}{N} \sum_{n \leq N} f(T^n x) \mu(an + b) \xrightarrow{N \rightarrow \infty} 0$$

for each $a, b \in \mathbb{N}$, $f \in C(X)$ and the convergence is uniform in x [64].⁵⁷

5 Glimpses of results on Sarnak's conjecture

The cases for which the Möbius disjointness has been proved, depend on the complexity of the deterministic system. They fit into two basic types. The first comes with sufficiently quantitative estimates for the disjointness sums which makes possible an analysis of the sums on primes yielding a PNT. This group includes Kronecker systems (Vinogradov [162]), nilsystems (Green and Tao [80]) and, perhaps the most striking, the Thue-Morse system (Mauduit and Rivat [127]) which resolved a conjecture of Gelfond [74]. When the systems are more complex, such as horocycles flows,⁵⁸ then at least to date they do not come with a PNT,⁵⁹ and for them the KBSZ criterion is used, in other words, the disjointness (perhaps in its weaker form, see Section 4) is achieved.

We now review most of important cases in which Möbius disjointness has been proved.

5.1 Systems of algebraic origin

5.1.1 Horocycle flows

Let $\Gamma \subset PSL_2(\mathbb{R})$ be a discrete subgroup with finite covolume.⁶⁰ Then the homogeneous space $X = \Gamma \backslash PSL_2(\mathbb{R})$ is the unit tangent bundle of a surface

⁵⁶By the *suspension flow* of R we mean the special flow over R under the constant function (equal to 1).

⁵⁷The same argument shows that if Sarnak's conjecture holds then (42) holds for each zero entropy (X, T) , $a, b \in \mathbb{N}$, $f \in C(X)$ uniformly in $x \in X$.

⁵⁸Horocycle flows are mixing of all orders, see [121].

⁵⁹In case of horocycle flows (Bourgain, Sarnak and Ziegler [31]) Ratner's theorems on joinings are used and these provide no rate.

⁶⁰We will tacitly assume that Γ is cocompact, so that the homogenous space $\Gamma \backslash PSL_2(\mathbb{R})$ is compact and the system is uniquely ergodic by [71]; otherwise, as in the modular case when $\Gamma = PSL_2(\mathbb{Z})$ we need to compactify our space. The proof of Theorem 5.1 in the modular case is slightly different than what we describe below.

M of constant negative curvature. Let us consider the corresponding *horocycle flow*⁶¹ $(h_t)_{t \in \mathbb{R}}$ and the *geodesic flow* $(g_s)_{s \in \mathbb{R}}$ on X . Since

$$(43) \quad g_s h_t g_s^{-1} = h_{e^{-2s}t} \text{ for all } s, t \in \mathbb{R},$$

the flows $(h_t)_{t \in \mathbb{R}}$ and $(h_{e^{-2s}t})_{t \in \mathbb{R}}$ are measure-theoretically isomorphic for each $s \in \mathbb{R}$. In order to show that $T := h_1$ is Möbius disjoint, the KBSZ criterion is used, and, given $x \in PSL_2(\mathbb{R})$, one studies limit points of $\frac{1}{N} \sum_{n \leq N} \delta_{(T^{pn}\Gamma x, T^{qn}\Gamma x)}$, $N \geq 1$. Now, the celebrated Ratner's rigidity theorem [143] tells us two important things: the point $(\Gamma x, \Gamma x)$ is generic for a measure ρ (which must be a joining by unique ergodicity: $\rho \in J(T^p, T^q)$) and moreover this joining is ergodic.⁶² Again using Ratner's theory (cf. [142]) such joinings are determined by the commensurator $Com(\Gamma)$ of the lattice Γ :

$$Com(\Gamma) := \{z \in PSL_2(\mathbb{R}) : z^{-1}\Gamma z \cap \Gamma \text{ has finite index in both } \Gamma \text{ and } z^{-1}\Gamma z\}.$$

Set $x_{p,q} := x g_{\frac{1}{2} \log(\frac{p}{q})} x^{-1}(\infty)$. The intersection of the stabilizer of $x_{p,q}$ with $Com(\Gamma)$ yields the correlator of $x_{p,q}$: it is a subgroup $C(\Gamma, x_{p,q}) \subset \mathbb{R}_+^*$ and if ρ is not the product measure then $\frac{p}{q} \in C(\Gamma, x_{p,q})$. The careful analysis of the arithmetic and non-arithmetic cases done in [31] shows that given $x \in PSL_2(\mathbb{R})$, $\frac{p}{q} \in C(\Gamma, x_{p,q})$ only for finitely many different primes p, q . Hence, the joining ρ has to be product measure for all but finitely many pairs $(p, q) \in \mathbb{P}^2$ with $p \neq q$ which, by Theorem 2.13, yields the following:

Theorem 5.1 ([31]). *All time-automorphisms of horocycle flows are Möbius disjoint.*

Remark 5.2. As noticed in [6], this is (43) which yields the absence⁶³ of AOP and makes the following questions of interest.

Question 5. Do we have the MOMO property for horocycle flows? Are all uniquely ergodic models of horocycle flows Möbius disjoint? Do we have uniform convergence in (1)?

Since the method to prove Möbius disjointness is through the KBSZ criterion (hence offers no rate of convergence), the following question is still open:

Question 6 (Sarnak). Do we have a PNT for horocycle flows?

For a partial answer, see [150], where it is proved that if Γx is a generic point for Haar measure μ_X of X then any limit point of $\left(\frac{1}{\pi(N)} \sum_{p \leq N} \delta_{T^p \Gamma x}\right)$ is a measure which is absolutely continuous with respect to μ_X .

Question 7 (Ratner). Are smooth time changes for horocycle flows Möbius disjoint?

As smooth time changes of horocycle flows enjoy so called Ratner's property, the above question can be asked in the larger context of flows possessing Ratner's property.

⁶¹We have $h_t(\Gamma x) = \Gamma \cdot \left(x \cdot \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}\right)$ and $g_s(\Gamma x) = \Gamma \cdot \left(x \cdot \begin{bmatrix} e^{-s} & 0 \\ 0 & e^s \end{bmatrix}\right)$; we identify g_s and h_t with the relevant matrices.

⁶²The measure ρ depends on p, q and x and it is so called algebraic measure, i.e. a Haar measure.

⁶³To be compared with Remark 4.12; the difference however is that when the ratio of p and q is close to 1, we can choose graph joinings in a compact set.

Added in September 2017: In the recent paper [100], a new criterion (of Ratner’s type) for disjointness of different time-automorphisms of flows has been proved. The criterion applies for some classes of flows with Ratner’s property, namely, in case of so called Arnold flows and for non-trivial smooth time changes of horocycle flows (in particular, the answer to Question 7 is positive).

5.1.2 Nilrotations, affine automorphisms

Green and Tao in [80] proved Möbius disjointnes in the following strong form:

Theorem 5.3 ([80]). *Let G be a simply-connected nilpotent Lie group with a discrete and cocompact subgroup Γ . Let $p: \mathbb{Z} \rightarrow G$ be any is polynomial sequence⁶⁴ and $f: G/\Gamma \rightarrow \mathbb{R}$ a Lipschitz function. Then*

$$\left| \frac{1}{N} \sum_{n \leq N} f(p(n)\Gamma) \boldsymbol{\mu}(n) \right| = O_{f,G,\Gamma,A} \left(\frac{N}{\log^A N} \right)$$

for all $A > 0$.

In particular, by considering $T_g(x\Gamma) = gx\Gamma$, we see that all nilrotations are Möbius disjoint with uniform Davenport’s estimate (3).

Also, a PNT holds for nilrotations: Let $2 = p_1 < p_2 < \dots$ denote the sequence of primes.

Theorem 5.4 ([80], Theorem 7.1). *Assume that a nil-rotation T_g is ergodic.⁶⁵ Then, for every $x \in G$, we have*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} f(T_g^{p_n} x\Gamma) = \int_{G/\Gamma} f d\lambda_{G/\Gamma}$$

for all continuous functions $f: G/\Gamma \rightarrow [-1, 1]$.

In [64], it is proved that all nil-rotations enjoy the AOP property (hence all uniquely ergodic models of nil-rotations are Möbius disjoint). In fact, the result is proved for all nil-affine automorphisms whose Möbius disjointness has been established earlier in [120]. Earlier, AOP has been proved for all quasi-discrete spectrum automorphism in [7], that is (following [82]) for all unipotent affine automorphisms $Tx = Ax + b$ of compact Abelian groups (A is a continuous group automorphism and b is an element of the group). The Möbius disjointness of the latter automorphisms has been established still earlier in [120].

The proof of the following corollary in [7] shows that Furstenberg’s proof [69] (see e.g. [55]) of Weyl’s uniform distribution theorem can be adapted to the short interval version.

Corollary 5.5 ([7]). *Assume that $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$, $\mathbf{u} \in \mathcal{M}$. Then, for each non constant polynomial $P \in \mathbb{R}[x]$ with irrational leading coefficient, we have*

$$\frac{1}{M} \sum_{M \leq m < 2M} \frac{1}{H} \left| \sum_{m \leq n < m+H} e^{2\pi i P(n)} \mathbf{u}(n) \right| \xrightarrow{H, M \rightarrow \infty, H=o(M)} 0.⁶⁶$$

⁶⁴I.e. $p(n) = a_1^{p_1(n)} \dots a_k^{p_k(n)}$, where $p_j: \mathbb{N} \rightarrow \mathbb{N}$ is a polynomial, $j = 1, \dots, k$. See, Section 6 in [81] for the equivalence with the classical definition of polynomials sequences in nilpotent Lie groups.

⁶⁵We assume that G is connected.

⁶⁶For degree 1 polynomials, the result is already in [125].

Recall that a sequence $(a_n) \subset \mathbb{C}$ is called a *nilsequence* if it is a uniform limit of basic nilsequences, i.e. of sequences of the form $(f(T_g^n x \Gamma))$, where $f \in C(G/\Gamma)$ (here, we do not assume that G/Γ is connected, neither that T_g is ergodic).

Corollary 5.6 ([64]). We have

$$\frac{1}{M} \sum_{M \leq m < 2M} \frac{1}{H} \left| \sum_{m \leq n < m+H} a_n \mathbf{u}(n) \right| \xrightarrow{H, M \rightarrow \infty, H=O(M)} 0.$$

It has been proved by Leibman [117] that all polynomial multicorrelation sequences⁶⁷ are limits in the Weyl pseudo-metric of nil-sequences, all such polynomial sequences are orthogonal to $\boldsymbol{\mu}$ on typical short interval, cf. Section 6.

The main problem connected with nilsequences is to prove the uniform version of convergence on short intervals as it is made precise in Conjecture C of Tao (see Section 3.3 and also Frantzikinakis' proofs [66]).

5.1.3 Other algebraic systems

For a more general zero entropy algebraic systems and their Möbius disjointness we refer the reader to [135], where in particular the Ad-unipotent translation case is treated.

5.2 Systems of measure-theoretic origin. Substitutions and interval exchange transformations

5.2.1 Systems whose powers are disjoint

We are interested in ergodic automorphisms $(Z, \mathcal{D}, \kappa, R)$ for which (sufficiently large) prime powers R^p are pairwise disjoint. Clearly, such automorphisms enjoy the AOP property. A typical automorphism has this property [98] but there are also large classes of rank one (we detail on this class below) automorphisms with this property [5, 30, 146]. Also minimal self-joining automorphisms [97] enjoy this property. Chaika and Eskin in [35] show that for a.e. 3-interval exchange transformation (we detail on interval exchange transformations below) there are sufficiently many prime powers that are disjoint. It follows that all uniquely ergodic models of these automorphisms are Möbius disjoint.

5.2.2 Adic systems and Bourgain's criterion

Let $(Z, \mathcal{D}, \kappa, R)$ be a measure-theoretic system.

Definition 5.1. In $(Z, \mathcal{D}, \kappa, R)$, a Rokhlin *tower* is a collection of disjoint measurable sets called *levels* $F, RF, \dots, R^{h-1}F$. If Z is equipped with a partition P such that each level $R^r F$ is contained in one atom $P_{w(r)}$, the *name* of the tower is the word $w(0) \dots w(h-1)$.

⁶⁷More precisely, given an automorphism T of a probability standard Borel space (X, \mathcal{B}, μ) , we consider

$$a_n = \int_X g_1 \circ T^{p_1(n)} \dots \cdot g_k \circ T^{p_k(n)} d\mu,$$

where $g_i \in L^\infty(X, \mu)$, $p_i \in \mathbb{Z}[x]$, $i = 1, \dots, k$ ($k \geq 1$).

Definition 5.2. A system $(Z, \mathcal{D}, \kappa, R)$ is of *rank one* if there exists a sequence of Rokhlin towers $(F_n, \dots, R^{h_n-1}F_n)$, $n \geq 1$, such that the whole σ -algebra is generated by the partitions $\{F_n, RF_n, \dots, R^{h_n-1}F_n, X \setminus \bigcup_{j=0}^{h_n-1} R^j F_n\}$.

For topological systems, there is no canonical notion of rank, but the useful notion is that of adic presentation [161], which we translate here from the original vocabulary into the one of Rokhlin towers.

Definition 5.3. An *adic presentation* of a topological system (X, T) is given, for each $n \geq 0$, by a finite collection \mathcal{Z}_n of Rokhlin towers such that:

- the levels of the towers in \mathcal{Z}_n partition X ,
- each level of a tower in \mathcal{Z}_n is a union of levels of towers in \mathcal{Z}_{n+1} ,
- the levels of the towers in $\bigcup_{n \geq 0} \mathcal{Z}_n$ form a basis of the topology of X .

In that case, the towers of \mathcal{Z}_{n+1} are built from the towers of \mathcal{Z}_n by cutting and stacking, following recursion rules: a given tower in \mathcal{Z}_{n+1} can be built by taking columns of successive towers in \mathcal{Z}_n and stacking them successively one above another. These rules are best seen by looking at the partition P into levels of the towers in \mathcal{Z}_0 ; possibly replacing Z_0 by some Z_k , we can always assume P has at least two atoms. The names of the towers in \mathcal{Z}_n form sets of words \mathcal{W}_n , and the cutting and stacking of towers gives a canonical decomposition of every $W \in \mathcal{W}_n$:

$$W = W_1^{k_1} \dots W_r^{k_r}$$

for r words $W_i \in \mathcal{W}_{n-1}$, $1 \leq i \leq r$, integers k_1, \dots, k_r ; all these parameters depend on the word W . These decompositions are called the *rules of cutting and stacking* of the system.

The following result is an improvement on Theorem 3.1 of [63], which itself can be found in [30], though it is not completely explicit in that paper (it is stated in full only in a particular case, as Theorem 3, and its proof is understated). The following effective bound stems from a closer reading of [30]:

Theorem 5.7. *Let (X, T) be a topological dynamical system admitting an adic presentation, as in Definition 5.3 and the comment just after.*

Suppose that for any n and W in \mathcal{W}_n , we have:

- *in the rules of cutting and stacking $r \leq C$, with $C \geq 2$,*
- *if we decompose W into words $W_\ell \in \mathcal{W}_{n-s}$ by iteration of the rules of cutting and stacking then for all ℓ and s large enough, we have*

$$|W| > C^{200s} |W_\ell|.$$

Then (X, T) is Möbius disjoint.

If such a system is uniquely ergodic and weakly mixing for its invariant probability, it satisfies also the following PNT: for any word $W = w_1 \dots w_N$ which is a factor of a word in any \mathcal{W}_n , we have

$$\sum_{i=1}^N \Lambda(i) w_i = \sum_{i=1}^N w_i + o(N).$$

Proof. We look at Theorem 2 of [30]. It requires a stronger assumption, denoted by relations (2.2) and (2.3) in p. 119 of [30], which is indeed the assumption of the present theorem with the estimate C^{200s} replaced by $\beta(s)$ for some function satisfying $\frac{\log \beta(s)}{s} \rightarrow \infty$ when $s \rightarrow \infty$ (note that the assumption in [30] that the words W_n are on the alphabet $\{0, 1\}$ is not used in the proof, which works for any finite alphabet). Then this theorem gives, for any word $w_1 \cdots w_N$ in some \mathcal{W}_m and N large enough, an estimate for

$$\int \left| \sum_{n \leq N} w_n e^{2\pi i n \theta} \right| \left| \sum_{n \leq N} \mu(n) e^{2\pi i n \theta} \right| d\theta,$$

and this, through the relation (1.62) on p. 118, implies that $\sum_{n \leq N} w_n \mu(n) = o(N)$.

Lacking space to rewrite the extensive computations in [30], we explain how to weaken the hypothesis. First, as suggested in the remark at the beginning of Section 2, p. 119, of that paper, we replace $\beta(s)$ by C_0^s for some constant C_0 , as yet unknown (the C_0 s written in the same p. 119 is a misprint). The relations (2.2) and (2.3) are used twice in the course of the proof: first, to get the relation (2.15), namely

$$\left(C \frac{\log |W|}{n} \right)^s < |W|^\epsilon$$

for a word W in \mathcal{W}_n , and then to get the estimate (2.42), which states that

$$\left(C \frac{\log K}{s} \right)^s < K^\epsilon,$$

where s is the number of stages such that a word of length N in \mathcal{W}_n is divided into words of \mathcal{W}_{n-s} , of lengths in the order of $\frac{N}{K}$. Under our hypothesis, in the first case, $|W|$ is in C_0^n , and in the second case K is in C_0^s . Thus both (2.15) and (2.42) are implied by the relation

$$\frac{\log \log C_0 + \log C}{\log C_0} < \epsilon.$$

The value of ϵ is dictated by relation (2.49), which requires $Q^\epsilon K^\epsilon (Q + K)^{-\frac{1}{4}} \leq (Q + K)^{-\frac{1}{5}}$ for some large numbers Q and K , thus we can take $\epsilon = \frac{1}{20}$. Then $\frac{\log \log C_0}{\log C_0}$ will be bounded if C_0 is large enough independently of C , while to bound $\frac{\log C}{\log C_0}$ we need to take $C_0 = C^a$; as $C \geq 2$, we see that $a = 200$ is convenient for the sum of the two terms.

Now, if we replace w_n by $u(n) = f(T^n(x_0))$, because of Definition 5.3 above, we can first assume that f is constant on all levels of the towers of some stage m , and then conclude by approximation. Such an f is also constant on all levels of all towers at stages $q > m$; fixing x_0 and N , except for some initial values $u(1)$ to $u(N_0)$ where N_0 is much smaller than N , we can replace $u(n)$ by w'_n , where w'_n is the value of f on the n -th level of some tower with name W in some \mathcal{W}_q for $q \geq m$. Then the $w'_1 \cdots w'_N$ are built by the same induction rules as the $w_1 \cdots w_N$, and the estimates using the w'_n are computed as those using the w_n in the proof of Theorem 2 of [30], thus we get the same result.

The PNT is in (3.4), (3.7), (3.14) of [30] ((3.14) is proved for the particular case of 3-interval exchanges but holds in the same way for the more general case). \square

Of course, the value of C_0 could be improved, but we need it to be at least some power of C .

5.2.3 Substitutions

We start with some basic notions.

Definition 5.4. A *substitution* σ is an application from an alphabet \mathcal{A} into the set \mathcal{A}^* of finite words on \mathcal{A} ; it extends to a morphism of \mathcal{A}^* for the concatenation. A *fixed point* of σ is an infinite sequence u with $\sigma u = u$. The associated symbolic dynamical system (X_σ, S) is (X_u, S) for a fixed point u .

Substitution σ has *constant length* q if $|\sigma a| = q$ for all a in \mathcal{A} .

The *Perron-Frobenius eigenvalue* is the largest eigenvalue of the matrix giving the number of occurrences of j in σa . A substitution σ is *primitive* if a power of this matrix has strictly positive entries.

For the class of constant length substitutions, there have been a lot of partial results on Möbius orthogonality:

- First for the most famous example, the Thue-Morse substitution $0 \rightarrow 01, 1 \rightarrow 10$, with Indlekofer and Kátai [94], Dartyge and Tenenbaum [41], Mauduit and Rivat [127], El Abdalaoui, Kasjan and Lemańczyk [2].⁶⁸
- The case of the Rudin - Shapiro substitution $0 \rightarrow 01, 1 \rightarrow 02, 2 \rightarrow 31, 3 \rightarrow 32$ was solved by Mauduit and Rivat [128]. Then Drmota [53], Deshouillers, Drmota and Müllner [48],
- Ferenczi, Kułaga-Przymus, Lemańczyk and Mauduit [60] extended these results to various subclasses of systems of arithmetic origin.⁶⁹

See also [122, 123] for a PNT for some digital functions.

But all this was superseded by the general result of Müllner [134], whose proof uses the arithmetic techniques of [128] together with a new structure theorem on the underlying *automata*:

Theorem 5.8 ([134]). *For any substitution of constant length, the associated symbolic system is Möbius disjoint. Moreover, a PNT holds if the substitution is primitive.*

The substitutions which are not of constant length are much less known:

- The most famous example is the Fibonacci substitution, $0 \rightarrow 01, 1 \rightarrow 0$: in that case, the associated symbolic system is a coding of an irrational rotation, hence it is Möbius disjoint as a uniquely ergodic model of a discrete spectrum automorphism, see Section 5.3.1.
- Drmota, Müllner and Spiegelhofer [54] have just shown Möbius disjointness for a new example, a substitution which generates $(-1)^{s_\phi(n)}$, where $s_\phi(n)$ is the Zeckendorf sum-of-digits function.⁷⁰

⁶⁸In [41, 94, 127] it is proved that the sequence $(-1)^{u(n)}$, $n \geq 1$ is orthogonal to μ .

⁶⁹While in [53, 60] Möbius disjointness is proved for the dynamical systems given by bijective substitutions, [48] treats the opposite case, so called synchronized. As noted in [21], this leads to dynamical systems given by rational sequences and such are Möbius disjoint. Note also that for the synchronized case, once the system is uniquely ergodic, it is automatically a uniquely ergodic model of an automorphism with discrete spectrum, cf. Corollary 3.20 and Remark 4.4.

⁷⁰This example has partly continuous spectrum.

- Also, we can exhibit a small subclass of examples which are Möbius disjoint, by a straightforward translation of Bourgain’s criterion above:

Theorem 5.9. *Suppose that σ is a primitive substitution satisfying*

- *for all $i \in \mathcal{A}$, $\sigma i = (j_1(i))^{a_1(i)} \dots (j_{q_i}(i))^{a_{q_i}(i)}$, $a_1(i) \in \mathcal{A}, \dots, a_{q_i}(i) \in \mathcal{A}$, $q_i \leq C$ (this can be expressed as: the multiplicative length of σ is smaller than C),*
- *the Perron-Frobenius eigenvalue of σ is larger than C^{200} ;*

then the associated symbolic dynamical system is Möbius disjoint. If (X_σ, S) is weakly mixing, the fixed points satisfy a PNT.

Proof. If all fixed points are periodic, the result is trivial. If σ has a non-periodic fixed point, it is well known (and proved by the methods of [138] together with the recognizability result of [133]) that the system has an adic presentation, where the names of the towers in Z_n are the words $\sigma^n a$, $a \in \mathcal{A}$. Thus the results come from Theorem 5.7 above and the properties of the matrix of σ . \square

Example 5.1. Here are some substitutions for which the above theorem applies, with a PNT: $0 \rightarrow 0^{k+1}12, 1 \rightarrow 12, 2 \rightarrow 0^k12, k + 2 > 3^{200}$.

Question 8. Are dynamical systems associated to substitutions Möbius disjoint?⁷¹

5.2.4 Interval exchanges

Definition 5.5. A k -interval exchange with probability vector $(\alpha_1, \alpha_2, \dots, \alpha_k)$, and permutation π is defined by

$$Tx = x + \sum_{\pi^{-1}(j) < \pi^{-1}(i)} \alpha_j - \sum_{j < i} \alpha_j.$$

when $x \in \Delta_i = \left[\sum_{j < i} \alpha_j, \sum_{j \leq i} \alpha_j \right)$.

Exchanges of 2 intervals are just rotations, thus Möbius disjointness holds for them by the Prime Number Theorem (on arithmetic progressions) when the rotation is rational and from a result of Davenport [43] – using a result of Vinogradov [163] – when the rotation is irrational, cf. (3) in Introduction.

Then [30] exhibits exchanges of 3 intervals which are Möbius disjoint, with a PNT if weak mixing holds: these use the criterion developed in Theorem 5.7 above, together with the adic presentation built in [62]. Generalizing these methods, it is shown in [63] that Möbius disjointness holds for examples of exchanges of k intervals for every $k \geq 2$ and every *Rauzy class*, with a PNT in the weak mixing case. A breakthrough came with [35], for a large subclass of exchange of 3 intervals:

Theorem 5.10 ([35]). *For (Lebesgue)-almost all (α_1, α_2) , Sarnak’s conjecture holds for exchanges of 3 intervals with permutation $\pi i = 3 - i$ and probability vector $(\alpha_1, \alpha_2, 1 - \alpha_1 - \alpha_2)$.*

⁷¹One can also ask about Möbius disjointness of related systems as tiling systems.

To prove this, Chaika and Eskin use first the well-known fact that such an exchange of 3 intervals, denoted by T , is the induced map of the rotation of angle $\alpha = \frac{1-\alpha_1}{1+\alpha_2}$ on the interval $[0, x)$ where $x = \frac{1}{1+\alpha_2}$. This approach, of course, does not generalize to 4 intervals or more.

In fact, in [35] two different results are proved. In the easier one, they deduce Möbius disjointness from the disjointness of powers of T ; they give a sufficient condition for T^m to be disjoint from T^n for all $m \neq n$, which is satisfied by almost all these T . Namely, if we take (a_1, \dots) to be the continued fraction of α and (b_1, \dots) the α -Ostrowski expansion of x , then it is enough that, for any ordered k -tuple of pairs $((c_1, d_1), \dots, (c_k, d_k))$ of natural numbers such that $d_i \leq c_i - 1$, there are infinitely many i with $a_i = c_1, \dots, a_{i+k-1} = c_k, b_i = d_1, \dots, b_{i+k-1} = d_k$.

Then most of the paper is used to give an explicit Diophantine condition on α and x , which implies a slightly weaker property than the disjointness of powers. Under that condition, there exists a constant C such that for all n , and $0 \leq m \leq n$, T^m is disjoint from T^n except maybe when m belongs to a sequence $m_i(n)$ in which any two consecutive terms satisfy $m_{i+1}(n) > C m_i(n)$, and this is proved to imply Möbius disjointness. The Diophantine condition holds for almost all T , and, as it is long, we refer the reader to Theorem 1.4 of [35]; it expresses the fact that the geodesic ray from a certain flat torus with two marked points, defined naturally from T and its inducing rotation, spends significant time in compact subsets of the space of such tori.

5.2.5 Systems of rank one

These systems form a measure-theoretic class defined in Definition 5.2 above. It is well known, but has been shown explicitly for all cases only in the recent [9], that each system of rank-one is measure-theoretically isomorphic to one of the topological systems we define now.

Definition 5.6. A *standard model of rank one* is the shift on the orbit closure of the sequence u which, for each $n \geq 0$, begins with the word B_n defined recursively by concatenation as follows. We take sequences of positive integers $q_n, n \geq 0$, with $q_n > 1$ for infinitely many n , and $a_{n,i}, n \geq 0, 0 \leq i \leq q_n - 1$, such that, if h_n are defined by $h_0 = 1, h_{n+1} = q_n h_n + \sum_{j=0}^{q_n-1} a_{n,i}$, then

$$\sum_{n=0}^{\infty} \frac{h_{n+1} - q_n h_n}{h_{n+1}} < \infty.$$

We define $B_0 = 0$,

$$B_{n+1} = B_n 1^{a_{n,0}} B_n \dots B_n 1^{a_{n,q_n-1}}$$

for $n \geq 0$.

In [30], Bourgain proved Möbius disjointness for a standard model of rank one if both the $q_n, n \in \mathbb{N} \cup \{0\}$ and $a_{n,i}, n \in \mathbb{N} \cup \{0\}$, are bounded by some constant C (we will refer to this as to a *bounded rank one construction*).

Note however that, in the same paper, the half-hidden criterion deduced from Theorem 2 or 3, see Theorem 5.7 above, is much more than an auxiliary to prove the supposedly main Theorem 1 of [30]; it applies to a much wider class of systems, and even for some famous rank one systems this criterion works while Theorem 1 does not apply.

Bourgain's result was improved in [5], where so called *recurrent rank one constructions* are considered with a stabilizing bounded subsequence of spacers (that is, of a subsequence of $(a_{n,i})$).⁷² One of main tools in [5] is a representation of each rank one transformation as an integral automorphism over an odometer with so called Morse-type roof function which goes back to [88]. See also [146] for a simpler proof of a generalization of Bourgain's result to a class of partially bounded rank one constructions.

Spectral approach In order to prove Möbius disjointness for standard models of rank one transformations, both papers [30] and [5] use a spectral approach. In [5], unitary operators U (of separable Hilbert spaces) are considered and weak limits of powers (U^{pm_k}) (for different primes p) are studied. Once such limits yield sufficiently different (for different p) analytic functions (of U), the powers U^p and U^q are spectrally disjoint.⁷³ If for a positive real number a we set $s_a(x) = ax \bmod 1$ on the additive circle $\mathbb{T} = [0, 1)$, then the above spectral disjointness means that

$$(44) \quad \sigma^{(p)} := (s_p)_*(\sigma) \text{ are mutually singular for different } p \in \mathbb{P},$$

where $\sigma = \sigma_U$ stands for the maximal spectral type of U .

In [30], a different spectral approach (sufficient for a use of the KBSZ criterion, hence, sufficient for Möbius disjointness) is used. Namely, if $r \geq 1$ is an integer, then by σ_r , we will denote the measure which is obtained first by taking the image of σ under the map $x \mapsto \frac{1}{r}x$, i.e. the measure $\sigma^{(1/r)}$, and then repeating this new measure periodically in intervals $[\frac{j}{r}, \frac{j+1}{r})$, that is:

$$\sigma_r := \frac{1}{r} \sum_{j=0}^{r-1} \sigma^{1/r} * \delta_{j/r}.$$

Bourgain [30] uses a representation of the maximal spectral type of a rank one transformation as a Riesz product and then shows the mutual disjointness of measures σ_p and σ_q for different $p, q \in \mathbb{P}$ (for more information about the measures σ_r , see e.g. [138], p. 196). Although, there seems not to be too much relation between the measures $\sigma^{(r)}$ and σ_r , the following observation⁷⁴ explains some equivalence of these both spectral approaches:

Lemma 5.11. *Assume that σ and η are two probability measures on the circle. Then:*

(a) *if $\sigma^{(r)} \perp \eta^{(s)}$ then $\sigma_s \perp \eta_r$;*

(b) *if $(r, s) = 1$ then $\sigma^{(r)} \perp \eta^{(s)}$ if and only if $\sigma_s \perp \eta_r$.*

⁷²Moreover, Möbius disjointness is established for some other famous classes of rank one transformations such as: Katok's α -weak mixing class (these are a special case of three interval exchange maps) or rigid generalized Chacon's maps.

⁷³Hence, T^p and T^q are disjoint in Furstenberg's sense, and, in fact, we even have AOP.

⁷⁴This has been proved, e.g. in an unpublished preprint of El Abadalaoui, Kułaga-Przymus, Lemańczyk and de la Rue.

5.2.6 Rokhlin extensions

Let T be a uniquely ergodic homeomorphism of a compact metric space and let $f: X \rightarrow \mathbb{R}$ be continuous. Set $T_f(x, t) := (Tx, f(x) + t)$ to obtain a skew product homeomorphism on $X \times \mathbb{R}$. Note that the latter space is not compact. But, if we take any continuous flow $\mathcal{S} = (S_t)_{t \in \mathbb{R}}$ acting on a compact metric space Y then the skew product $T_{f, \mathcal{S}}$ acting on $X \times Y$ by the formula:

$$T_{f, \mathcal{S}}(x, y) = (Tx, S_{f(x)}(y)), \quad (x, y) \in X \times Y$$

is a homeomorphism of the compact space $X \times Y$ and it is called a *Rokhlin extension* of T . To get a good theory, usually one has to put some further assumptions on f (considered as a cocycle taking values in a locally compact but not compact group, see e.g. [118, 151]). It is proved in [111] that there are irrational rotations $Tx = x + \alpha$ and continuous $f: \mathbb{T} \rightarrow \mathbb{R}$ (even smooth) such that $T_{\varphi, \mathcal{S}}$ has the AOP property for each uniquely ergodic \mathcal{S} .⁷⁵

We would like to emphasize that the Rokhlin skew product construction are usually relatively weakly mixing [118], so the class we consider here is drastically different from the distal class which is our next object to give account.

This approach leads in [111] to so called random sequences⁷⁶ $(a_n) \subset \mathbb{R}$ such that

$$\frac{1}{N} \sum_{n \leq N} g(S_{a_n} y) \mu(n) \rightarrow 0$$

for each uniquely ergodic flow \mathcal{S} acting on a compact metric space Y , each $g \in C(Y)$ and (due to [4]) uniformly in $y \in Y$.

5.3 Distal systems

Assume that R is an ergodic automorphism of a probability standard Borel space (Z, \mathcal{D}, κ) . R is called (measurably) *distal* if it can be represented as transfinite sequence of consecutive isometric extensions, where in case of a limit ordinal, we take the corresponding inverse limit (i.e. we start with the one-point dynamical system, the first isometric extension is a rotation and then we take a further isometric extension of it etc.). Recall that by a *separating sieve* we mean a sequence

$$Z \supset A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$$

of sets of positive measure such that $\mu(A_n) \rightarrow 0$ and there exists $Z_0 \subset Z$, $\mu(Z_0) = 1$, such that for each $z, z' \in Z_0$ if for each $n \geq 1$ there is $k_n \in \mathbb{Z}$ such that $R^{k_n} z, R^{k_n} z' \in A_n$, then $z = z'$. A theorem by Zimmer [171] says that T is distal if and only if it has a separating sieve.

Distal automorphisms play a special role in ergodic theory: each automorphism has a maximal distal factor and is relatively weakly mixing over it [73, 170, 171]. Hence, many problems in ergodic theory can be reduced to study the two opposite cases: the distal and the weak mixing one.⁷⁷ Recall that distal automorphisms have entropy zero.

⁷⁵If \mathcal{S} preserves a measure ν then $T_{\varphi, \mathcal{S}}$ preserves measure $\mu \otimes \nu$, the AOP property is considered with respect to this measure.

⁷⁶Such a sequence (a_n) is of the form $(\varphi^{(n)}(x))$ with $\varphi^{(n)}(x) = \varphi(x) + \varphi(Tx) + \dots + \varphi(T^{n-1}x)$, $n \geq 0$.

⁷⁷See the most prominent example of such a reduction, namely, Furstenberg's ergodic proof of Szemerédi theorem on the existence of arbitrarily long arithmetic progressions in subsets of integers of positive upper Banach density [73].

There is also a notion of distality in topological dynamics. A homeomorphism T of a compact metric space X is called *distal* if the orbit $(T^n x, T^n x')$, $n \in \mathbb{Z}$, is bounded away from the diagonal in $X \times X$ for each $x \neq x'$. Some of topologically distal classes already appeared in previous sections. Indeed, zero entropy (minimal) affine transformations are examples of distal homeomorphisms. Another natural class of distal (uniquely ergodic) homeomorphisms is given by nil-translations and, more generally, affine unipotent diffeomorphisms of nilmanifolds. A theorem by Lindenstrauss [119] says that a measurably distal automorphism R has a minimal⁷⁸ model (X, T) together with $\mu \in M^e(X, T)$ of full support (and (X, μ, T) is isomorphic to (Z, κ, R)) in which T is topologically distal.

The following (still open) question seems to be a natural and important step in proving Sarnak's conjecture:

Question 9 (Liu and Sarnak [120]). Are all topologically distal systems Möbius disjoint?

As transformations with discrete spectrum are measurably distal and Theorem 5.12 holds, we can of course ask whether given a measurably distal automorphism, all of its uniquely ergodic models are Möbius disjoint.⁷⁹

We now focus on the famous class of Anzai skew products. This is the class of transformations defined on \mathbb{T}^2 by the formula:

$$T_\varphi: \mathbb{T}^2 \rightarrow \mathbb{T}^2, T_\varphi(x, y) = (x + \alpha, \varphi(x) + y).$$

In other words, Anzai skew products are given by $Tx = x + \alpha$ an irrational rotation on the (additive) circle, and a measurable $\varphi: \mathbb{T} \rightarrow \mathbb{T}$; the skew product T_φ preserves the Lebesgue measure. If φ is continuous, T_φ is a homeomorphism of \mathbb{T}^2 . If we cannot solve the functional equations

$$(45) \quad k\varphi(x) = \xi(x) - \xi(Tx)$$

($k \in \mathbb{N}$) in continuous functions $\xi: \mathbb{T} \rightarrow \mathbb{T}$, then T_φ is minimal, but if for one $k \in \mathbb{N}$ we have a measurable solution then T_φ is not uniquely ergodic. In [120], we find examples of Anzai skew products which are minimal not uniquely ergodic but are Möbius disjoint,⁸⁰ moreover it is proved that if φ is analytic with an additional condition on the decay (from below) of Fourier coefficients then T_φ is Möbius disjoint for each irrational α . In [110], it is proved that if φ is of class $C^{1+\delta}$ then for a typical (in topological sense) α , we have Möbius disjointness of T_φ .⁸¹ A remarkable result is proved by Wang [165]: all analytic Anzai skew products are Möbius disjoint. The proofs in all these papers are using Fourier analysis techniques but in [165], it is also a short interval argument from [125] used in one crucial case.

⁷⁸In general, there is no uniquely ergodic model (X, T) of R with T topologically distal.

⁷⁹As a matter of fact, such a question remains open even for 2-point extensions of irrational rotations.

⁸⁰As a matter of fact, in [4] it is proved that if a uniquely ergodic homeomorphism T satisfies the strong MOMO property (see Definition 4.1 on page 30) and (continuous) $\varphi: X \rightarrow G$ (G is a compact Abelian group) satisfies $\varphi := \xi - \xi \circ T$ has a measurable solution $\xi: X \rightarrow G$, then the homeomorphism T_φ of $X \times G$ is Möbius disjoint. This applies if (45) has a measurable solution for $k = 1$. It is however an open question whether we have Möbius disjointness when there is no measurable solution for $k = 1$ but there is such a solution for some $k \geq 2$.

⁸¹It follows from a subsequent paper [111] that the Anzai skew products considered in [110] enjoy the AOP property.

Nothing seems to be proved about a PNT in the class of distal systems (except for rotations).

5.3.1 Discrete spectrum automorphisms

The simplest examples of (measurably) distal automorphisms are those with discrete spectrum. Recall that a measure-theoretic system $(Z, \mathcal{D}, \kappa, R)$ is said to *have discrete spectrum* if the L^2 -space is generated by the eigenfunctions of the Koopman operator $Tf := f \circ T$. The classical Halmos-von Neumann theorem tells us that each ergodic automorphism with discrete spectrum has a uniquely ergodic model being a rotation on a compact Abelian (monothetic) group.

Theorem 5.12. *All uniquely ergodic models of automorphisms with discrete spectrum are Möbius disjoint.*

This result was first proved in [7] for totally ergodic discrete spectrum automorphisms (as they have the AOP property) and in full generality by Huang, Wang and Zhang in [91]. In fact, the latter result is stronger:

Theorem 5.13 ([91]). *Let (X, T) be a dynamical system, $x \in X$ and $N_i \rightarrow \infty$. Assume that $\frac{1}{N_i} \sum_{n \leq N_i} \delta_{T^n x} \xrightarrow{i \rightarrow \infty} \mu$. Assume that μ is a convex combination of countably many ergodic measures, each of which yields a system with discrete spectrum. Then $\lim_{i \rightarrow \infty} \frac{1}{N_i} \sum_{n \leq N_i} f(T^n x) \mu(n) = 0$ for each $f \in C(X)$.*

Note that Theorem 5.12 also follows from Theorem 4.3 because ergodic rotations enjoy the strong MOMO property [4] (see Remark 4.4). As a matter of fact, as we have already noticed in Corollary 3.20, Theorem 5.12 follows from [125].

5.4 Sub-polynomial complexity

Let T be a homeomorphism of a compact metric space (X, d) and let $\mu \in M(X, T)$. Assume also that $a: \mathbb{N} \rightarrow \mathbb{R}$ is increasing with $\lim_{n \rightarrow \infty} a(n) = \infty$. In the spirit of [61], we say that the measure complexity of μ is weaker than a if

$$\liminf_{n \rightarrow \infty} \frac{\min\{m \geq 1 : \mu(\bigcup_{j=1}^m B_{d_n}(x_j, \varepsilon)) > 1 - \varepsilon \text{ for some } x_1, \dots, x_m \in X\}}{a(n)} = 0$$

for each $\varepsilon > 0$ (here $d_n(y, z) = \frac{1}{n} \sum_{j=1}^n d(T^j y, T^j z)$).

The main result of the recent article [90] states the following:

Theorem 5.14 ([90]). *If (X, T) is a topological system for which all its invariant measures have sub-polynomial complexity, i.e. their complexity is weaker than n^δ for each $\delta > 0$, then (X, T) is Möbius disjoint.*

As shown in [90], Theorem 5.14 applies to: topological systems whose all invariant measures yield systems with discrete spectrum (cf. Corollary 3.20), Anzai skew products of C^∞ -class (over each irrational rotation), $K(\mathbb{Z})$ -sequences introduced by Veech [160] and tame systems.⁸²

⁸²For the latter two classes all invariant measures yield discrete spectrum.

5.5 Systems of number-theoretic origin

Recall that a sequence $x \in \{0, 1\}^{\mathbb{N}}$ is called a *generalized Morse sequence* [105] if

$$(46) \quad x = b^0 \times b^1 \times \dots$$

with $b^i \in \{0, 1\}^{\ell_i}$, $\ell_i \geq 2$, $b^i(0) = 0$ for each $i \geq 0$.⁸³ The following question still remains open.

Question 10 (Mauduit (2014)). Are dynamical systems arising from generalized Morse sequences Möbius disjoint?

Consider the simplest subclass of the class of generalized Morse sequences, for which in (46) we have $|b^i| = 2$ for all $i \geq 0$ (in other words, either $b^i = 01$ or $b^i = 00$). Such sequences are called *Kakutani sequences* [114]. A particular case of Sarnak's conjecture, namely:

$$(47) \quad \frac{1}{N} \sum_{n=1}^N (-1)^{x(n)} \mu(n) \rightarrow 0,$$

for the classical Thue-Morse sequence $x = 01 \times 01 \times \dots$ follows from [94, 104] (see also [41] where, additionally, the speed of convergence to zero is given and [127], where, additionally, a PNT has been proved). Then (47) has been proved for some subclass of Kakutani sequences in [79]. As a matter of fact, in [79], the problem whether $\frac{1}{N} \sum_{n=1}^N (-1)^{s_E(n)} \mu(n) \rightarrow 0$ is considered. Here $E \subset \mathbb{N}$ is fixed and $s_E(n) := \sum_{i \in E} n_i$, where $n = \sum_{i=0}^{\infty} n_i 2^i$ ($n_i \in \{0, 1\}$). To see a relationship with Kakutani sequences define a Kakutani sequence $x = b^0 \times b^1 \times \dots$ with $b^n = 01$ iff $n + 1 \in E$; it is now not hard to see that $s_E(n) = x(n) \pmod{2}$. Finally, using some methods from [127], Bourgain [29] completed the result from [79] so that (47) holds in the whole class of Kakutani sequences (moreover, in [29, 79] a relevant PNT has been proved). One can show that the methods used in the aforementioned papers allow us to have (47) with x replaced by every $y \in \mathcal{O}(x)$ (as shown in [60] in Lemma 6.5 therein, this can be sufficient to show Möbius disjointness for the simple spectrum case; for example, this approach works for the Thue-Morse system).

The problem of Möbius disjointness is also studied (and solved) in the class of (generalized) Kakutani sequences taking values in compact (even non-Abelian) groups, see [160].

5.6 Other research around Sarnak's conjecture

As all periodic observable sequences are orthogonal to μ , one could think that a limit of periodic constructions of type of Toeplitz sequences⁸⁴ also yields systems that are Möbius disjoint.⁸⁵ However, in [2] (and then [51]) there are examples of Toeplitz systems which are not Möbius orthogonal. These examples have positive entropy [3, 51]. Karagulyan in [101] shows Möbius disjointness of zero

⁸³If $B \in \{0, 1\}^k$ and $C = C(0)C(1) \dots C(\ell - 1) \in \{0, 1\}^\ell$ then we define $B \times C := (B + C(0))(B + C(1)) \dots (B + C(\ell - 1))$.

⁸⁴A sequence $x \in A^{\mathbb{N}}$ is called Toeplitz if for each $n \in \mathbb{N}$ there is $q_n \in \mathbb{N}$ such that $x(n + jq_n) = x(n)$ for each $j = 0, 1, \dots$

⁸⁵So called regular Toeplitz sequences are treated in [2] and [51], these are however uniquely ergodic models of odometers.

entropy continuous maps of the interval and (orientation preserving) homeomorphisms of the circle. In [56], Eisner proposes to study a polynomial version of Sarnak’s conjecture (in the minimal case). See also [1, 8, 46, 50, 59, 89].

6 Related research: \mathcal{B} -free numbers

6.1 Introduction

Sets of multiples We have already seen that some properties of the Möbius function μ can be investigated by looking at its square μ^2 , i.e. the characteristic function of the set of square-free numbers $Q := \{n \in \mathbb{Z} : p^2 \nmid n \text{ for all primes } p\}$. A natural generalization comes when we study sets of integers that are not divisible by elements of a given set. Let $\mathcal{B} \subset \mathbb{N}$ and let $\mathcal{M}_{\mathcal{B}}$ be the corresponding set of multiples, i.e. $\mathcal{M}_{\mathcal{B}} = \bigcup_{b \in \mathcal{B}} b\mathbb{Z}$ and the associated set of \mathcal{B} -free numbers $\mathcal{F}_{\mathcal{B}} := \mathbb{Z} \setminus \mathcal{M}_{\mathcal{B}}$ (for convenience, we will deal now with subsets of \mathbb{Z} instead of subsets of \mathbb{N} – the Möbius function μ is not defined for negative arguments, but its square has a natural extension to negative integers). By $\eta = \eta_{\mathcal{B}}$ we will denote the characteristic function of $\mathcal{F}_{\mathcal{B}}$. It is not hard to show that a symmetric subset $F \subset \mathbb{Z}$ is a \mathcal{B} -free set (for some \mathcal{B}) if and only if F is closed under taking divisors.

Historical remarks Sets of multiples were an object of intensive studies already in the 1930s [25, 37, 42, 58]. The basic motivating example there was the set of *abundant numbers* ($n \in \mathbb{Z}$ is *abundant* if $|n|$ is smaller than the sum of its (positive) proper divisors, i.e. $|n| < \sigma(|n|$)), see also more recent [93, 96, 108] on that subject. Also many natural questions on general \mathcal{B} -free sets emerged. Besicovitch [24] showed that the asymptotic density of $\mathcal{M}_{\mathcal{B}}$ may fail to exist. It turned out that it was more natural to use the notion of logarithmic density (denoted by δ) which always exists in this case and equals the lower density. More precisely, we have the following result of Davenport and Erdős:

Theorem 6.1 ([44, 45]). *For any \mathcal{B} , the logarithmic density $\delta(\mathcal{M}_{\mathcal{B}})$ of $\mathcal{M}_{\mathcal{B}}$ exists. Moreover, $\delta(\mathcal{M}_{\mathcal{B}}) = \underline{d}(\mathcal{M}_{\mathcal{B}}) = \lim_{n \rightarrow \infty} d(\mathcal{M}_{\{b \in \mathcal{B} : b \leq n\}})$.*

In the so-called *Erdős case* when \mathcal{B} consists of pairwise coprime elements whose sum of reciprocals converges, the density does exist, cf. [83] (in particular, $\mathbb{1}_{\mathcal{F}_{\mathcal{B}}}$ is rational). We refer the reader to [83, 84] for a coherent, self-contained introduction to the theory of sets of multiples from the analytic and probabilistic number theory viewpoint.

Dynamics comes into play Sarnak in [148], suggested to study μ^2 from the dynamical viewpoint and he announced the following results:

- (i) μ^2 is *generic* for an ergodic S -invariant measure ν_{μ^2} on $\{0, 1\}^{\mathbb{Z}}$ such that the measure-theoretical dynamical system $(X_{\mu^2}, \nu_{\mu^2}, S)$ has zero measure-theoretic entropy;⁸⁶
- (ii) the topological entropy of (X_{μ^2}, S) is equal to $6/\pi^2$;

⁸⁶This is clearly a refinement of the fact that the asymptotic density of square-free integers exists (it is given by $6/\pi^2 = 1/\zeta(2)$). It follows that μ^2 is a completely deterministic point.

(iii) $X_{\mu^2} = X_{\{p^2: p \in \mathbb{P}\}}$ (see the definition of admissibility below);

(iv) (X_{μ^2}, S) is *proximal*.

This triggered intensive research in analogous direction for dynamical systems given by other \mathcal{B} -free sets. In [6], Abdalaoui, Lemańczyk and de la Rue developed the necessary tools in the Erdős case and covered (i)-(iii) from the above list. Given $\mathcal{B} = \{b_k : k \geq 1\}$, In particular, they defined a function $\varphi: G = \prod_{k \geq 1} \mathbb{Z}/b_k\mathbb{Z} \rightarrow \{0, 1\}^{\mathbb{Z}}$ given by

$$\varphi(g)(n) = 1 \iff g_k + n \not\equiv 0 \pmod{b_k} \text{ for all } k \geq 1.$$

Note that $\eta_{\mathcal{B}} = \varphi(0)$ and φ is the coding of points under the translation by $(1, 1, \dots)$ on G with respect to a two-set partition $\{W, W^c\}$, where

$$(48) \quad W = \{h \in G : h_b \neq 0 \text{ for all } b \in \mathcal{B}\}.$$

This study was continued in a general setting in [17] and the first obstacle was that it was no longer clear which subshift to study – it turned out that the most important role is played by the following three subshifts, which coincide in the Erdős case (for the square-free, case see [136] by Peckner and for the Erdős case, see [6]):

- X_{η} is the closure of the orbit of $\eta_{\mathcal{B}}$ under S (*\mathcal{B} -free subshift*),
- \tilde{X}_{η} is the smallest hereditary subshift containing X_{η} (a subshift (X, S) is *hereditary*, whenever $x \in X$ and $y \leq x$ coordinatewise, then $y \in X$),
- $X_{\mathcal{B}}$ is the set of *\mathcal{B} -admissible* sequences, i.e. of $x \in \{0, 1\}^{\mathbb{Z}}$ such that, for each $b \in \mathcal{B}$, the support $\text{supp } x := \{n \in \mathbb{Z} : x(n) = 1\}$ of x taken modulo b is a proper subset of $\mathbb{Z}/b\mathbb{Z}$ (*\mathcal{B} -admissible subshift*).

Remark 6.2. As $X_{\mathcal{B}}$ is hereditary, we have $X_{\eta} \subset \tilde{X}_{\eta} \subset X_{\mathcal{B}}$. In the Erdős case, we have $X_{\eta} = X_{\mathcal{B}}$ [6] (for the square-free system [148]).

Also the group G turned out to be too large for the studies – it is natural to consider its closed subgroup

$$(49) \quad H := \overline{\{(n, n, \dots) \in G : n \in \mathbb{Z}\}}.$$

In the Erdős case we have $H = G$. Certain special cases more general than the Erdős one were considered in [17]:

- we say that \mathcal{B} is *taut* whenever $\delta(\mathcal{F}_{\mathcal{B}}) < \delta(\mathcal{F}_{\mathcal{B} \setminus \{b\}})$ for each $b \in \mathcal{B}$;
- we say that \mathcal{B} has *light tails*, i.e. $\bar{d}(\sum_{b > K} b\mathbb{Z}) \rightarrow 0$ as $K \rightarrow \infty$.

Following [84], we also say that \mathcal{B} is *Besicovitch* if $d(\mathcal{M}_{\mathcal{B}})$ exists (equivalently, $d(\mathcal{F}_{\mathcal{B}})$ exists). A set $\mathcal{B} \subset \mathbb{N} \setminus \{1\}$ is called *Behrend* if $\delta(\mathcal{M}_{\mathcal{B}}) = 1$. Throughout, we will tacitly assume that \mathcal{B} is *primitive*, i.e. does not contain $b \neq b'$ with $b \mid b'$. Recall that \mathcal{B} is taut if and only if \mathcal{B} does not contain $d\mathcal{A}$, where $\mathcal{A} \subset \mathbb{N} \setminus \{1\}$ is Behrend and $d \in \mathbb{N}$.

Further generalizations Several further generalizations of \mathcal{B} -free integers were discussed in the literature from the dynamical viewpoint. Let us briefly recall them here:

- Pleasants and Huck [137] considered *k-free lattice points* $\mathcal{F}_k = \mathcal{F}_k(\Lambda) := \Lambda \setminus \bigcup_{p \in \mathbb{P}} p^k \Lambda$, where Λ is a lattice in \mathbb{R}^d (the corresponding dynamical system given by the orbit closure of $\mathbb{1}_{\mathcal{F}_k} \in \{0, 1\}^\Lambda$ under the multidimensional shift).
- Cellarosi and Vinogradov [34] considered *k-free integers in number fields* $\mathcal{F}_k = \mathcal{F}_k(\mathcal{O}_K) := \mathcal{O}_K \setminus \bigcup_{\mathfrak{p} \in \mathfrak{P}} \mathfrak{p}^k$. Here K is a finite extension of \mathbb{Q} , $\mathcal{O}_K \subset K$ is the ring of integers, \mathfrak{P} stands for the family of all prime ideals in \mathcal{O}_K and \mathfrak{p}^k stands for $\mathfrak{p} \dots \mathfrak{p}$ (\mathfrak{p} is taken k times).
- Baake and Huck in their survey [13] considered *\mathcal{B} -free lattice points* $\mathcal{F}_{\mathcal{B}} = \mathcal{F}_{\mathcal{B}}(\Lambda) := \Lambda \setminus \bigcup_{b \in \mathcal{B}} b\Lambda$. Here Λ is a lattice in \mathbb{R}^d and $\mathcal{B} \subseteq \mathbb{N} \setminus \{1\}$ is an infinite pairwise coprime set with $\sum_{b \in \mathcal{B}} 1/b^d < \infty$.
- Finally, one can consider \mathfrak{B} -free integers $\mathcal{F}_{\mathfrak{B}}$ in number fields as suggested in [13]. Here K is a finite extension of \mathbb{Q} , $\mathcal{O}_K \subset K$ is the ring of integers and \mathfrak{B} is a family of pairwise coprime ideals in \mathcal{O}_K such that the sum of reciprocals of their norms converges.

We will recall some of the main results from the above papers in the relevant sections below.

6.2 Invariant measures and entropy

Mirsky measure Cellarosi and Sinai proved (i) in [33]: they showed that ν_{μ^2} is generic for a shift-invariant measure ν_{μ^2} on $\{0, 1\}^{\mathbb{Z}}$, and that $(X_{\mu^2}, \nu_{\mu^2}, S)$ is isomorphic to a rotation on the compact Abelian group $\prod_{p \in \mathbb{P}} \mathbb{Z}/p^2\mathbb{Z}$. In particular, $(X_{\mu^2}, \nu_{\mu^2}, S)$ is of zero Kolmogorov entropy.⁸⁷ In case of k -free lattice points and k -free integers in number fields an analogous result can be found in [137] and [34], respectively and for \mathcal{B} -free lattice points it was announced in [13]. Recently, Huck [92] showed that in case of \mathfrak{B} -free integers in number fields, the logarithmic density of $\mathcal{F}_{\mathfrak{B}}$ always exists and equals the lower density, thus extending Theorem 6.1 in the (1-dimensional) Erdős case.

Since $\mathcal{F}_{\mathcal{B}}$ may fail to have asymptotic density, the more η may fail to be a generic point. However (Proposition E in [17]), for any $\mathcal{B} \subset \mathbb{N}$, η is always a quasi-generic point for a natural ergodic S -invariant measure ν_{η} on $\{0, 1\}^{\mathbb{Z}}$ (the relevant Mirsky measure). Moreover, \mathcal{B} is Besicovitch if and only if η is generic for ν_{η} . Now, if we additionally assume that \mathcal{B} is taut, then $(X_{\eta}, \nu_{\eta}, S)$ is isomorphic to an ergodic rotation on a compact metric group (Theorem F in [17]).⁸⁸ In particular, $(X_{\eta}, \nu_{\eta}, S)$ has zero entropy.

Finally, for a generalization to so-called weak model sets, see [15], and for some results related to the distribution of \mathcal{B} -free integers, see [10, 11].

⁸⁷The frequencies of blocks on μ^2 were first studied by Mirsky [130, 131] and that is why we refer to ν_{μ^2} (and the analogous measure in case of general \mathcal{B} -free systems) as the *Mirsky measure*.

⁸⁸More precisely, it is isomorphic to (H, \mathbb{P}, T) , where H is the closure of $\{(n \bmod b_k)_{k \geq 1} : n \in \mathbb{Z}\}$ in $\prod_{k \geq 1} \mathbb{Z}/b_k\mathbb{Z}$ and $Tg = g + (1, 1, \dots)$, cf. (49).

Entropy The topological entropy of X_{μ^2} is positive and equals $6/\pi^2 = \prod_{p \in \mathbb{P}} (1 - 1/p^2) = d(\mathcal{F}_{\mathcal{B}})$ for $\mathcal{B} = \{p^2 : p \in \mathbb{P}\}$, see [136]. This extends to the Erdős case, where the topological entropy of $X_\eta = \tilde{X}_\eta = X_{\mathcal{B}}$ equals $\prod_{b \in \mathcal{B}} (1 - 1/b) = d(\mathcal{F}_{\mathcal{B}})$, see [6]. In the general case of \mathcal{B} -free systems, we have $h_{top}(\tilde{X}_\eta, S) = h_{top}(X_{\mathcal{B}}, S) = \delta(\mathcal{F}_{\mathcal{B}})$ (see Theorem K in [17]). The formula for the topological entropy of k -free lattice points is provided in [137].

In view of the variational principle, the positivity of the topological entropy evokes two problems: whether the system under consideration is intrinsically ergodic (i.e. whether there is a unique measure of maximal entropy) and to describe the set of all invariant measures. We address them next.

Maximal entropy measure In the square-free case, the intrinsic ergodicity is proved by Peckner in [136]. This extends to the Erdős case, see [112] by Kułaga-Przymus, Lemańczyk and Weiss. Finally, for any $\mathcal{B} \subset \mathbb{N}$, the subshift (\tilde{X}_η, S) is intrinsically ergodic, see Theorem J in [17]. In particular, if \mathcal{B} has light tails and contains an infinite pairwise coprime subset then $(X_{\mathcal{B}}, S)$ is intrinsically ergodic.

All invariant measures Notice that for each \mathcal{B} , the map $M: X_\eta \times \{0, 1\}^{\mathbb{Z}} \rightarrow \tilde{X}_\eta$ given by the coordinatewise multiplication of sequences is well-defined and each $S \times S$ -invariant measure ρ on $X_\eta \times \{0, 1\}^{\mathbb{Z}}$ yields an S -invariant measure on \tilde{X}_η . In particular, this applies to those ρ whose projection on the first coordinate is ν_η . It turns out that the converse is also true: for any S -invariant measure ν on \tilde{X}_η there exists an $S \times S$ -invariant measure ρ on $X_\eta \times \{0, 1\}^{\mathbb{Z}}$ whose projection on the first coordinate is ν_η and such that $M_*(\rho) = \nu$. For the Erdős case see [112] and for general \mathcal{B} -free systems, see Theorem I in [17] (for further generalizations of \mathcal{B} -free systems listed before (see page 51) no analogous description of the set of all invariant measures is known).

It turns out that a special role is played by \mathcal{B} that are taut. We have the following: for any \mathcal{B} , there exists a unique taut set $\mathcal{B}' \subset \mathbb{N}$ such that $\mathcal{F}_{\mathcal{B}'} \subset \mathcal{F}_{\mathcal{B}}$, $\tilde{X}_{\mathcal{B}'} \subset \tilde{X}_\eta$ and all S -invariant measures on \tilde{X}_η are in fact supported on $\tilde{X}_{\mathcal{B}'}$ (Theorem C in [17]).

More subtle properties of the simplex of invariant measures of the \mathcal{B} -shift have been studied in [113] by Kułaga-Przymus, Lemańczyk and Weiss – it was shown that in the positive entropy case the simplex of S -invariant measures on \tilde{X}_η is *Poulsen*, i.e. the ergodic measures are dense. In particular, if we additionally know that X_η is hereditary (and has positive entropy), then its simplex of invariant measures is Poulsen. However, this is no longer true for a general (not necessarily \mathcal{B} -free) hereditary system. On the other hand, Konieczny, Kupsa and Kwietniak [109] showed that the set of ergodic invariant measures of a hereditary shift is always arcwise connected (when endowed with the d -bar metric).

6.3 Topological results

A lot can be said about the topological properties of (X_η, S) . E.g. for any $\mathcal{B} \subset \mathbb{N}$ the subshift X_η has a unique minimal subset that is the orbit closure of a Toeplitz system (Theorem A in [17]). In particular, X_η is minimal if and

only if X_η is a Toeplitz system.⁸⁹ In fact, η itself can be a Toeplitz sequence (see Example 3.1 in [17]) and it was shown in [103] that η is a Toeplitz sequence different from $\dots 0.00\dots$ if and only if \mathcal{B} does not contain a subset of the form $d\mathcal{A}$, where $d \in \mathbb{N}$ and $\mathcal{A} \subset \mathbb{N} \setminus \{1\}$ is infinite and pairwise coprime. Moreover, if η is Toeplitz then \mathcal{B} is necessarily taut [103].

On the other hand, the proximality of X_η is equivalent to $\{\dots 0.00\dots\}$ being the unique minimal subset of X_η . Moreover, X_η is proximal if and only if \mathcal{B} contains an infinite pairwise coprime subset (Theorem B in [17]).

Some of the properties of the \mathcal{B} -free subshift X_η can be characterized via properties of a set W called the *window*: $W = \{h \in H : h_b \neq 0 \text{ for all } b \in \mathcal{B}\}$, cf. (48). This name has its origins in the theory of weak model sets (for more details see [12]); $\mathcal{F}_\mathcal{B}$ is an example of such a set. Again a special role is played by sets \mathcal{B} that are taut. In [103], Kasjan, Keller and Lemańczyk show the following:

- \mathcal{B} is taut if and only if W is Haar regular, i.e. the topological support of Haar measure restricted to W is the whole W ;
- if \mathcal{B} is primitive then X_η is a Toeplitz system if and only if W is topologically regular;
- X_η is proximal if and only if W has empty interior.

In [103] there is also a detailed description of the maximal equicontinuous factor of X_η (with no extra assumptions on \mathcal{B}). See also [107].

Clearly, if X_η is hereditary, i.e. $X_\eta = \tilde{X}_\eta$ then $(\dots 0.00\dots) \in X_\eta$ and hence X_η is proximal. If we assume that \mathcal{B} is taut then the converse is true: proximality yields heredity (Theorem D in [17]). However, $\tilde{X}_\eta = X_\mathcal{B}$ may fail to hold, even under quite strong assumptions on \mathcal{B} . Indeed, the set of abundant numbers \mathbb{A} is the corresponding set of multiples $\mathcal{M}_\mathcal{B}$ for a certain set \mathcal{B} with the property that $\sum_{b \in \mathcal{B}} 1/b < \infty$. Here, $\tilde{X}_\eta \neq X_\mathcal{B}$, see Section 11 in [17].

More subtle results on heredity were recently obtained by Keller in [106]. He shows that whenever X_η is proximal then it is contained in a slightly larger subshift that is hereditary (there is no need to make extra assumptions on \mathcal{B}). He also generalizes the concept of heredity to the non-proximal case.

It is also interesting to ask about the (invertible) centralizer of (S, X_η) . In the Erdős case it was proved by Mentzen⁹⁰ in [129] that the group of homeomorphisms commuting with the shift (S, X_η) consists only of the powers of S . In case of some Toeplitz \mathcal{B} -free systems an analogous result was proved by Bartnicka in [16].

Question 11. Is the invertible centralizer trivial for each \mathcal{B} -free subshift?

6.4 Ergodic Ramsey theory

We will now see some connections of the theory of \mathcal{B} -free sets with the theory uniform distribution and ergodic Ramsey theory.

⁸⁹This has been recently improved in [103] and by A. Bartnicka: X_η is minimal if and only if η is Toeplitz.

⁹⁰Mentzen's result is extended in [14] to every hereditary \mathcal{B} -free subshift.

Polynomial recurrence and divisibility Recall that Szemerédi showed [153] that any set $S \subset \mathbb{N}$ with positive upper density contains arbitrarily long arithmetic progressions and Furstenberg [72, 73] introduced an ergodic approach to this result that proved very fruitful from the point of view of various generalizations. In particular, it allowed one to prove the following: for any probability space (X, \mathcal{B}, μ) , invertible measure preserving transformation $T: X \rightarrow X$, $A \in \mathcal{B}$ with $\mu(A) > 0$ and any polynomials $p_i \in \mathbb{Q}[t]$ satisfying $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$, $1 \leq i \leq \ell$, there exists arbitrarily large $n \in \mathbb{N}$ such that

$$(50) \quad \mu(A \cap T^{-p_1(n)}A \cap \dots \cap T^{-p_\ell(n)}A) > 0.$$

In fact, we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(A \cap T^{-p_1(n)}A \cap \dots \cap T^{-p_\ell(n)}A) > 0$$

[18, 86, 116]. One can now restrict attention to a specific subset R of $n \in \mathbb{N}$ for which we ask whether (50) holds or even demand

$$(51) \quad \lim_{N \rightarrow \infty} \frac{1}{|R \cap [1, N]|} \sum_{n=1}^N \mathbb{1}_R(n) \mu(A \cap T^{-p_1(n)}A \cap \dots \cap T^{-p_\ell(n)}A) > 0.$$

If (51) holds for any invertible measure preserving system (X, \mathcal{B}, μ, T) , $A \in \mathcal{B}$ with $\mu(A) > 0$, $\ell \in \mathbb{N}$ and any polynomials $p_i \in \mathbb{Q}[t]$, $i = 1, \dots, \ell$, with $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$ for all $i \in \{1, \dots, \ell\}$, we say (cf. [20, Definition 1.5]) that $R \subset \mathbb{N}$ is *averaging set of polynomial multiple recurrence*. If $\ell = 1$, we speak of an *averaging set of polynomial single recurrence*.

We will be interested in polynomial recurrence for \mathcal{B} -free sets. Before we get there, let us direct our attention to so-called rational sets. Recall that $R \subset \mathbb{N}$ is rational if it can be approximated in density by finite unions of arithmetic progressions, cf. footnote 18. Note that the rationality of $\mathcal{F}_{\mathcal{B}}$ is equivalent to \mathcal{B} being Besicovitch. An easy necessary condition for $R \subset \mathbb{N}$ to be an averaging set of polynomial recurrence is that the density of $R \cap u\mathbb{N}$ exists and is positive for any $u \in \mathbb{N}$ (indeed, otherwise consider the cyclic rotation on $\mathbb{Z}/u\mathbb{Z}$ to see that even usual recurrence fails). If the latter holds, we will say that R is *divisible*. It turns out that in case of rational sets, divisibility is not only necessary but also sufficient for polynomial recurrence. More precisely, we have the following:

Theorem 6.3 ([21]). *Let $R \subset \mathbb{N}$ be rational and of positive density. The following conditions are equivalent:*

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

Recall that it was proved in [19] that every self-shift $Q - r$, $r \in \mathbb{Q}$, of the set of square-free numbers Q is divisible and these are the only divisible shifts of Q . For general \mathcal{B} -free sets the situation is more complicated and we have the following result:

Theorem 6.4 ([21]). *Given $\mathcal{B} \subset \mathbb{N}$ that is Besicovitch, there exists a set $D \subset \mathcal{F}_{\mathcal{B}}$ with $d(\mathcal{F}_{\mathcal{B}} \setminus D) = 0$ such that the set $\mathcal{F}_{\mathcal{B}} - r$ is an averaging set of polynomial multiple recurrence if and only if $r \in D$. Moreover, $D = \mathcal{F}_{\mathcal{B}}$ if and only if the set \mathcal{B} is taut.*

This can be generalized to \mathcal{B} that are not Besicovitch by considering divisibility and recurrence along a certain subsequence $(N_k)_{k \geq 1}$. As a combinatorial application, one obtains in [21] the following result: Suppose that $(N_k)_{k \geq 1}$ is such that the density of $\mathcal{F}_{\mathcal{B}}$ along $(N_k)_{k \geq 1}$ exists and is positive. Then there exists $D \subset \mathcal{F}_{\mathcal{B}}$ which equals $\mathcal{F}_{\mathcal{B}}$ up to a set of zero density along $(N_k)_{k \geq 1}$ such that for all $r \in D$ and for all $E \subset \mathbb{N}$ with positive upper density, for any polynomials $p_i \in \mathbb{Q}[t]$, $i = 1, \dots, \ell$, which satisfy $p_i(\mathbb{Z}) \subset \mathbb{Z}$ and $p_i(0) = 0$, for all $1 \leq i \leq \ell$, there exists $\beta > 0$ such that the set

$$\left\{ n \in \mathcal{F}_{\mathcal{B}} - r : \bar{d}\left(E \cap (E - p_1(n)) \cap \dots \cap (E - p_\ell(n))\right) > \beta \right\}$$

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

Results of similar flavor as above have been also obtained in [23] in the context of level sets of multiplicative functions. In particular, if E is a level set of a multiplicative function and has positive density then every self-shift of E is an averaging set of polynomial multiple recurrence (Corollary C in [23]). The key tool here is (17) that provides an important link between level sets of multiplicative functions and rational sets. See also [22].

Acknowledgments

The research resulting in this survey was carried out during the Research in Pairs Program of CIRM, Luminy, France, 15-19.05.2017. J. Kułaga-Przymus and M. Lemańczyk also acknowledge the support of Narodowe Centrum Nauki grant UMO-2014/15/B/ST1/03736. J. Kułaga-Przymus was also supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 647133 (ICHAOS)).

The authors special thanks go to N. Frantzikinakis and P. Sarnak for a careful reading of the manuscript, numerous remarks and suggestions to improve presentation. We also thank M. Baake, V. Bergelson, B. Green, D. Kwietniak, C. Mauduit and M. Radziwiłł for some useful comments on the subject.

References

- [1] E. H. El Abdalaoui and M. Disertori, *Spectral properties of the Möbius function and a random Möbius model*, Stoch. Dyn. **16** (2016), no. 1, 1650005, 25.
- [2] E. H. El Abdalaoui, S. Kasjan, and M. Lemańczyk, *0-1 sequences of the Thue-Morse type and Sarnak's conjecture*, Proc. Amer. Math. Soc. **144** (2016), no. 1, 161–176.
- [3] E. H. 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** (2017), no. 6, 2899–2944.

- [4] ———, *Möbius disjointness for models of an ergodic system and beyond*, preprint, <https://arxiv.org/abs/1704.03506>.
- [5] E. H. El Abdalaoui, M. Lemańczyk, and T. de la Rue, *On spectral disjointness of powers for rank-one transformations and Möbius orthogonality*, *J. Funct. Anal.* **266** (2014), no. 1, 284–317.
- [6] ———, *A dynamical point of view on the set of \mathcal{B} -free integers*, *Int. Math. Res. Not.* (2015), no. 16, 7258–7286.
- [7] ———, *Automorphisms with Quasi-discrete Spectrum, Multiplicative Functions and Average Orthogonality Along Short Intervals*, *Int. Math. Res. Not.* (2017), no. 14, 4350–4368.
- [8] E. H. El Abdalaoui and X. Ye, *A cubic non-conventional ergodic average with multiplicative or von Mangoldt weights*, preprint, <https://arxiv.org/abs/1606.05630>.
- [9] T. Adams, S. Ferenczi, and K. Petersen, *Constructive symbolic presentations of rank one measure-preserving systems*, to appear in *Coll. Math.*
- [10] M. Avdeeva, *Variance of \mathcal{B} -free integers in short intervals*, preprint, <https://arxiv.org/abs/1512.00149>.
- [11] M. Avdeeva, F. Cellarosi, and Y. G. Sinai, *Ergodic and statistical properties of \mathcal{B} -free numbers*, *Teor. Veroyatnost. i Primenen.* **61** (2016), 805–829.
- [12] M. Baake and U. Grimm, *Aperiodic Order. Vol 1. A Mathematical Invitation*, *Encyclopedia of Mathematics and its Applications*, vol. 149, Cambridge University Press, Cambridge, September 2013.
- [13] M. Baake and C. Huck, *Ergodic properties of visible lattice points*, *Proceedings of the Steklov Institute of Mathematics* **288** (2015), no. 1, 165–188 (English).
- [14] M. Baake, C. Huck, and M. Lemańczyk, *Positive entropy shifts with small centraliser and large normaliser*, in preparation.
- [15] M. Baake, C. Huck, and N. Strungaru, *On weak model sets of extremal density*, *Indagationes Mathematicae* **28** (2017), no. 1, 3 – 31, Special Issue on Automatic Sequences, Number Theory, and Aperiodic Order.
- [16] A. Bartnicka, *Automorphisms of Toeplitz \mathcal{B} -free systems*, preprint, <http://arxiv.org/abs/1705.07021>.
- [17] A. Bartnicka, S. Kasjan, J. Kułaga-Przymus, and M. Lemańczyk, *\mathcal{B} -free sets and dynamics*, to appear in *Trans. Amer. Math. Soc.*, <http://arxiv.org/abs/1509.08010>.
- [18] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden’s and Szemerédi’s theorems*, *J. Amer. Math. Soc.* **9** (1996), no. 3, 725–753.
- [19] V. Bergelson and I. Ruzsa, *Squarefree numbers, IP sets and ergodic theory*, *Paul Erdős and his mathematics, I* (Budapest, 1999), *Bolyai Soc. Math. Stud.*, vol. 11, János Bolyai Math. Soc., Budapest, 2002, pp. 147–160.
- [20] V. Bergelson and I. J. Håland, *Sets of recurrence and generalized polynomials*, *Convergence in ergodic theory and probability* (Columbus, OH, 1993), *Ohio State Univ. Math. Res. Inst. Publ.*, vol. 5, de Gruyter, Berlin, 1996, pp. 91–110.

- [21] V. Bergelson, J. Kułaga-Przymus, M. Lemańczyk, and F. K. Richter, *Rationally almost periodic sequences, polynomial multiple recurrence and symbolic dynamics*, to appear in Erg. Theory Dyn. Syst., <https://arxiv.org/abs/1611.08392>.
- [22] ———, *A generalization of Kátai’s orthogonality criterion with applications*, preprint, <https://arxiv.org/abs/1705.07322>.
- [23] ———, *A structure theorem for level sets of multiplicative functions and applications*, preprint, <https://arxiv.org/abs/1708.02613>.
- [24] A. S. Besicovitch, *On the density of certain sequences of integers*, Math. Ann. **110** (1935), no. 1, 336–341.
- [25] E. Bessel-Hagen, *Zahlentheorie*, Teubner, 1929.
- [26] J. Bourgain, *An approach to pointwise ergodic theorems*, Geometric aspects of functional analysis (1986/87), Lecture Notes in Math., vol. 1317, Springer, Berlin, 1988, pp. 204–223.
- [27] ———, *On the maximal ergodic theorem for certain subsets of the integers*, Israel J. Math. **61** (1988), no. 1, 39–72.
- [28] ———, *On the pointwise ergodic theorem on L^p for arithmetic sets*, Israel J. Math. **61** (1988), no. 1, 73–84.
- [29] ———, *Möbius-Walsh correlation bounds and an estimate of Mauduit and Rivat*, J. Anal. Math. **119** (2013), 147–163.
- [30] ———, *On the correlation of the Moebius function with rank-one systems*, J. Anal. Math. **120** (2013), 105–130.
- [31] J. Bourgain, P. Sarnak, and T. Ziegler, *Disjointness of Möbius from horocycle flows*, From Fourier analysis and number theory to Radon transforms and geometry, Dev. Math., vol. 28, Springer, New York, 2013, pp. 67–83.
- [32] D. Carmon and Z. Rudnick, *The autocorrelation of the Möbius function and Chowla’s conjecture for the rational function field*, Q. J. Math. **65** (2014), no. 1, 53–61.
- [33] F. Cellarosi and Y. G. Sinai, *Ergodic properties of square-free numbers*, J. Eur. Math. Soc. **15** (2013), no. 4, 1343–1374.
- [34] F. Cellarosi and I. Vinogradov, *Ergodic properties of k -free integers in number fields*, J. Mod. Dyn. **7** (2013), no. 3, 461–488.
- [35] J. Chaika and A. Eskin, *Möbius disjointness for interval exchange transformations on three intervals*, preprint, <https://arxiv.org/abs/1606.02357>.
- [36] S. Chowla, *The Riemann hypothesis and Hilbert’s tenth problem*, Mathematics and Its Applications, Vol. 4, Gordon and Breach Science Publishers, New York, 1965.
- [37] S. Chowla, *On abundant numbers*, J. Indian Math. Soc., New Ser. **1** (1934), 41–44 (English).
- [38] I. P. Cornfeld, S. V. Fomin, and Y. G. Sinai, *Ergodic theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 245, Springer-Verlag, New York, 1982.

- [39] C. Cuny and M. Weber, *Ergodic theorems with arithmetical weights*, Israel J. Math. **217** (2017), no. 1, 139–180.
- [40] H. Daboussi and H. Delange, *On multiplicative arithmetical functions whose modulus does not exceed one*, J. London Math. Soc. (2) **26** (1982), no. 2, 245–264.
- [41] C. Dartyge and G. Tenenbaum, *Sommes des chiffres de multiples d’entiers*, Ann. Inst. Fourier (Grenoble) **55** (2005), no. 7, 2423–2474.
- [42] H. Davenport, *Über numeri abundantes*, Sitzungsber.Preuss.Akad.Wiss. (1933), 830–837.
- [43] ———, *On some infinite series involving arithmetical functions. II*, Quart. J. Math. Oxford **8** (1937), 313–320.
- [44] H. Davenport and P. Erdős, *On sequences of positive integers*, Acta Arithmetica **2** (1936), no. 1, 147–151 (eng).
- [45] ———, *On sequences of positive integers*, J. Indian Math. Soc. (N.S.) **15** (1951), 19–24.
- [46] R. de la Breteche and G. Tenenbaum, *A remark on Sarnak’s conjecture*, preprint, <https://arxiv.org/abs/1709.01194>.
- [47] M. Denker, C. Grillenberger, and K. Sigmund, *Ergodic theory on compact spaces*, Lecture Notes in Mathematics, Vol. 527, Springer-Verlag, Berlin-New York, 1976.
- [48] J.-M. Deshouillers, M. Drmota, and C. Müllner, *Automatic sequences generated by synchronizing automata fulfill the Sarnak conjecture*, Studia Math. **231** (2015), no. 1, 83–95.
- [49] H. G. Diamond, *Elementary methods in the study of the distribution of prime numbers*, Bull. Amer. Math. Soc. (N.S.) **7** (1982), no. 3, 553–589.
- [50] T. Downarowicz and E. Glasner, *Isomorphic extensions and applications*, Topol. Methods Nonlinear Anal. **48** (2016), no. 1, 321–338.
- [51] T. Downarowicz and S. Kasjan, *Odometers and Toeplitz systems revisited in the context of Sarnak’s conjecture*, Studia Math. **229** (2015), no. 1, 45–72.
- [52] T. Downarowicz and J. Serafin, *Almost full entropy subshifts uncorrelated to the Möbius function*, preprint, <https://arxiv.org/abs/1611.02084>.
- [53] M. Drmota, *Subsequences of automatic sequences and uniform distribution*, Uniform distribution and quasi-Monte Carlo methods, Radon Ser. Comput. Appl. Math., vol. 15, De Gruyter, Berlin, 2014, pp. 87–104.
- [54] M. Drmota, C. Müllner, and L. Spiegelhofer, *Möbius orthogonality for the Zeckendorf sum-of-digits function*, preprint, <http://arxiv.org/abs/1706.09680>.
- [55] M. Einsiedler and T. Ward, *Ergodic theory with a view towards number theory*, Graduate Texts in Mathematics, vol. 259, Springer-Verlag London, Ltd., London, 2011.
- [56] T. Eisner, *A polynomial version of Sarnak’s conjecture*, C. R. Math. Acad. Sci. Paris **353** (2015), no. 7, 569–572.

- [57] P. D. T. A. Elliott, *Probabilistic number theory. I*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science], vol. 239, Springer-Verlag, New York-Berlin, 1979, Mean-value theorems.
- [58] P. Erdős, *On the Density of the Abundant Numbers*, J. London Math. Soc. **9** (1934), no. 4, 278–282.
- [59] A. Fan, *Weighted Birkhoff ergodic theorem with oscillating weights*, preprint, <https://arxiv.org/abs/1705.02501>.
- [60] S. Ferenczi, J. Kułaga-Przymus, M. Lemańczyk, and C. Mauduit, *Substitutions and Möbius disjointness*, Ergodic Theory, Dynamical Systems, and the Continuing Influence of John C. Oxtoby, Contemp. Math., vol. 678, Amer. Math. Soc., Providence, RI, 2016, pp. 151–173.
- [61] S. Ferenczi, *Measure-theoretic complexity of ergodic systems*, Israel J. Math. **100** (1997), 189–207.
- [62] S. Ferenczi, C. Holton, and L. Q. Zamboni, *Structure of three-interval exchange transformations. II. A combinatorial description of the trajectories*, J. Anal. Math. **89** (2003), 239–276.
- [63] S. Ferenczi and C. Mauduit, *On Sarnak’s conjecture and Veech’s question for interval exchanges*, to appear in J. d’Analyse Math.
- [64] L. Flaminio, K. Frączek, J. Kułaga-Przymus, and M. Lemańczyk, *Approximate orthogonality of powers for ergodic affine unipotent diffeomorphisms on nilmanifolds*, preprint, <https://arxiv.org/abs/1609.00699>.
- [65] N. Frantzikinakis, *An averaged Chowla and Elliott conjecture along independent polynomials*, <https://arxiv.org/abs/1606.08420>.
- [66] ———, *Ergodicity of the Liouville system implies the Chowla conjecture*, <https://arxiv.org/abs/1611.09338>.
- [67] N. Frantzikinakis and B. Host, *Higher order Fourier analysis of multiplicative functions and applications*, J. Amer. Math. Soc. **30** (2017), no. 1, 67–157.
- [68] N. Frantzikinakis and B. Host, *The logarithmic Sarnak conjecture for ergodic weights*, preprint, <https://arxiv.org/abs/1708.00677>.
- [69] H. Furstenberg, *Strict ergodicity and transformation of the torus*, Amer. J. Math. **83** (1961), 573–601.
- [70] ———, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory **1** (1967), 1–49.
- [71] ———, *The unique ergodicity of the horocycle flow*, (1973), 95–115. Lecture Notes in Math., Vol. 318.
- [72] ———, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256.
- [73] ———, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, Princeton, N.J., 1981, M. B. Porter Lectures.
- [74] A. O. Gel’fond, *Sur les nombres qui ont des propriétés additives et multiplicatives données*, Acta Arith. **13** (1967/1968), 259–265.

- [75] E. Glasner, *Ergodic theory via joinings*, Mathematical Surveys and Monographs, vol. 101, American Mathematical Society, Providence, RI, 2003.
- [76] A. Górnika, D. Kwietniak, and M. Lemańczyk, *Sarnak’s conjecture implies the chowla conjecture along a subsequence*, Proceedings of the Chair Morlet semester “Ergodic Theory and Dynamical Systems in their Interactions with Arithmetic and Combinatorics” 1.08.2016-31.01.2017, Lecture Notes in Math., Springer, to appear.
- [77] W. T. Gowers, *A new proof of Szemerédi’s theorem*, *Geom. Funct. Anal.* **11** (2001), no. 3, 465–588.
- [78] K. Granville, A. Soundararajan, *Multiplicative number theory: The pretentious approach.*, Book manuscript in preparation.
- [79] B. Green, *On (not) computing the Möbius function using bounded depth circuits*, *Combin. Probab. Comput.* **21** (2012), no. 6, 942–951.
- [80] B. Green and T. Tao, *The Möbius function is strongly orthogonal to nilsequences*, *Ann. of Math. (2)* **175** (2012), no. 2, 541–566.
- [81] ———, *The quantitative behaviour of polynomial orbits on nilmanifolds*, *Ann. of Math. (2)* **175** (2012), no. 2, 465–540.
- [82] F. Hahn and W. Parry, *Some characteristic properties of dynamical systems with quasi-discrete spectra*, *Math. Systems Theory* **2** (1968), 179–190.
- [83] H. Halberstam and K. F. Roth, *Sequences*, second ed., Springer-Verlag, New York-Berlin, 1983.
- [84] R. R. Hall, *Sets of multiples*, Cambridge Tracts in Mathematics, vol. 118, Cambridge University Press, Cambridge, 1996.
- [85] A. Hildebrand, *Introduction to analytic number theory*, <http://www.math.uiuc.edu/~ajh/531.fall105/>.
- [86] B. Host and B. Kra, *Convergence of polynomial ergodic averages*, *Israel J. Math.* **149** (2005), 1–19, Probability in mathematics.
- [87] ———, *Nonconventional ergodic averages and nilmanifolds*, *Ann. of Math. (2)* **161** (2005), no. 1, 397–488.
- [88] B. Host, J.-F. Méla, and F. Parreau, *Nonsingular transformations and spectral analysis of measures*, *Bull. Soc. Math. France* **119** (1991), no. 1, 33–90.
- [89] W. Huang, Z. Lian, S. Shao, and X. Ye, *Sequences from zero entropy noncommutative toral automorphisms and Sarnak Conjecture*, *J. Differential Equations* **263** (2017), no. 1, 779–810.
- [90] W. Huang, Z. Wang, and X. Ye, *Measure complexity and Möbius disjointness*, <https://arxiv.org/abs/1707.06345>.
- [91] W. Huang, Z. Wang, and G. Zhang, *Möbius disjointness for topological models of ergodic systems with discrete spectrum*, preprint, <https://arxiv.org/abs/1608.08289>.
- [92] C. Huck, *On the logarithmic probability that a random integral ideal is relatively \mathcal{A} -free*, From Chowla’s conjecture: from the Liouville function to the Moebius function, Lecture Notes in Math., Springer, to appear.

- [93] D. E. Iannucci, *On the smallest abundant number not divisible by the first k primes*, Bull. Belg. Math. Soc. Simon Stevin **12** (2005), no. 1, 39–44.
- [94] K.-H. Indlekofer and I. Kátai, *Investigations in the theory of q -additive and q -multiplicative functions. I*, Acta Math. Hungar. **91** (2001), no. 1-2, 53–78.
- [95] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.
- [96] E. Jennings, P. Pollack, and L. Thompson, *Variations on a theorem of Davenport concerning abundant numbers*, Bull. Aust. Math. Soc. **89** (2014), no. 3, 437–450.
- [97] A. del Junco and D. Rudolph, *On ergodic actions whose self-joinings are graphs*, Ergodic Theory Dynam. Systems **7** (1987), no. 4, 531–557.
- [98] A. del Junco, *Disjointness of measure-preserving transformations, minimal self-joinings and category*, Ergodic theory and dynamical systems, I (College Park, Md., 1979–80), Progr. Math., vol. 10, Birkhäuser, Boston, Mass., 1981, pp. 81–89.
- [99] T. Kamae, *Subsequences of normal sequences*, Israel J. Math. **16** (1973), 121–149.
- [100] A. Kanigowski, M. Lemańczyk, and C. Ulcigrai, *On disjointness properties of some parabolic flows*, in preparation.
- [101] D. Karagulyan, *On Möbius orthogonality for interval maps of zero entropy and orientation-preserving circle homeomorphisms*, Ark. Mat. **53** (2015), no. 2, 317–327.
- [102] ———, *On Möbius orthogonality for subshifts of finite type with positive topological entropy*, Studia Math. **237** (2017), no. 3, 277–282.
- [103] S. Kasjan, G. Keller, and M. Lemańczyk, *Dynamics of \mathcal{B} -free sets: a view through the window*, to appear in Int. Math. Res. Not., <https://arxiv.org/abs/1702.02375>.
- [104] I. Kátai, *A remark on a theorem of H. Daboussi*, Acta Math. Hungar. **47** (1986), no. 1-2, 223–225.
- [105] M. Keane, *Generalized Morse sequences*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **10** (1968), 335–353.
- [106] G. Keller, *Generalized heredity in \mathcal{B} -free systems*, preprint, <https://arxiv.org/abs/1704.04079>.
- [107] G. Keller and C. Richard, *Dynamics on the graph of the torus parametrisation*, Ergod. Th. Dynam. Sys. first online (2015), <https://arxiv.org/abs/1511.06137v3.pdf>.
- [108] M. Kobayashi, *A new series for the density of abundant numbers*, Int. J. Number Theory **10** (2014), no. 1, 73–84.
- [109] J. Konieczny, M. Kupsa, and D. Kwietniak, *Arcwise connectedness of the set of ergodic measures of hereditary shifts*, preprint, <https://arxiv.org/abs/1610.00672>.
- [110] J. Kułaga-Przymus and M. Lemańczyk, *The Möbius function and continuous extensions of rotations*, Monatsh. Math. **178** (2015), no. 4, 553–582.

- [111] ———, *Möbius disjointness along ergodic sequences for uniquely ergodic actions*, preprint, <https://arxiv.org/abs/1703.02347>.
- [112] J. Kułaga-Przymus, M. Lemańczyk, and B. Weiss, *On invariant measures for \mathcal{B} -free systems*, Proc. Lond. Math. Soc. (3) **110** (2015), no. 6, 1435–1474.
- [113] ———, *Hereditary subshifts whose simplex of invariant measures is Poulsen*, Ergodic theory, dynamical systems, and the continuing influence of John C. Oxtoby, Contemp. Math., vol. 678, Amer. Math. Soc., Providence, RI, 2016, pp. 245–253.
- [114] J. Kwiatkowski, *Spectral isomorphism of Morse dynamical systems*, Bull. Acad. Polon. Sci. Sér. Sci. Math. **29** (1981), no. 3-4, 105–114.
- [115] E. Lehrer, *Topological mixing and uniquely ergodic systems*, Israel J. Math. **57** (1987), no. 2, 239–255.
- [116] A. Leibman, *Convergence of multiple ergodic averages along polynomials of several variables*, Israel J. Math. **146** (2005), 303–315.
- [117] ———, *Multiple polynomial correlation sequences and nilsequences*, Ergodic Theory Dynam. Systems **30** (2010), no. 3, 841–854.
- [118] M. Lemańczyk and F. Parreau, *Lifting mixing properties by Rokhlin cocycles*, Ergodic Theory Dynam. Systems **32** (2012), no. 2, 763–784.
- [119] E. Lindenstrauss, *Measurable distal and topological distal systems*, Ergodic Theory Dynam. Systems **19** (1999), no. 4, 1063–1076.
- [120] J. Liu and P. Sarnak, *The Möbius function and distal flows*, Duke Math. J. **164** (2015), no. 7, 1353–1399.
- [121] B. Marcus, *The horocycle flow is mixing of all degrees*, Invent. Math. **46** (1978), no. 3, 201–209.
- [122] B. Martin, C. Mauduit, and J. Rivat, *Théorème des nombres premiers pour les fonctions digitales*, Acta Arith. **165** (2014), no. 1, 11–45.
- [123] ———, *Fonctions digitales le long des nombres premiers*, Acta Arith. **170** (2015), no. 2, 175–197.
- [124] K. Matomäki and M. Radziwiłł, *Multiplicative functions in short intervals*, Ann. of Math. (2) **183** (2016), no. 3, 1015–1056.
- [125] K. Matomäki, M. Radziwiłł, and T. Tao, *An averaged form of Chowla’s conjecture*, Algebra & Number Theory **9** (2015), 2167–2196.
- [126] ———, *Sign patterns of the Liouville and Möbius functions*, Forum Math. Sigma **4** (2016), e14, 44.
- [127] C. Mauduit and J. Rivat, *Sur un problème de Gelfond: la somme des chiffres des nombres premiers*, Ann. of Math. (2) **171** (2010), no. 3, 1591–1646.
- [128] ———, *Prime numbers along Rudin–Shapiro sequences*, J. Eur. Math. Soc. (JEMS) **17** (2015), no. 10, 2595–2642.
- [129] M. K. Mentzen, *Automorphisms of subshifts defined by \mathcal{B} -free sets of integers*, Colloq. Math. **147** (2017), no. 1, 87–94.

- [130] L. Mirsky, *Note on an asymptotic formula connected with r -free integers*, Quart. J. Math., Oxford Ser. **18** (1947), 178–182.
- [131] ———, *Arithmetical pattern problems relating to divisibility by r th powers*, Proc. London Math. Soc. (2) **50** (1949), 497–508.
- [132] H. L. Montgomery and R. C. Vaughan, *Exponential sums with multiplicative coefficients*, Invent. Math. **43** (1977), no. 1, 69–82.
- [133] B. Mossé, *Puissances de mots et reconnaissabilité des points fixes d’une substitution*, Theoret. Comput. Sci. **99** (1992), no. 2, 327–334.
- [134] C. Müllner, *Automatic sequences fulfill the Sarnak conjecture*, to appear in Duke Math., <https://arxiv.org/abs/1602.03042>.
- [135] R. Peckner, *Two dynamical perspectives on the randomness of the Möbius function*, ProQuest LLC, Ann Arbor, MI, 2015, Thesis (Ph.D.)—Princeton University.
- [136] ———, *Uniqueness of the measure of maximal entropy for the squarefree flow*, Israel J. Math. **210** (2015), no. 1, 335–357.
- [137] P. A. B. Pleasants and C. Huck, *Entropy and diffraction of the k -free points in n -dimensional lattices*, Discrete Comput. Geom. **50** (2013), no. 1, 39–68.
- [138] M. Queffélec, *Substitution dynamical systems—spectral analysis*, second ed., Lecture Notes in Mathematics, vol. 1294, Springer-Verlag, Berlin, 2010.
- [139] M. Ram Murty and A. Vatwani, *A remark on a conjecture of Chowla*, to appear in Journal of the Ramanujan Mathematical Society, 2017.
- [140] ———, *Twin primes and the parity problem*, J. Number Theory **180** (2017), 643–659.
- [141] O. Ramaré, *From Chowla’s conjecture: from the Liouville function to the Möbius function*, Proceedings of the Chair Morlet semester “Ergodic Theory and Dynamical Systems in their Interactions with Arithmetic and Combinatorics” 1.08.2016-31.01.2017, Lecture Notes in Math., Springer, to appear.
- [142] M. Ratner, *Horocycle flows, joinings and rigidity of products*, Ann. of Math. (2) **118** (1983), no. 2, 277–313.
- [143] ———, *On Raghunathan’s measure conjecture*, Ann. of Math. (2) **134** (1991), no. 3, 545–607.
- [144] J. Rivat, *Analytic number theory*, Proceedings of the Chair Morlet semester “Ergodic Theory and Dynamical Systems in their Interactions with Arithmetic and Combinatorics” 1.08.2016-31.01.2017, Lecture Notes in Math., Springer, to appear.
- [145] T. de la Rue, *La fonction de Möbius à la rencontre des systèmes dynamiques*, Gaz. Math. (2016), no. 150, 31–40.
- [146] V. V. Ryzhikov, *Bounded ergodic constructions, disjointness, and weak limits of powers*, Trans. Moscow Math. Soc. (2013), 165–171.
- [147] P. Sarnak, *Möbius randomness and dynamics six years later*, <http://www.youtube.com/watch?v=LXXOntxrkb0>.
- [148] ———, *Three lectures on the Möbius function, randomness and dynamics*, <http://publications.ias.edu/sarnak/>.

- [149] ———, *Möbius randomness and dynamics*, Not. S. Afr. Math. Soc. **43** (2012), no. 2, 89–97.
- [150] P. Sarnak and A. Ubis, *The horocycle flow at prime times*, J. Math. Pures Appl. (9) **103** (2015), no. 2, 575–618.
- [151] K. Schmidt, *Cocycles on ergodic transformation groups*, Macmillan Company of India, Ltd., Delhi, 1977, Macmillan Lectures in Mathematics, Vol. 1.
- [152] K. Soundararajan, *The Liouville function in short intervals [after Matomäki and Radziwiłł]*, Séminaire Boubaki, 68ème année (2015-2016), no. 1119.
- [153] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245, Collection of articles in memory of Juriĭ Vladimirovič Linnik.
- [154] T. Tao, *The logarithmically averaged Chowla and Elliot conjectures for two-point correlations*, Forum Math. Pi **4** (2016).
- [155] ———, *Equivalence of the logarithmically averaged Chowla and Sarnak conjectures*, Number Theory – Diophantine Problems, Uniform Distribution and Applications: Festschrift in Honour of Robert F. Tichy’s 60th Birthday (C. Elsholtz and P. Grabner, eds.), Springer International Publishing, Cham, 2017, pp. 391–421.
- [156] ———, *The Bourgain-Sarnak-Ziegler orthogonality criterion*, What’s new, <http://terrytao.wordpress.com/2011/11/21/the-bourgain-sarnak-ziegler-orthogonality-criterion/>.
- [157] ———, *The Chowla and the Sarnak conjecture*, What’s new, <http://terrytao.wordpress.com/2012/10/14/the-chowla-conjecture-and-the-sarnak-conjecture/>.
- [158] T. Tao and J. Teräväinen, *The structure of logarithmically averaged correlations of multiplicative functions, with applications to the Chowla and Elliott conjectures*, preprint, <https://arxiv.org/abs/1708.02610>.
- [159] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics, vol. 46, Cambridge University Press, Cambridge, 1995, Translated from the second French edition (1995) by C. B. Thomas.
- [160] W. A. Veech, *Möbius orthogonality for generalized Morse-Kakutani flows*, to appear in Amer. J. Math.
- [161] A. M. Vershik and A. N. Livshits, *Adic models of ergodic transformations, spectral theory, substitutions, and related topics*, Representation theory and dynamical systems, Adv. Soviet Math., vol. 9, Amer. Math. Soc., Providence, RI, 1992, pp. 185–204.
- [162] I. M. Vinogradov, *Some theorems concerning the theory of primes*, Rec. Math. Moscou **2** (1937), no. 44, 179–195.
- [163] ———, *The method of trigonometrical sums in the theory of numbers*, Dover Publications, Inc., Mineola, NY, 2004, Translated from the Russian, revised and annotated by K. F. Roth and Anne Davenport, Reprint of the 1954 translation.
- [164] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, vol. 79, Springer-Verlag, New York, 1982.

- [165] Z. Wang, *Möbius disjointness for analytic skew products*, to appear in *Inventiones Math.*, <https://arxiv.org/abs/1509.03183>.
- [166] B. Weiss, *Normal sequences as collectives*, Proc. Symp. on Topological Dynamics and ergodic theory, Univ. of Kentucky, 1971.
- [167] ———, *Strictly ergodic models for dynamical systems*, Bull. Amer. Math. Soc. (N.S.) **13** (1985), no. 2, 143–146.
- [168] ———, *Single orbit dynamics*, CBMS Regional Conference Series in Mathematics, vol. 95, American Mathematical Society, Providence, RI, 2000.
- [169] M. Wierdl, *Pointwise ergodic theorem along the prime numbers*, Israel J. Math. **64** (1988), no. 3, 315–336 (1989).
- [170] R. J. Zimmer, *Ergodic actions with generalized discrete spectrum*, Illinois J. Math. **20** (1976), no. 4, 555–588.
- [171] ———, *Extensions of ergodic group actions*, Illinois J. Math. **20** (1976), no. 3, 373–409.

Sebastien Ferenczi
 AIX MARSEILLE UNIVERSITÉ, CNRS, CENTRALE MARSEILLE, INSTITUT DE MATHÉMATIQUES
 DE MARSEILLE, I2M – UMR 7373
 13453 MARSEILLE
 FRANCE
 ssferenczi@gmail.com

Joanna Kułaga-Przymus
 AIX-MARSEILLE UNIVERSITÉ, CNRS, CENTRALE MARSEILLE, INSTITUT DE MATHÉMATIQUES
 DE MARSEILLE, I2M – UMR 7373
 13453, MARSEILLE, FRANCE
 FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, NICOLAUS COPERNICUS UNIVERSITY,
 CHOPINA 12/18, 87-100 TORUŃ, POLAND
 joanna.kulaga@gmail.com

Mariusz Lemańczyk
 FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, NICOLAUS COPERNICUS UNIVERSITY,
 CHOPINA 12/18, 87-100 TORUŃ, POLAND
 mlem@mat.umk.pl