FURSTENBERG DISJOINTNESS, RATNER PROPERTIES, AND SARNAK'S CONJECTURE

MARIUSZ LEMAŃCZYK

ABSTRACT

A recent progress on Sarnak's conjecture on Möbius orthogonality is discussed with the main focus on the proof of Veech's conjecture and its consequences.

MATHEMATICS SUBJECT CLASSIFICATION 2020

Primary 37A44; Secondary 11N37, 37A05, 37A17, 37A30, 37B10, 37B40

INTERNATIONAL CONGRESS OF MATHEMATICIANS 2022 JULY 6–14

Proc. Int. Cong. Math. 2022, Vol. 5, pp. 3508–3528
DOI 10.4171/ICM2022/51

Published by EMS Press

1. INTRODUCTION. STATE-OF-THE-ART

This article deals with Sarnak's conjecture [34] from 2010, also called the Möbius Orthogonality Conjecture, MOC for short, see (1.2) below. More precisely, the aim of this article is to give an account on the progress concerning MOC caused by the proof of Veech's conjecture [38] from 2016 in the recent article [24]. In order to do it, we will present a panorama of earlier concepts and results concerning the new interactions between ergodic theory and analytic number theory caused by MOC, especially relating MOC to the celebrated Chowla conjecture [5] from 1965. We will be concentrated on the directions of research on the ergodic theory side, with a special focus on projects in which the author of the article took part. In this extended introduction, we present the state-of-the-art of the subject. To keep this presentation reasonably short, some elementary definitions and facts from dynamics are postponed to Sections 2 and 3, while some facts (especially around entropy) are treated as "commonly" known and can be found in the ergodic theory literature, see, e.g., [9,16,39]. On the analytic number theory side, basic facts can be found, e.g., in [19,23].

Möbius function. Each natural number $n \in \mathbb{N} := \{1,2,\ldots\}$ has its (unique) decomposition into a product of primes which is the basic fact about the not finitely generated *multiplicative structure* of $\mathbb{N}$. The set $\mathbb{P}$ of primes is believed to behave like a "random" subset of natural numbers. This randomness should be reflected in the properties of arithmetic functions $\boldsymbol{u}:\mathbb{N}\to\mathbb{C}$ that preserve the multiplicative structure of $\mathbb{N}$, that is, they are themselves *multiplicative*, $\boldsymbol{u}(mn)=\boldsymbol{u}(m)\boldsymbol{u}(n)$ whenever $(m,n)=1$ (in other words, they are determined by their values on the powers of the primes). One of most prominent multiplicative functions is the *Möbius function* $\boldsymbol{\mu}:\mathbb{N}\to\{-1,0,1\}$ for which $\boldsymbol{\mu}(1)=1$, $\boldsymbol{\mu}(n)=(-1)^k$, with n being the product of k distinct primes, and $\boldsymbol{\mu}(n)=0$ for the remaining $n\in\mathbb{N}$. Is this function "random"? If so, this should be reflected in the phenomenon of cancelations of ± 1s like for a sample from an independent process. It is then classical that the Prime Number Theorem (PNT), i.e., $|\{p\leq N : p\in\mathbb{P}\}|\sim\frac{N}{\log N}$, is equivalent to $\lim_{N\to\infty}\frac{1}{N}\sum_{n\leq N}\boldsymbol{\mu}(n)=0$, and the Riemann Hypothesis is equivalent to a quantitative version of the above, namely $\sum_{n\leq N}\boldsymbol{\mu}(n)=O(N^{\frac{1}{2}+\varepsilon})$ for each $\varepsilon>0$.

The Chowla conjecture. Another way to express the randomness of $\boldsymbol{\mu}$ is the Chowla conjecture [5] from 1965 claiming that the autocorrelations of the Möbius function, unless they are correlations of $\boldsymbol{\mu}^2$, vanish:[1]

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leq N}\boldsymbol{\mu}^{s_1}(n+a_1)\cdots\boldsymbol{\mu}^{s_k}(n+a_k)=0 \tag{1.1}$$

1 Originally, the Chowla conjecture was formulated for the Liouville function $\boldsymbol{\lambda}:\mathbb{N}\to\{-1,1\}$ and concerned *all* correlations. Liouville function depends on the parity of all prime factors (counted with multiplicities) and clearly satisfies $\boldsymbol{\mu}=\boldsymbol{\lambda}\cdot\boldsymbol{\mu}^2$.

for each $k \geq 1$, $0 \leq a_1 < \cdots < a_k$, and $s_j \in \{1, 2\}$, not all s_j being 2.[2] We will see later (see Section 3) that the Chowla conjecture is precisely the fact that $\boldsymbol{\mu}$ is a generic point for a kind of Bernoulli measure over the so-called Mirsky measure, for which $\boldsymbol{\mu}^2$ is generic, i.e., intuitively, relative to the positions of zeros, in $\boldsymbol{\mu}$ we observe a replacement of 1s in $\boldsymbol{\mu}^2$ by ± 1s with equal probability.

Sarnak's conjecture (MOC). Another way of talking about the randomness of $\boldsymbol{\mu}$ could be in terms of correlations with other sequences. In [23], this is expressed by the Möbius Randomness Law: $\boldsymbol{\mu}$ is so random that it does not correlate with any "reasonable" bounded sequence. In 2010, P. Sarnak [34] formulated a much more precise form of this vague randomness principle, namely

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n \leq N} f(T^n x)\boldsymbol{\mu}(n) = 0, \tag{1.2}$$

for each *zero (topological) entropy* homeomorphism T of a compact metric space X, all $f \in C(X)$, and all $x \in X$.[3] In other words, $\boldsymbol{\mu}$ is random because $\boldsymbol{\mu}$ does not correlate with *any* deterministic sequence,[4] we also say that $\boldsymbol{\mu}$ *is orthogonal* to all (bounded) deterministic sequences or to all deterministic systems. Of course, in (1.2), we can consider other (always bounded though) arithmetic functions $\boldsymbol{u} : \mathbb{N} \to \mathbb{C}$ and consider an analogous problem of orthogonality to a selected class of topological systems (for example, it is well known that if we replace the Möbius function $\boldsymbol{\mu}$ with the Liouville function $\boldsymbol{\lambda}$, then the corresponding Sarnak's conjectures are equivalent; see, e.g., [10]).[5]

Proposition 1.1 (Sarnak [34], for proofs see [1, 35]). *Chowla conjecture implies Sarnak's conjecture.*

Whether the converse is true remains open, but we have the following:

Proposition 1.2 ([17]). *Sarnak's conjecture implies the Chowla conjecture along a subsequence (i.e., in* (1.1) *we need to consider* (N_s) *instead of* N[6]*).*

---

2 If all $s_j = 2$ then the limit exists but need not be zero: $\boldsymbol{\mu}^2$ is the characteristic function of the set of square-free numbers whose natural density is $6/\pi^2$. In fact, the frequencies of all blocks of 0, 1s on $\boldsymbol{\mu}^2$ exist – $\boldsymbol{\mu}^2$ is a generic point for a shift-invariant measure $\nu_{\boldsymbol{\mu}^2}$ called the *Mirsky measure* – see Section 3 for details.

3 It is "for all $x \in X$" which is the core of Sarnak's conjecture: the "x-almost every" version of (1.2) holds for *every* dynamical system (in particular, regardless of the entropy [34], see also [1]).

4 In the theory of dynamical systems, zero-entropy systems are also called *deterministic* and the corresponding continuous observables $(f(T^n x))_{n \in \mathbb{Z}}$ are precisely deterministic sequences.

5 The reader can notice that if $\boldsymbol{u}$ is orthogonal to all deterministic sequences, then for *any* bounded deterministic sequence $(a(n))$, the arithmetic function $\boldsymbol{v}(n) := \boldsymbol{u}(n)a(n)$ is also orthogonal to all deterministic sequences. Of course, such an operation in general "kills" the multiplicativity of $\boldsymbol{u}$.

6 Tao in [37] strengthened this result by showing that (N_s) can be selected to have full logarithmic density.

We will detail more on that at the end of the introduction when we consider the logarithmic versions of the Chowla and Sarnak's conjectures.

Two strategies to "attack" Sarnak's conjecture. Returning to the original Sarnak's conjecture, we can view (1.2) as a classical Cesàro (ergodic) sum with **m**ultiplicative **w**eights (this point of view leads to the MW-strategy) or we can reverse the roles and consider Cesàro sums of $\boldsymbol{\mu}$[7] with **e**rgodic **w**eights (this point of view leads to the EW-strategy). Both strategies lead sooner or later to an interplay between analytic number theory and the theory of joinings in dynamics, in particular, the disjointness theory of Furstenberg in ergodic theory. Let us now say a few words on these strategies; more details, especially on the EW-strategy, will be provided later.

MW-strategy. DDKBSZ criterion. The core of the MW-strategy (in which we only use the fact that $\boldsymbol{\mu}$ is multiplicative) is the following numerical DDKBSZ criterion:[8]

Theorem 1.3 ([4, 27]). *Assume that $(f_n)_{n\in\mathbb{N}}\subset\mathbb{C}$ is bounded. If*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leq N} f_{pn}\overline{f}_{qn}=0 \tag{1.3}$$

for all distinct, sufficiently large primes $p,q\in\mathbb{P}$, then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leq N} f_n\boldsymbol{u}(n)=0$$

for each bounded multiplicative *function $\boldsymbol{u}:\mathbb{N}\to\mathbb{C}$.*

Then, the DDKBSZ criterion is used in the following manner. Take *any* dynamical system (X,T) with a unique invariant measure μ (which must also be unique for all nonzero powers of T) and let $x\in X$. In the space $M(X\times X)$ of probability measures on $X\times X$ consider the sequence of *empiric measures* $(\frac{1}{N}\sum_{n\leq N}\delta_{(T^{pn}x,T^{qn}x)})$ (see Section 3 for more details). By the compactness of the weak-$*$-topology, there exists a subsequence (N_k) such that

$$\lim_{k\to\infty}\frac{1}{N_k}\sum_{n\leq N_k}\delta_{(T^{pn}x,T^{qn}x)}=\rho,$$

where necessarily the measure ρ is $T^p\times T^q$-invariant and if $f\in C(X)$ then

$$\lim_{k\to\infty}\frac{1}{N_k}\sum_{n\leq N_k} f(T^{pn}x)\overline{f(T^{qn}x)}=\int_{X\times X} f\otimes\overline{f}\,d\rho.$$

By our assumption, the two projections of ρ on X are μ. *If* we are able to show that ρ is the product measure $\mu\otimes\mu$, then we can easily apply the DDKBSZ criterion for the continuous zero-mean (for the measure μ) functions f and the sequences $(f(T^nx))$ (i.e., $f_n=f(T^nx)$ in Theorem 1.3). A spectacular case when this approach works was first demonstrated in the

7 In analytic number theory, the problem of studying means of multiplicative functions is classical, cf. Halász theorem.

8 DDKBSZ stands for Daboussi, Delange, Kátai, Bourgain, Sarnak, and Ziegler.

case of horocycle flows in [4] using Ratner's theory. However, a "typical" playground for the MW-strategy is when the automorphisms T^p and T^q considered with the same T-invariant measure μ are disjoint (see Section 2 for definition) in the Furstenberg sense, as in this case $\mu\otimes\mu$ is simply the *only* $T^p\times T^q$-invariant measure. (Note that this is not applicable to the horocycle flows themselves as all positive time automorphisms are isomorphic, so there are many $T^p\times T^q$-invariant measures.) As we can see, the MW-strategy leads to pure ergodic theory problems. While a list of papers in which the disjointness of powers has been proved can be found in the surveys [10,28], in Section 4, we will detail more on the answer to Ratner's question about the validity of MOC for smooth time changes of horocyclic flows. Namely, while for the algebraic actions the underlying configuration space is homogeneous, once the time is changed (in a nontrivial way), the configuration space becomes nonhomogeneous for the action of the time-changed flow which surprisingly leads to another extreme: the new flows enjoy formidable internal disjointness properties which allows us to use Theorem 1.3 to answer positively Ratner's question and go beyond.

MW-Strategy. The AOP property. There is a pure ergodic theory counterpart of DDKBSZ criterion, namely the notion of AOP (Asymptotic Orthogonality of Powers) introduced in [3]: a measure-theoretic (ergodic) dynamical system $(X,\mathcal{B},\mu,T)$ has the AOP property ($J^e(T^p,T^q)$ below stands for the set of ergodic joinings between T^p and T^q) if

$$\limsup_{p\neq q,\mathbb{P}\ni p,q\to\infty}\ \sup_{\rho\in J^e(T^p,T^q)}\left|\int_{X\times X} f\otimes g\,d\rho\right|=0$$

for each $f,g\in L^2_0(X,\mu)$. Obviously, AOP takes place if the prime powers of an automorphisms are disjoint but an AOP automorphism can have all nonzero powers isomorphic. AOP implies zero entropy and total ergodicity, i.e., all nonzero powers are ergodic. All ergodic quasidiscrete spectrum automorphisms [3], and also ergodic nil-automorphisms enjoy this property [12]. Let us see why AOP is useful when proving Möbius orthogonality. Suppose that we want to prove Möbius orthogonality for *all* (uniquely ergodic) models of totally ergodic rotations. Well, the Möbius orthogonality can first be easily established in *some* models of such rotations.

Namely, let X stand for any compact Abelian monothetic group with Haar measure λ_X. Then λ_X is the only (ergodic) invariant measure for any rotation $Tx=x+x_0$, with $\{nx_0:n\in\mathbb{Z}\}$ being dense in X. If $\chi:X\to\mathbb{S}^1$ is any nontrivial character of X then

$$\frac{1}{N}\sum_{n\leq N}\chi(T^{pn}x)\overline{\chi(T^{qn}x)}=\frac{1}{N}\sum_{n\leq N}\big(\chi(x_0)^{p-q}\big)^n\to 0$$

whenever $p\neq q$ (remembering that, by total ergodicity, $\chi(x_0)$ is not a root of unity). Since the dual group $\hat{X}$ is linearly dense in $C(X)$, all totally ergodic rotations are Möbius orthogonal by virtue of Theorem 1.3. But now take *any* topological system (Z,R) which is uniquely ergodic (with a unique invariant measure κ) and suppose that (Z,κ,R) and (X,λ_X,T) are measure-theoretically isomorphic. Since the eigenfunctions in $L^2(Z,\kappa)$ need not be continuous,[9] using DDKBSZ criterion does not seem to be possible. We can prove, however, (see [3]) that the AOP property holds. Moreover, once a uniquely ergodic system satisfies AOP, it must be orthogonal to any (bounded) multiplicative function [3]. Now, AOP being a measure-theoretic invariant must be satisfied in all uniquely ergodic models of totally ergodic rotations. In fact, AOP implies something which looks much stronger.

MW-strategy. The strong MOMO property. A dynamical system (X, T) is said to satisfy the strong MOMO[10] property (see [2]) if for each increasing sequence (b_k) of natural numbers, $b_{k+1} - b_k \to \infty$, and each $f \in C(X)$, we have

$$\lim_{K\to\infty} \frac{1}{b_K} \sum_{k<K} \left\| \sum_{b_k \le n < b_{k+1}} \mu(n) f \circ T^n \right\|_{C(X)} = 0.$$

We have the following:

Theorem 1.4 ([2]). *The following holds:*

(i) *The strong MOMO property of a topological systems (X, T) implies its Möbius orthogonality.*

(ii) *The strong MOMO property of a topological system (X, T) implies uniformity (in $x \in X$) in the definition of Möbius orthogonality property.*

(iii) *Sarnak's conjecture is equivalent to the fact that all zero-entropy systems satisfy the strong MOMO property.*

(iv) *No system with positive entropy satisfies the strong MOMO property.*[11]

Moreover, we have the following:

Proposition 1.5 ([2]). *Let $(Z, \mathcal{D}, \kappa, R)$ be any totally ergodic measure-theoretic system. If it satisfies the AOP property then each of its uniquely ergodic models satisfies the strong MOMO property. In particular, all such uniquely ergodic models are Möbius orthogonal.*

Short interval behavior. The reader certainly noticed that by putting $f = 1$ in the definition of the strong MOMO property, we obtain that, whenever $b_{k+1} - b_k \to \infty$,

$$\lim_{K\to\infty} \frac{1}{b_K} \sum_{k<K} \left| \sum_{b_k \le n < b_{k+1}} \mu(n) \right| = 0.$$

It is not hard to see that this property is equivalent to the following:

$$\frac{1}{M} \sum_{M \le m < 2M} \frac{1}{H} \left| \sum_{m \le n < m+H} \mu(n) \right| \to 0,$$

---

9 Topological system (Z, R) can be even topologically mixing, which excludes the possibility of continuous eigenfunctions.

10 The acronym comes from *Möbius Orthogonality of Moving Orbits*.

11 In [8] there are examples of positive entropy systems which are Möbius orthogonal. As Theorem 1.4 (iv) shows, this cannot happen for the strong MOMO property (assuming the Chowla conjecture).

whenever $H \to \infty$ and $H = o(M)$. This tells us that on a "typical" short interval (i.e., of length H) we have cancelations of 1s and -1s. This is a special property of the Möbius function proved in the breakthrough paper [30] by Matomäki and Radziwiłł in 2015. Together with the subsequent paper [31], it allowed in [3] to prove Sarnak's conjecture for all uniquely ergodic models of *finite* rotations[12] and all totally ergodic rotations. The fact, that all dynamical systems whose all invariant measures yield automorphisms with discrete spectrum satisfy Sarnak's conjecture was first proved in [20,21].

EW-strategy. Let us now pass to the second strategy which consists in the following. Using combinatorial properties of $\boldsymbol{\mu}$, we count on deriving special ergodic properties of the Furstenberg systems (see Section 3.1 for this crucial definition) of $\boldsymbol{\mu}$ (we recall that the Chowla conjecture predicts that there is only one Furstenberg system of $\boldsymbol{\mu}$ given by $\hat{\nu}_{\boldsymbol{\mu}^2}$, i.e., by the relatively independent extension of the Mirsky measure of $\boldsymbol{\mu}^2$). We then expect that Furstenberg systems will display "enough" of disjointness with at least a subclass of zero entropy systems to advance on MOC or else, expressing MOC as some ergodic property of them, when translated it back to $\boldsymbol{\mu}$, will tell us which new combinatorial properties of $\boldsymbol{\mu}$ are needed to prove Sarnak's conjecture. So, of course, the crucial question is whether Sarnak's conjecture can be expressed in the language of Furstenberg systems of $\boldsymbol{\mu}$. This was conjectured by Veech [38] and finally proved in [24] ($\Pi(\kappa)$ below stands for the Pinsker σ-algebra of $(X_{\boldsymbol{\mu}}, \mathcal{B}(X_{\boldsymbol{\mu}}), \kappa, S)$, i.e., the largest zero-entropy factor of the system, while $\pi_0 : \{-1, 0, 1\}^{\mathbb{Z}} \to \mathbb{R}$, $\pi_0(y) = y_0$).

Theorem 1.6 ([24]). *Sarnak's conjecture holds if and only if, for all Furstenberg systems κ of $\boldsymbol{\mu}$, we have $\pi_0 \perp L^2(\Pi(\kappa))$.*

(This theorem also holds in the logarithmic case.) The above put across the intuition that the Chowla conjecture in ergodic theory corresponds to the Bernoulli property (maximal chaos), while Sarnak's conjecture is rather related to the weaker property, namely the Kolmogorov property (K-property) of a measure-preserving system, meant "locally", i.e., for the single function π_0. It is classical in ergodic theory that the K-property is equivalent to K-mixing (called also uniform mixing). We will see in Section 5.4 that K-mixing property applied "locally" to π_0 yields a combinatorial condition on $\boldsymbol{\mu}$ equivalent to MOC. Roughly, this condition is about cancelations of $+1$s and -1s along larger and larger shifts of the sets of return times of blocks which can also be interpreted as the intuition that the multiplicative and additive structures of $\mathbb{N}$ are independent. While due to the MW-strategy many examples of classes of zero-entropy systems for which the MOC holds were given, to apply the DDKBSZ criterion, the arguments were provided ad hoc, depending on the class under consideration which shows its certain weakness.[13] In contrast, the EW-strategy aims at general results and some spectacular successes were obtained for the logarithmic versions of the Chowla and Sarnak's conjectures which we now present.

12 In other words, before 2015, we had no chances to prove Sarnak's conjecture, as we were already stuck in a relatively simple class of dynamical systems with zero entropy.

13 On the other hand, this strategy leads to the study of internal disjointness properties of measure-preserving systems which is of independent interest in ergodic theory.

The logarithmic versions of the Chowla and Sarnak's conjectures. When we replace Cesàro sums in (1.2) and (1.1) by their logarithmic versions,

$$\lim_{N\to\infty}\frac{1}{L_N}\sum_{n\le N}\frac{1}{n}f(T^nx)\boldsymbol{\mu}(n)=0$$

and

$$\lim_{N\to\infty}\frac{1}{L_N}\sum_{n\le N}\frac{1}{n}\boldsymbol{\mu}^{s_1}(n+a_1)\cdots\boldsymbol{\mu}^{s_k}(n+a_k)=0,$$

where $L_N=\sum_{n\le N}\frac{1}{n}$, we obtain *logarithmic* Sarnak's and the Chowla conjectures, respectively. The first striking result was obtained by Tao in 2015 (cf. the corresponding knowledge about MOC, i.e., Proposition 1.2):

Theorem 1.7 ([36]). *The logarithmic Chowla conjecture and the logarithmic Sarnak's conjecture are equivalent.*

Hence Sarnak's conjecture implies the logarithmic Chowla conjecture, and in [17] it is proved that the logarithmic Chowla conjecture implies the Chowla conjecture along a subsequence, hence Proposition 1.2 follows. The logarithmic Sarnak's conjecture is still open, but a significant progress has been achieved by Frantzikinakis and Host in [14] (in 2018). In that paper, the authors were able to relate logarithmic Furstenberg systems of the Möbius function (and many other strongly aperiodic multiplicative functions) to the theory of strongly stationary processes. They basically observed the following principle: either such a system is ergodic and then it must be $\hat{\nu}_{\boldsymbol{\mu}^2}$ or the corresponding Furstenberg system is disjoint from all ergodic systems. By proving new disjointness theorems in ergodic theory, this led them to the following remarkable result:

Theorem 1.8 ([14]). *All zero-entropy systems (X,T) for which the set $M^e(X,T)$ of ergodic invariant measures is countable are logarithmically Möbius orthogonal. In particular, all zero-entropy uniquely ergodic systems are logarithmically Möbius orthogonal.*

In Tao's proof of Theorem 1.7 an important step was to show the equivalence with the third condition which resembles the strong MOMO property (which we discussed above) uniformly with respect to all nil-rotations of a fixed nil-manifold. In fact, one of surprising consequences of Theorem 1.6, which also uses previous results by Tao [35] and Frantzikinakis [13] reduces the logarithmic Sarnak's conjecture to "merely" algebraic situation.

Theorem 1.9 ([24]). *The logarithmic Sarnak's conjecture holds if and only if all systems (X,T) for which each member of $M^e(X,T)$ yields a nil-system are logarithmically Möbius orthogonal.*

Sarnak's conjecture – where are we stuck? Returning to the original MOC, we would like first to notice that there is no result comparable to Frantizkinakis–Host's theorem (Theorem 1.8). In fact, only few general results concerning large classes of zero entropy systems which are Möbius orthogonal are known, namely, besides the already mentioned discrete spectrum case, MOC holds for systems whose all invariant measures yield rigid systems (with some arithmetic limitations on the arithmetics of rigidity sequences) [25] (the polynomial mean complexity characterization from [22] is for the logarithmic case).

A quick look at [24] shows that, at the moment, we are stuck with MOC since (surprisingly) we are not able to prove the strong MOMO property for zero-entropy algebraic automorphisms of the tori. We speak about very special unipotent systems, the simplest (nontrivial) one being $X = \mathbb{T}^2$ and $T(x,y) = (x, x+y)$.[14] It is almost obvious that such systems are Möbius orthogonal (apply, for example, the DDKBSZ criterion) but the situation changes dramatically if we try to prove the strong MOMO property for T (cf. Theorem 1.4 (iii)). In fact, the (potential) strong MOMO property applied to the function $(x,y) \mapsto e^{2\pi i y}$ gives the following:

$$\lim_{K\to\infty} \frac{1}{b_K} \sum_{k<K} \sup_{x\in\mathbb{T}} \left| \sum_{b_k \le n < b_{k+1}} \boldsymbol{\mu}(n) e^{2\pi i n x} \right| = 0,$$

for each sequence (b_k) satisfying $b_{k+1} - b_k \to \infty$. This reminds of a version of the classical Davenport's estimate[15] [7] but the sup inside makes it completely open (see also the discussion on the averaged form of Chowla conjecture in [31]).

2. ERGODIC THEORY – BASIC CONCEPTS

Given a standard Borel probability space $(X, \mathcal{B}, \mu)$, we consider *automorphisms* T of it and the quadruple $(X, \mathcal{B}, \mu, T)$ is often called a *dynamical system*. That is, $T : X \to X$ is invertible, bi-measurable,[16] and $\mu(A) = \mu(T^{-1}A) = \mu(TA)$ for each $A \in \mathcal{B}$. If S is another automorphism (acting on $(Y, \mathcal{C}, \nu)$) then S is a *factor* of T if there exists a measurable $\phi : X \to Y$ which is equivariant, i.e., $\phi \circ T = S \circ \phi$, pushing forward μ onto ν, i.e., $\phi_*(\mu) = \nu$.[17] Then, we obtain a (unique) disintegration of μ over ν:

$$\mu = \int_Y \mu_y \, d\nu(y) \tag{2.1}$$

with μ_y being probability measures on $(X, \mathcal{B})$ concentrated on $\phi^{-1}(y)$ (it is not hard to see that $T_*(\mu_y) = \mu_{Sy}$ for ν-a.e. $y \in Y$). If ϕ is invertible, then T and S are *isomorphic*.

An automorphism T is called *ergodic* if whenever $T^{-1}A = A$ (a.e.) then $\mu(A)$ equals zero or one. But in general, of course, T is nonergodic. In this situation, we consider its ergodic decomposition, which is simply the distintegration (2.1) of μ over the factor $(X/\mathcal{I}, \mathcal{I}, \mu|_{\mathcal{I}}, \mathrm{Id})$, where $\mathcal{I}$ stands for the σ-algebra of invariant sets.

---

14 The reader can notice that the ergodic measures for T yield either irrational rotations or finite (cyclic) rotations. There are uncountably many ergodic measures.

15 The estimate is $\sup_{t\in\mathbb{T}} |\sum_{n\le N} \boldsymbol{\mu}(n) e^{2\pi i n t}| = O(N/\log^A N)$ for each $A > 0$.

16 More precisely, if needed, we complete $\mathcal{B}$, and we can also assume that T is well defined only on a T-invariant subset $X_0 \subset X$ of full measure. Generally, in what follows we do not distinguish between sets, functions, etc., if they differ on a subset of measure zero.

17 Note that, setting $\mathcal{A} = \phi^{-1}(\mathcal{C})$, we can represent S as T acting on $(X/\mathcal{A}, \mathcal{A}, \mu|_{\mathcal{A}})$, where "points" in $X/\mathcal{A}$ are cosets of the relation on X of being indistinguishable by the sets of $\mathcal{A}$. By that reason, factors of T are identified with T-invariant sub-σ-fields of $\mathcal{B}$.

With T we can associate a unitary operator U_T, called Koopman operator, acting on $L^2(X, \mathcal{B}, \mu)$ by the formula $U_T(f) = f \circ T$. Studying the properties of Koopman operators is the *spectral theory* of dynamical systems. It is not hard to see that ergodicity means precisely that the only invariant functions of U_T are the constants. An automorphism T is called *weakly mixing* if its Cartesian square $T \times T$ acting on $(X \times X, \mathcal{B} \otimes \mathcal{B}, \mu \otimes \mu)$ is ergodic. This is equivalent to the fact that the *spectral measure*[18] of each zero mean f, i.e., of $f \in L_0^2(X, \mathcal{B}, \mu)$, is atomless. If the Fourier transforms of elements from L_0^2 vanish at infinity, we speak about *mixing* of T.

Given two automorphisms T and S acting on $(X, \mathcal{B}, \mu)$, $(Y, \mathcal{C}, \nu)$, respectively, by a *joining* between them we mean any measure ρ on $(X \times Y, \mathcal{B} \otimes \mathcal{C})$ with the coordinate projections μ, ν, respectively, and being $T \times S$-invariant. Denote the set of joinings by $J(T, S)$ which is always nonempty as $\mu \otimes \nu \in J(T, S)$. If T and S are additionally ergodic, we can ask about the subset $J^e(T, S)$ of ergodic joinings. This set is nonempty as the ergodic decomposition of any joining consists (a.e.) of joinings. A crucial concept here is that of disjointness introduced by Furstenberg **[15]** in 1967: we say that T and S are *disjoint*, $T \perp S$, if $J(T, S) = \{\mu \otimes \nu\}$. One should stress that to have disjointness of T and S, at least one of these automorphisms must be ergodic. Note also that if T and S are disjoint then they cannot have a nontrivial common factor (the converse to this implication does not hold). It is not hard to see that if T and S are *spectrally* disjoint, that is, if their maximal spectral types on the corresponding L_0^2-spaces are mutually singular, then $T \perp S$. This yields, in particular, classical disjointness results: identity Id is disjoint with all ergodic automorphisms, discrete spectrum automorphisms (i.e., those whose Koopman operators possess an orthonormal basis consisting of eigenvectors) are disjoint from weakly mixing automorphisms. For more classical examples of automorphisms and instances of disjointness, see **[16]**.

We can repeat the above concepts almost word for word in case of actions of the group $\mathbb{R}$ (or other locally compact Abelian groups) on $(X, \mathcal{B}, \mu)$, remembering that we consider only *measurable* actions of $\mathbb{R}$, called *flows* $\mathcal{T} = (T_t)_{t \in \mathbb{R}}$: the map

$$X \times \mathbb{R} \ni (x, t) \mapsto T_t x \in X$$

is measurable. This assumption yields that $t \mapsto U_{T_t} f$ is continuous in the strong topology for each $f \in L^2(X, \mathcal{B}, \mu)$. We also recall that the spectral measures of the corresponding Koopman representations are defined on the dual of the acting group, hence on $\mathbb{R}$ in case of flows.

Given $p \in \mathbb{R}^+$, the flow $\mathcal{T}_p := (T_{pt})_{t \in \mathbb{R}}$ is a *rescaling* of the original flow $\mathcal{T}$. It is not hard to see that the disjointness of the rescaling flows $\mathcal{T}_p$ and $\mathcal{T}_q$ $(0 < p < q)$ is equivalent to the disjointness of the time-p and time-q automorphisms, i.e., of T_p and T_q.

18 A spectral measure σ_f is a finite (nonnegative) Borel measure on the circle $\mathbb{S}^1$ whose Fourier transform is given by

$$\hat{\sigma}_f(n) := \int z^n \, d\sigma_f(z) = \langle U_T^n f, f \rangle = \int_X f(T^n x) \overline{f(x)} \, d\mu(x)$$

for each $n \in \mathbb{Z}$. Among the spectral measures there are the *maximal* ones (in the sense of absolute continuity of measures); each of them is a measure of *maximal spectral type*.

3. MEASURE-THEORETIC DYNAMICAL SYSTEMS – CONSTRUCTIONS AND EXAMPLES

3.1. Topological dynamics. Subshifts. Invariant measures for homeomorphisms

In topological dynamics we study homeomorphisms T acting on compact metric spaces X; (X,T) is a *topological dynamical system*. Such a system is called *transitive* if there is a point $x_0 \in X$ whose orbit $\{T^n x_0 : n \in \mathbb{Z}\}$ is dense. When all orbits are dense, the system is called *minimal*. The latter is equivalent to the fact that (X,T) has no proper subsystems. If A is a compact metric space, then $A^{\mathbb{Z}}$ considered with the product metric is also compact and $(A^{\mathbb{Z}}, S)$ with S the left *shift*, $S((x_n)_{n\in\mathbb{Z}}) = (x_{n+1})_{n\in\mathbb{Z}}$, is a topological dynamical system called the *full shift* (with the set of states A). Then every closed subset $X \subset A^{\mathbb{Z}}$ which is S-invariant yields a subsystem (X,S) of the full shift, so called *subshift*. By taking first $y = (y_n)_{n\in\mathbb{Z}} \in A^{\mathbb{Z}}$ and setting

$$X_y := \overline{\{S^n y : n \in \mathbb{Z}\}},$$

we obtain (X_y, S) a transitive subshift. In particular, we obtain $(X_{\boldsymbol{\mu}^2}, S)$, called the *square-free system*, where $X_{\boldsymbol{\mu}^2} \subset \{0,1\}^{\mathbb{Z}}$, and $(X_{\boldsymbol{\mu}}, S)$, called the *Möbius subshift*, where $X_{\boldsymbol{\mu}} \subset \{-1,0,1\}^{\mathbb{Z}}$.

The notions of a (topological) factor and isomorphism (conjugacy) are defined similarly to the measure-theoretic category remaining in the class of continuous maps. An important invariant of topological conjugacy is that of entropy $h(T) = h(X,T)$. We refer the reader to [39] for general definitions, however, if A is finite and (X,S) is a subshift, then

$$h(X,S) = \lim_{N\to\infty} \frac{1}{N} \log\left|\mathcal{L}(X) \cap A^N\right|,$$

where $\mathcal{L}(X)$ is the *language* of X, i.e., the set of all words (blocks) appearing in $x \in X$. Clearly, if $X = X_y$, it is enough to compute only words appearing in y.

A topological dynamical system (X,T) yields measure-theoretic dynamical systems through Borel T-invariant measures: if $M(X,T)$ stands for the set of Borel T-invariant measures and $\nu \in M(X,T)$, then it yields a measure-theoretic dynamical system $(X, \mathcal{B}, \nu, T)$, where $\mathcal{B} = \mathcal{B}(X)$ denotes the σ-algebra of Borel subsets of X. We will detail slightly on that. Let $M(X)$ denote the space of probability measures on X. With the weak-$*$-topology, it becomes a metrizable compact space: $\mu_n \to \mu$ if and only $\lim_{n\to\infty} \int_X f\,d\mu_n = \int_X f\,d\mu$ for each $f \in C(X)$. If $x \in X$ then the measures of the form $\frac{1}{N}\sum_{n<N} \delta_{T^n x}$ are called *empiric* measures. Note that any limit ν of a convergent subsequence of empiric measures,

$$\frac{1}{N_k} \sum_{n<N_k} \delta_{T^n x} \to \nu, \tag{3.1}$$

must be a T-invariant measure (one says also that x is *quasi-generic* for ν). This is the classical Krylov–Bogoljubov theorem which tells us that the set $M(X,T)$ is nonempty. It automatically yields that the set $M^e(X,T)$ of ergodic measures is also nonempty. Note that another way to obtain an invariant measure is to change the Cesàro way of summation into the logarithmic one,

$$\frac{1}{L_{N_k}}\sum_{n<N_k}\frac{1}{n}\delta_{T^n x}\to\nu,$$

where the limit measure is also T-invariant. We say that $x\in X$ is *generic for* $\nu\in M(X,T)$ *along* (N_k) if (3.1) holds. For example, $\boldsymbol{\mu}^2$ is generic (along the whole sequence of natural numbers) for the so-called Mirsky measure $\nu_{\boldsymbol{\mu}^2}$ and if the Chowla conjecture holds then $\boldsymbol{\mu}$ is generic for the relatively independent extension $\hat{\nu}_{\boldsymbol{\mu}^2}$ of the Mirsky measure, where

$$\hat{\nu}_{\boldsymbol{\mu}^2}(C)=\frac{1}{2^{\mathrm{supp}(C)}}\nu_{\boldsymbol{\mu}^2}(C^2)$$

for each block C of -1, 0, 1s. We recall that each ergodic measure has a generic point. The set of measures (called also "visible") for which x is quasi-generic is denoted by $V(x)$ and it is compact. Note that either $|V(x)|=1$ (we say then that x is generic for this unique measure) or $V(x)$ is uncountable as it is also connected. Note also that

$$M^e(X,T)\subset V(X,T):=\bigcup_{x\in X}V(x).$$

If $y\in A^{\mathbb{Z}}$ and $\nu\in V(y)$ then the measure-theoretic system $(X_y,\mathcal{B}(X_y),\nu,S)$ is called a *Furstenberg system of* y.

We can introduce similar notions (and prove similar facts) for the logarithmic way of averaging. In general, there is no relation between $V(x)$ and $V^{\log}(x)$ unless x is generic for $\nu\in V(x)$ (it is then logarithmically generic for the same measure).

A topological system (X,T) is *uniquely ergodic* if $|M(X,T)|=1$. The unique invariant measure is then necessarily ergodic. Uniquely ergodic and minimal systems are called *strictly ergodic*. The classical Jewett–Krieger theorem tells us that each ergodic system has a strictly ergodic model.

3.2. Flows, special flows, change of time

Let us first see how, given a flow, to produce new flows with the same orbits but (potentially) representing completely different (even disjoint!) dynamics. Assume that $\mathcal{R}=(R_t)$ is a flow on $(Z,\mathcal{D},\kappa)$ and let $v:Z\to\mathbb{R}$, $v\geq\varepsilon_0>0$ and $v\in L^1(Z,\mathcal{D},\kappa)$. Then, for κ-a.e. $z\in Z$ and all $t\in\mathbb{R}$, there is a unique solution $u=u(t,z)$ of

$$\int_0^u v(R_s z)\,ds=t.$$

Then we set $\tilde{R}^v_t(z)=R_{u(t,z)}(z)$ and obtain a new flow $\widetilde{\mathcal{R}^v}=(\tilde{R}^v_t)$ which preserves the measure $(\frac{v}{\int v\,d\kappa})d\kappa$ for which $u(t,z)$ is a cocycle. On the other hand, $(u,x)\mapsto\int_0^u v(R_s x)\,ds$ defines a cocycle for $\mathcal{R}$. If $v':Z\to\mathbb{R}$ is another time change and, for some measurable $\xi:Z\to\mathbb{R}$,

$$\int_0^u v(R_s z)\,ds=\int_0^u v'(R_s z)\,ds-\xi(z)+\xi(R_u z)$$

for κ-a.e $z\in Z$ and all $u\in\mathbb{R}$ (that is, the two cocycles for $\mathcal{R}$ are cohomologous), then the two time changes $\widetilde{\mathcal{R}^v}$, $\widetilde{\mathcal{R}^{v'}}$ are isomorphic. If $v'=c$ is additionally a constant (that is, the cocycle given by v' is a quasi-coboundary), then $\widetilde{\mathcal{R}^{v'}} = \widetilde{\mathcal{R}^c} = (R_{t/c})_{t\in\mathbb{R}}$, so this time change is isomorphic to a rescaling of the original flow $\mathcal{R}$.

We now invoke a construction transforming $\mathbb{Z}$-actions (automorphisms) into $\mathbb{R}$-actions (flows) which is a kind of inducing representation.[19] Assume that $(X, \mathcal{B}, \mu, T)$ is a dynamical system and let $f : X \to \mathbb{R}^+$ be an L^1-function. Consider the probability space $(X^f, \mathcal{B}^f, \mu^f)$, where

$$X^f = \{(x,r) \in X \times \mathbb{R} : x \in X, 0 \le r < f(x)\}$$

with $\mathcal{B}^f$ being the restriction of the product σ-algebra, and $\mu^f := (\mu \otimes \lambda_{\mathbb{R}})|_{X^f} / \int_X f\, d\mu$. We now define the *special flow over T under the roof function f* by setting

$$T_t^f(x,r) = \big(T^n x, r + t - f^{(n)}(x)\big),$$

where $n \in \mathbb{Z}$ is unique such that

$$f^{(n)}(x) \le r + t < f^{(n+1)}(x)$$

and $f^{(n)}(x) = f(x) + f(Tx) + \cdots + f(T^{n-1}x)$ if $n > 0$, $f^{(0)}(x) = 0$ and $f^{(m+n)}(x) = f^{(m)}(x) + f^{(n)}(T^m x)$ for $m, n \in \mathbb{Z}$.

If $f = 1$ then we speak about the *suspension flow* $\hat{T}$ over T,

$$\hat{T}_t(x,r) = \big(T^{[t+r]}x, \{t + r\}\big),$$

for $(x,r) \in X \times [0,1)$. Note that for $k \in \mathbb{N}$,

$$\int_0^k f\big(\hat{T}_s(x,0)\big)\, ds = \int_0^k f\big(T^{[s]}x\big)\, ds = f^{(k)}(x)$$

allows us to see the special flow T^f as a time change of the suspension flow over T. It follows that, given two special flows over the same automorphism T, we can obtain one from the other by a time change.

The Kakutani–Ambrose theorem tells us that each flow has a special representation. Representing a flow as a special flow (over a "known" automorphism) is a useful operation, and finding T, especially in the smooth case, leads to seeking a good transversal to orbits of the original flow. For example, in the case of smooth flows on surfaces it often leads to the study of special flows over interval exchange transformations and interesting roof functions having "controllable" singularities.

4. RATNER'S QUESTION, MW-STRATEGY, AND MOC FOR SMOOTH TIME CHANGES OF HOROCYCLE FLOWS

4.1. Horocycle flows and MOC

One of the most important zero-entropy classes in dynamics is given by horocycle flows whose definition we now recall. Let $\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$ be a discrete subgroup with finite

19 The reader can check that the Koopman representation of the special flow defined below is indeed the genuine induced representation of the Koopman operator associated to the automorphism.

covolume, in fact, we consider only the case Γ is cocompact, so that the homogeneous space $M = \Gamma\backslash \mathrm{PSL}_2(\mathbb{R})$ is compact and then the system is uniquely ergodic. Let us consider the corresponding *horocycle flow* $(h_t)_{t\in\mathbb{R}}$ and the *geodesic flow* $(g_t)_{t\in\mathbb{R}}$ on M given by

$$h_t(\Gamma x) = \Gamma \cdot \left(x \cdot \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}\right) \quad \text{and} \quad g_t(\Gamma x) = \Gamma \cdot \left(x \cdot \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{t} \end{bmatrix}\right).$$

Since

$$g_s h_t g_s^{-1} = h_{e^{-2s}t} \quad \text{for all } s, t \in \mathbb{R}, \tag{4.1}$$

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 particular, all positive time automorphisms are isomorphic). In 2011, Bourgain, Sarnak, and Ziegler proved the following:

Theorem 4.1 ([4]). *Each time-t automorphism h_t is Möbius orthogonal.*

The main idea is to use the MW-strategy, and to show that, in fact, these time-t automorphisms are orthogonal to *any* (zero mean, bounded) multiplicative function. It works here because of the famous Ratner's theory: given $x \in \mathrm{PSL}_2(\mathbb{R})$, any point $(\Gamma x, \Gamma x)$ is generic for a measure ρ (which must be a joining by unique ergodicity: $\rho \in J(T^p, T^q)$, where $T = h_t$) and, moreover, this joining is ergodic and of algebraic nature. As shown in [4], this algebraic nature yields that, perhaps except for finitely many primes, we must obtain the product measure.[20] The proof really depends on some algebraic properties of horocycle flows and because of that M. Ratner asked in 2013 what happens if we (smoothly) change time and study MOC in this class.

4.2. Time changes of horocycle flows and MOC

In general, especially for flows which are mixing, it is difficult to decide whether or not they are disjoint. Horocycle flows are mixing and so are their smooth time changes. In 1983, M. Ratner [32] discovered a new property of horocycle flows which basically gave a quadratic way of divergence of distinct orbits of nearby points, which allows one to observe some drift of these orbits. This geometric property has surprisingly strong rigidity joining consequences. Ratner also showed that smooth time changes of horocycle flows enjoy this divergence property [33]. It took more than 20 years to understand how to variate her original property (keeping the joining consequences) to now commonly called Ratner's properties, to observe quantitatively the drift phenomenon also beyond the horocyclic world, in particular to see it in dimension 2 (e.g., for some smooth flows on surfaces). A kind of a breakthrough new disjointness criterion has been recently proved in [26]. It is tailored for flows (with a Ratner's property) having different speed of divergence (polynomial or subpolynomial) of distinct orbit of close points. It fits to nontrivial smooth time changes $(\tilde{h}_t^v)$ of horocycle flows as one of the main results of [26] shows:

---

20 This might suggest that we have the AOP property, but, in fact, as noticed in [3], (4.1) applied to some compact regions implies that AOP fails for horocycle flows.

Theorem 4.2 ([26]). *Assume that the cocycle determined by a positive $v \in W^6(M)$ has a nontrivial support outside of the discrete series*[21] *and is not a quasi-coboundary. Then, for any real numbers $0 < p < q$, the rescalings $(\tilde{h}^v_{pt})_{t\in\mathbb{R}}$ and $(\tilde{h}^v_{qt})_{t\in\mathbb{R}}$ are disjoint.*

The situation looks a little bit paradoxical as, for the horocycle flows themselves, we know that they are Möbius orthogonal, but the problem of whether the convergence in (1.2) is uniform (in $x \in M$) is open, the strong MOMO property is open, and we also do not know whether the Möbius orthogonality takes place in *all* uniquely ergodic models of horocycle flows. On the other hand, when we change time (as above), for the flows whose dynamics intuitively become more complicated, the answers to these questions are simply positive due to Theorem 4.2, Proposition 1.5, and Theorem 1.4 (ii).

Using the same disjointness criterion, other disjointness results concerning some mixing locally Hamiltonian flows (on surfaces), considered most often in their special representations over irrational rotations (Arnol'd special flows) are proved in [26] to enjoy similar internal disjointness properties. Hence, the Möbius orthogonality for them is also established.

5. SARNAK'S CONJECTURE AND FURSTENBERG SYSTEMS

It is not clear at all that the MOC can be expressed in terms of Furstenberg systems of the Möbius function. In fact, following [24], we will consider a general problem in which $\boldsymbol{\mu}$ is replaced by a function $\boldsymbol{u} : \mathbb{N} \to \mathbb{D}$ (the unit disc). We want to characterize those $\boldsymbol{u}$ which are orthogonal to all zero-entropy systems,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n\leq N} f(T^n x)\boldsymbol{u}(n) = 0, \tag{5.1}$$

for each (X, T) of zero entropy, all $f \in C(X)$, and $x \in X$. One can now wonder what is special in the zero (topological) entropy class. For that we need to recall some classical facts, namely the variational principle, which tells us that $h(X, T) = 0$ if and only if the measure-theoretic entropy of each T-invariant measure is zero. By a convexity property of the entropy, this is still equivalent to the fact that all ergodic measures have zero entropy. In this way, we replaced the original assumption on (X, T) by an assumption on the *measure-theoretic* systems determined by invariant measures. More than that, while thinking about problem (5.1), we only care about properties of systems determined by *visible* measures $\mu \in V(X, T)$. Finally, one can wonder what is special in the class of measure-theoretic systems with zero entropy. Classical ergodic theory tells us that this is a class which is closed under taking joinings and factors and for each automorphism $(Z, \mathcal{D}, \kappa, R)$ there exists a largest factor $\Pi(\kappa) \subset \mathcal{D}$ of zero entropy, called the Pinsker factor of R. All this leads us to the concept of a characteristic class and the problem of orthogonality to such.

21 This assumption is dropped in [11]. Flaminio and Forni used more directly Ratner's work [33] and show that the cocycles determined by v and $v \circ g_r$, where $r = -\frac{1}{2}\log(q/p)$ are not jointly cohomologous.

5.1. Characteristic classes and the problem of orthogonality

A class $\mathcal{F}$ of automorphisms (it is implicit that this class is closed under isomorphism) is called *characteristic* if it is closed under taking (countable) joinings and factors. Classical classes, like the zero-entropy class, the class of systems with discrete spectrum, and the class of automorphisms which are distal, are characteristic classes (many more classes are listed in [24]). Once $\mathcal{F}$ is fixed, we consider the class $\mathscr{C}_{\mathcal{F}}$ of those topological systems (X,T) for which $(X,\mathcal{B}(X),\nu,T)\in\mathcal{F}$ for each $\nu\in V(X,T)$. The following theorem establishes the most useful ergodic properties following the concept of a characteristic class.

Theorem 5.1 ([24]). *Assume that $\mathcal{F}$ is a characteristic class then, for each automorphism $(Z,\mathcal{D},\kappa,R)$, there exists a largest factor $\mathcal{D}_{\mathcal{F}}\subset\mathcal{D}$ belonging to $\mathcal{F}$. Moreover, any joining of $(Z,\mathcal{D},\kappa,R)$ with an automorphism from $\mathcal{F}$ is uniquely determined by its restriction to a joining with $\mathcal{D}_{\mathcal{F}}$.*

With each class $\mathcal{F}$, we can associate the class $\mathcal{F}_{\mathrm{ec}}$ consisting of the automorphisms whose all ergodic components are in $\mathcal{F}$.

Proposition 5.2 ([24]). *If $\mathcal{F}$ is characteristic, then also $\mathcal{F}_{ec}$ is characteristic.*

In general, we then have $\mathscr{C}_{\mathcal{F}}\subset\mathscr{C}_{\mathcal{F}_{\mathrm{ec}}}$, but the reader can check that if $\mathcal{F}$ is the zero-entropy class then we have equality. Moreover,

$$\mathscr{C}_{\mathcal{F}_{\mathrm{ec}}}=\big\{(X,T):\big(X,\mathcal{B}(X),\nu,T\big)\in\mathcal{F}\text{ for each }\nu\in M^e(X,T)\big\}.$$

The zero-entropy class turns out to be special in the family of characteristic classes:

Proposition 5.3 ([24]). *The zero-entropy class is the largest proper characteristic class.*

It is also shown in [24] that there exists the smallest nontrivial characteristic class (it consists of all identities of standard Borel probability spaces).

5.2. Orthogonality to characteristic classes. Veech's conjecture

Given a class $\mathscr{C}$ of topological systems, we can now consider the problem of orthogonality of $\boldsymbol{u}:\mathbb{N}\to\mathbb{D}$ to $\mathscr{C}$, that is,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}f(T^nx)\boldsymbol{u}(n)=0 \tag{5.2}$$

for each $(X,T)\in\mathscr{C}$, all $f\in C(X)$ and $x\in X$. If orthogonality takes place, we write $\boldsymbol{u}\perp\mathscr{C}$. The central result of [24] is the following:

Theorem 5.4 ([24]). *Assume that $\mathcal{F}$ is a characteristic class and $\boldsymbol{u}:\mathbb{N}\to\mathbb{D}$. Then $\boldsymbol{u}\perp\mathscr{C}_{\mathcal{F}_{ec}}$ if and only if*

$$\pi_0\perp L^2\big(\big(\mathcal{B}(X_{\boldsymbol{u}}),\kappa\big)_{\mathcal{F}_{ec}}\big)\quad\textit{for each Furstenberg system }\kappa\in V(\boldsymbol{u}). \tag{5.3}$$

Remark 5.5. Condition (5.3) will be called the *Veech condition* as Veech formulated it in [38], in a form of a conjecture, as a statement equivalent to MOC in case of $\boldsymbol{u}=\boldsymbol{\mu}$ and $\mathcal{F}$

equal to the (measure-theoretic) zero-entropy class. Theorem 5.4 proves in particular Veech's conjecture but, clearly, goes beyond it.

In the subsequent subsections we will say a few words on the tools that are employed for the proof of Theorem 5.4 and briefly indicate some consequences of it.

5.3. Proof of Theorem 5.4

The sufficiency in Theorem 5.4 follows from a more general result:

Theorem 5.6 ([24]). *If $\boldsymbol{u} : \mathbb{N} \to \mathbb{D}$ satisfies the Veech condition with respect to a characteristic class $\mathscr{F}$ then $\boldsymbol{u} \perp \mathscr{C}_{\mathscr{F}}$.*

The proof of this theorem is purely ergodic, belongs to joining theory, and is based on a fundamental non-disjointness lemma [29].[22]

The necessity requires more tools. The first relies on the existence of the so-called Hansel's models, being a counterpart of the classical Jewett–Kieger theorem in the nonergodic case. Namely, if $(Z, \mathcal{D}, \nu, R)$ is a measure-theoretic dynamical system and we fix a set of full measure of ergodic components then a Hansel model [18] of it is any topological system (X, T) for which there exists $\nu \in M(X, T)$ yielding a measure-theoretic isomorphic copy of R and such that *each* $x \in X$ is generic for one of the chosen ergodic components. The next major step is a new lifting lemma[23] (going largely beyond the context considered in [6]) on quasi-generic points for joinings, and tailored to be applicable for the strong MOMO property:

Lemma 5.7 ([24]). *Assume that (Y, S) and (X, T) are topological systems. Let $\nu \in M(X, T)$, $u \in Y$ be generic along an increasing sequence (N_m) for $\kappa \in M(Y, S)$, and $\rho \in J(\kappa, \nu)$. Then there exist a sequence $(x_n) \subset X$ and a subsequence (N_{m_ℓ}) such that $(S^n u, x_n)$ is generic along (N_{m_ℓ}) for ρ and the set $\{n \geq 0 : x_{n+1} \neq T x_n\}$ is of the form (b_k) with $b_{k+1} - b_k \to \infty$ when $k \to \infty$.*

We then use some joining techniques and Lemma 5.7 to Hansel models of the largest $\mathscr{F}_{\mathrm{ec}}$-factors of Furstenberg systems of $\boldsymbol{u}$. Finally, the reason why we use $\mathscr{F}_{\mathrm{ec}}$ (and not $\mathscr{F}$ itself) is that the orthogonality of $\boldsymbol{u}$ to $\mathscr{C}_{\mathscr{F}_{\mathrm{ec}}}$ is *equivalent* to the strong MOMO property (relative to $\boldsymbol{u}$) of all systems in $\mathscr{C}_{\mathscr{F}_{\mathrm{ec}}}$ (such a result, in full generality, is unknown for $\mathscr{F}$).

5.4. Some consequences of Theorem 5.4

We now come back to the problem of orthogonality of $\boldsymbol{u} : \mathbb{N} \to \mathbb{D}$ to the zero (topological) entropy class (MOC is a particular case of this). In this case, Theorem 5.4 describes a kind of relative Kolmogorov property which, by some ergodic considerations, can

---

22 Veech in [38], only for the zero entropy class, gives a rather complicated proof based on the concept of quasi-factors by Glasner and Weiss. For this particular class, the proof is also implicit in [1].

23 The lemma is also valid for the logarithmic way of averaging, which seems to be the first result of that type in the literature.

be replaced by so called (relative) K-mixing. We will now write a combinatorial reformulation of the latter property assuming (for sake of simplicity) that there is only one Furstenberg system of $\boldsymbol{u}$:

Corollary 5.8 ([24]). *If $\boldsymbol{u} : \mathbb{N} \to \mathbb{D}$ is generic then $\boldsymbol{u}$ is orthogonal to all zero-entropy systems if and only if*

$$\lim_{m\to\infty} \lim_{N\to\infty} \left| \frac{1}{N} \sum_{n\le N} \boldsymbol{u}(n) 1_{\boldsymbol{u}(m+n),\boldsymbol{u}(m+n+1),\dots,\boldsymbol{u}(m+n+\ell-1)\in C} \right| = 0$$

uniformly in $\ell \ge 1$ and in C, a set of blocks of length ℓ.

The above corollary is, of course, about cancelations of $+1$s and -1s along larger and larger shifts of return times to a fixed set of blocks (of a fixed length). By a rather standard argument, it can be replaced with a *conditional* cancelation phenomenon for a single "typical" block.

Another consequence of Theorem 5.4 is a purely ergodic proof of the so-called averaged Chowla property shown first (even in the quantitative version) in [31] for the Möbius function: for each $\boldsymbol{u} : \mathbb{N} \to \mathbb{D}$, for which all circle rotations satisfy the strong MOMO (relative to $\boldsymbol{u}$) property,[24] we have

$$\lim_{H\to\infty} \frac{1}{H^k} \sum_{h_1,\dots,h_k\le H} \lim_{k\to\infty} \frac{1}{N_k} \left| \sum_{n\le N_k} \boldsymbol{u}(n) \prod_{i=1}^{k} c_i(n+h_i) \right| = 0$$

for all sequences $c_i : \mathbb{N} \to \mathbb{D}$, $i = 1, \dots, k$.

The strength of Theorem 5.4 also follows from the fact that it is valid in the logarithmic context which is better understood. As we have already mentioned in the introduction, using it together with some earlier results by Tao and Frantzikinakis yields Theorem 1.9.

FUNDING

This work was partially supported by Narodowe Centrum Nauki grant UMO-2019/33/B/ST1/00364.

REFERENCES

[1] 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), 2899–2944.

[2] H. El Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk, and T. de la Rue, Möbius disjointness for models of an ergodic system and beyond. *Israel J. Math.* **228** (2018), 707–751.

[3] H. El Abdalaoui, M. Lemańczyk, and T. de la Rue, Automorphisms with quasi-discrete spectrum, multiplicative functions and average orthogonality along short intervals. *Int. Math. Res. Not. IMRN* **2017** (2017), no. 14, 4350–4368.

24 We recall that this takes place for the Möbius function.

[4] J. Bourgain, P. Sarnak, and T. Ziegler, Disjointness of Möbius from horocycle flows. In *From Fourier and number theory to radon transforms and geometry, in memory of Leon Ehrenpreiss*, pp. 67–83, Dev. Math. 28, Springer, 2012.

[5] S. Chowla, *The Riemann hypothesis and Hilbert's tenth problem*. Math. Appl. 4, Gordon and Breach Science Publishers, New York, 1965.

[6] J.-P. Conze, T. Downarowicz, and J. Serafin, Correlation of sequences and of measures, generic points for joinings and ergodicity of certain cocycles. *Trans. Amer. Math. Soc.* **369** (2017), 3421–3441.

[7] H. Davenport, On some infinite series involving arithmetical functions. II. *Quart. J. Math. Oxford* **8** (1937), 313–320.

[8] T. Downarowicz and J. Serafin, Almost full entropy subshifts uncorrelated to the Möbius function. *Int. Math. Res. Not. IMRN* **2019** (2019), no. 11, 3459–3472.

[9] M. Einsiedler and T. Ward, *Ergodic theory with a view towards number theory*. Grad. Texts in Math. 259, Springer, London, 2011.

[10] S. Ferenczi, J. Kułaga-Przymus, and M. Lemańczyk, Sarnak's conjecture – what's new. In *Ergodic theory and dynamical systems in their interactions with arithmetics and combinatorics, CIRM Jean-Morlet Chair, Fall 2016*, edited by S. Ferenczi, J. Kułaga-Przymus, and M. Lemańczyk, p. 418, Lecture Notes in Math. 2213, Springer, 2018.

[11] L. Flaminio and G. Forni, Orthogonal powers and Möbius conjecture for smooth time-changes of horocycle flows. *Electron. Res. Announc. Math. Sci.* **26** (2019), 16–23.

[12] L. Flaminio, K. Frączek, J. Kułaga-Przymus, and M. Lemańczyk, Approximate orthogonality of powers for ergodic affine unipotent diffeomorphisms. *Studia Math.* **244** (2019), 43–97.

[13] N. Frantzikinakis, *Ergodicity of the Liouville system implies the Chowla conjecture*. Discrete Analysis 2017, paper 8, p. 23.

[14] N. Frantzikinakis and B. Host, The logarithmic Sarnak conjecture for ergodic weights. *Ann. of Math. (2)* **187** (2018), no. 3, 869–931.

[15] H. Furstenberg, Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Syst. Theory* **1** (1967), 1–49.

[16] E. Glasner, *Ergodic theory via joinings*. American Mathematical Society, Providence, 2003.

[17] A. Gomilko, D. Kwietniak, and M. Lemańczyk, Sarnak's conjecture implies the Chowla conjecture along a subsequence. In *Ergodic theory and dynamical systems in their interactions with arithmetics and combinatorics*, pp. 237–247, Lecture Notes in Math. 2213, Springer, Cham, 2018.

[18] G. Hansel, Strict uniformity in ergodic theory. *Math. Z.* **135** (1974), 221–248.

[19] A. Hildebrand, *Introduction to analytic number theory*. http://www.math.uiuc.edu/~ajh/531.fall05/.

[20] W. Huang, Z. Wang, and X. Ye, Measure complexity and Möbius disjointness. *Adv. Math.* **347** (2019), 827–858.

[21] W. Huang, Z. Wang, and G. Zhang, Möbius disjointness for topological models of ergodic systems with discrete spectrum. *J. Mod. Dyn.* **14** (2019), 277–290.

[22] W. Huang, L. Xu, and X. Ye, Polynomial mean complexity and logarithmic Sarnak conjecture. 2020, arXiv:2009.02090.

[23] H. Iwaniec and E. Kowalski, *Analytic number theory*. Amer. Math. Soc. Colloq. Publ. 53, American Mathematical Society, Providence, RI, 2004.

[24] A. Kanigowski, J. Kułaga-Przymus, M. Lemańczyk, and T. de la Rue, On arithmetic functions orthogonal to deterministic sequences. 2021, arXiv:2105.11737.

[25] A. Kanigowski, M. Lemańczyk, and M. Radziwiłł, Rigidity in dynamics and Möbius disjointness. *Fund. Math.* **255** (2021), 309-336.

[26] A. Kanigowski, M. Lemańczyk, and C. Ulcigrai, On disjointness of some parabolic flows. *Invent. Math.* **221(1)** (2020), 1–111.

[27] I. Kátai, A remark on a theorem of H. Daboussi. *Acta Math. Hungar.* **47** (1986), no. 1–2, 223–225.

[28] J. Kułaga-Przymus and M. Lemańczyk, Sarnak's conjecture from the ergodic theory point of view. In *Encyclopedia for Complexity and Systems Science*, published online 2021, arXiv:2009.04757.

[29] M. Lemańczyk, F. Parreau, and J.-P. Thouvenot, Gaussian automorphisms whose ergodic self-joinings are Gaussian. *Fund. Math.* **164** (2000), 253–293.

[30] K. Matomäki and M. Radziwiłł, Multiplicative functions in short intervals. *Ann. of Math. (2)* **183** (2016), 1015–1056.

[31] K. Matomäki, M. Radziwiłł, and T. Tao, An averaged form of Chowla's conjecture. *Algebra Number Theory* **9** (2015), 2167–2196.

[32] M. Ratner, Horocycle flows, joinings and rigidity of products. *Ann. of Math. (2)* **118** (1983), 277–313.

[33] M. Ratner, Rigid reparametrizations and cohomology for horocycle flows. *Invent. Math.* **88** (1987), no. 2, 341–374.

[34] P. Sarnak, *Three lectures on the Möbius function, randomness and dynamics*. http://publications.ias.edu/sarnak/.

[35] T. Tao, *The Chowla and the Sarnak conjecture*, What's new. http://terrytao.wordpress.com/2012/10/14/the-chowla-conjecture-and-the-sarnak-conjecture/.

[36] T. Tao, Equivalence of the logarithmically averaged Chowla and Sarnak conjectures. In *Number theory – diophantine problems, uniform distribution and applications: Festschrift in honour of Robert F. Tichy's 60th birthday*, edited by C. Elsholtz and P. Grabner, pp. 391–421, Springer, Cham, 2017.

[37] T. Tao, *The logarithmically averaged and non-logarithmically averaged Chowla conjectures*. https://terrytao.wordpress.com/2017/10/20/the-logarithmically-averaged-and-non-logarithmically-averaged-chowla-conjectures/.

[38] W. A. Veech, *Möbius dynamics*. Unpublished lecture notes, Spring semester 2016.

[39] P. Walters, *An introduction to ergodic theory*. Grad. Texts in Math. 79, Springer, New York, 1982.

MARIUSZ LEMAŃCZYK
Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, Chopin street 12/18, 87-100 Toruń, Poland, mlem@mat.umk.pl