



HAL
open science

Sarnak's Möbius disjointness for dynamical systems with singular spectrum and dissection of Möbius flow

Houcein El Abdalaoui, Mahesh Nerurkar

► **To cite this version:**

Houcein El Abdalaoui, Mahesh Nerurkar. Sarnak's Möbius disjointness for dynamical systems with singular spectrum and dissection of Möbius flow. 2020. hal-02867657

HAL Id: hal-02867657

<https://hal.science/hal-02867657>

Preprint submitted on 15 Jun 2020

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 Möbius disjointness for dynamical systems with singular spectrum and dissection of Möbius flow

El Houcein El Abdalaoui
Department of Mathematics
Universite of Rouen, France

and

Mahesh Nerurkar
Department of Mathematics
Rutgers University, Camden NJ 08102 USA

Abstract

It is shown that Sarnak's Möbius orthogonality conjecture is fulfilled for the compact metric dynamical systems for which every invariant measure has singular spectra. This is accomplished by first establishing a special case of Chowla conjecture which gives a correlation between the Möbius function and its square. Then a computation of W. Veech, followed by an argument using the notion of 'affinity between measures', (or the so called 'Hellinger method'), completes the proof. We further present an unpublished theorem of Veech which is closely related to our main result. This theorem asserts if for any probability measure in the closure of the Cesaro averages of the Dirac measure on the shift of the Möbius function, the first projection is in the orthocomplement of its Pinsker algebra then Sarnak's Möbius disjointness conjecture holds. Among other consequences, we obtain a simple proof of Matomäki-Radziwiłł-Tao's theorem and Matomäki-Radziwiłł's theorem on the correlations of order two of the Liouville function.

Keywords Topological dynamics, singular spectrum, Sarnak's Möbius disjointness conjecture, Pinsker algebra, affinity and Hellinger distance.

Mathematics Subject Classification (2000) 37A30, 54H20, 47A35

1 Introduction

The purpose of this article is to establish the validity of Sarnak's Möbius orthogonality conjecture for the dynamical systems with singular spectra. More precisely, this means that Sarnak's conjecture holds for any compact metric dynamical system for which every invariant measure has a singular spectrum. We remark that it is observed in [1] that, Chowla conjecture of order two implies that Möbius disjointness holds for any dynamical system with singular spectra. We also recall that Sarnak in his seminal paper, states that Chowla conjecture on the multiple correlations of Möbius function implies Sarnak's Möbius disjointness conjecture. For more details about the connection between Chowla conjecture and Möbius orthogonality conjecture, we refer to [3].

Now we briefly describe our strategy of the proof. We first obtain a special case of Chowla conjecture which gives a correlation between the Möbius function and its square, (Theorem 3.8). Next, using a computation of W. Veech we establish a key Proposition (Proposition 3.22). Another ingredient to the proof involves understanding Pinsker sigma algebras for any - (what we call) - potential spectral

measure for the ‘Möbius flow’, (i.e. for the shift dynamical system generated by the Möbius function). In this understanding, a ‘generating partition’ type result of Rokhlin and Sinai plays an important role. Our argument also uses existence of -(what we call)- ‘the Chowla-Sarnak-Veech measure’, (Theorem 3.21). Once this is done, the rest of the argument uses the notion of ‘affinity between measures’, or the ‘Hellinger method’. This method was developed by Below and Losert [5] and may go back to Coquet-Mendes-France-Kamae [9]. We shall recall the basic ingredients of this method along with results on the spectral measure associated with a sequence in Section 2. We shall also present a proof of an unpublished theorem due to W. Veech, (Theorem 3.24). In the light of the proof of this Theorem the underlying issues become much clearer and will help us to formulate our result in the language of spectral isomorphisms, (Corollary 4.7).

Before beginning formally with the proof, we comment summarizing various cases where this conjecture is proved by various techniques. We remark that the only ingredients used until now, broadly fall into three groups.

- (1) The first and foremost ingredient is the prime number theorem (PNT) and the Dirichlet PNT.
- (2) The result of Matomaki-Radziwiłł-Tao on the validity of average Chowla of order two [17]. This was used to establish the conjecture for systems with discrete spectra. It was used by el Abdalaoui-Lemańczyk-de-la-Rue [2], Huang-Wang and Zhang [15], Huang-Wang-Ye [16].
- (3) The Daboussi-Katai-Bourgain-Sarnak-Ziegler criterion, which says that if a sequence (a_n) satisfy for a large prime p and q , the orthogonality of (a_{np}) and (a_{nq}) , then the orthogonality holds between (a_n) and any multiplicative function. In fact this criterion is also based on PNT, that is, the PNT implies Daboussi-Katai-Bourgain-Sarnak-Ziegler criterion. We remark that however the converse implication is false. For example, take the trivial sequence $a_n = 1$. This sequence does not satisfy the criterion. But the Möbius sequence is orthogonal to the constant sequence $a_n = 1$, $n \in \mathbb{N}$ is exactly the PNT. This criterion was used first by Bourgain-Sarnak-Ziegler [7] and then by many other authors. Of course it is not possible to prove the conjecture just using this criterion because one need to verify the conjecture for the constant function in any system and for the eventual periodic sequence, this is where again PNT and Dirichlet PNT come into play.

We mention that our method does not use any of the above techniques, (from Number Theory, we only need the Davenport’s Möbius disjointness theorem which insure that the Möbius is orthogonal to any rotation, and the rest is ‘dynamics-ergodic’ theory). On the other hand we shall strengthen some results obtained by above techniques, proved in [18], [26], [32] as corollaries to our result, (e.g. see Corollaries 3.5 and 3.6).

2 Definitions and tools

We start by recalling the definition of the Möbius function which is intimately linked to the Liouville function, denoted by λ . This later function is defined to be 1 if the number of prime factors (counting with multiplicity) is even and -1 if not. The Möbius function is given by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1; \\ \lambda(n) & \text{if } n \text{ is square-free}; \\ 0 & \text{if not} \end{cases} \quad (2.1)$$

We recall that n is square-free if n has no factor in the subset $\mathcal{P}_2 \stackrel{\text{def}}{=} \{p^2/p \in \mathcal{P}\}$, where as customary, \mathcal{P} denote the set of prime numbers.

A topological dynamical system is a pair (X, T) where X is a compact metric space and T is a homeomorphism. The topological entropy $h_{\text{top}}(T)$ of T is given by

$$h_{\text{top}}(T) = \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow +\infty} \frac{1}{n} \log \text{sep}(n, T, \varepsilon).$$

where for n integer and $\varepsilon > 0$, $\text{sep}(n, T, \varepsilon)$ is the maximal possible cardinality of an (n, T, ε) -separated set in X , this later means that for every two distinct points of it, there exists $0 \leq j < n$ with $d(T^j(x), T^j(y)) > \varepsilon$, where T^j denotes the j -th iterate of T . For more details, we refer to [30, Chap. 7].

We are now able to state Sarnak's Möbius disjointness conjecture. It states that for any compact metric, topological dynamical system (X, T) with topological entropy zero, we have, for any $x \in X$ and any $f \in C(X)$,

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

The popular Chowla conjecture on the correlation of the Möbius function state that, for any $r \geq 0$, $1 \leq a_1 < \dots < a_r$, $i_s \in \{1, 2\}$ not all equal to 2, we have

$$\sum_{n \leq N} \mu^{i_0}(n) \mu^{i_1}(n + a_1) \cdot \dots \cdot \mu^{i_r}(n + a_r) = o(N). \quad (2.2)$$

This conjecture implies a weaker conjecture stated by Chowla in [10, Problem 56. pp. 95-96].

The Chowla conjecture has a dynamical interpretation. Indeed, one can see μ as a point in a compact space $X_3 = \{0, \pm 1\}^{\mathbb{N}}$. As customary, we consider the shift map S on it, and by taking the map $s : (x_n) \in X_3 \mapsto (x_n^2) \in X_2 \stackrel{\text{def}}{=} \{0, 1\}^{\mathbb{N}}$, it follows that the topological dynamical system (X_2, S) is a factor of (X_3, S) . We will denote by pr_1 the projection map on the first coordinate, that is, $\text{pr}_1(x) = x_1$, where $x \in X_i$, $i = 2, 3$.

For a careful study of Chowla conjecture, one also need to introduce the notion of admissibility. This later notion is due to Mirsky [22] (see also [28]).

Definition 2.1

(I) The subset $A \subset \mathbb{N}$ is admissible if the cardinality $t(p, A)$ of classes modulo p^2 in A given by

$$t(p, A) \stackrel{\text{def}}{=} |\{z \in \mathbb{Z}/p^2\mathbb{Z} : \exists n \in A, n \equiv z \pmod{p^2}\}|,$$

satisfies

$$\forall p \in \mathcal{P}, t(p, A) < p^2. \quad (2.3)$$

In other words, for every prime p the image of A under reduction mod p^2 is a proper subset of $\mathbb{Z}/p^2\mathbb{Z}$.

(II) An infinite sequence $x = (x_n)_{n \in \mathbb{N}} \in X_3$ is said to be admissible if its support $\{n \in \mathbb{N} : x_n \neq 0\}$ is admissible. In the same way, a finite block $x_1 \dots x_N \in \{0, \pm 1\}^N$ is admissible if $\{n \in \{1, \dots, N\} : x_n \neq 0\}$ is admissible. In the same manner, we define the admissible sets in X_2 .

For each $i = 2, 3$, we denote by $X_{\mathcal{A}_i}$ the set of all admissible sequences in X_i . Since a set is admissible if and only if each of its finite subsets is admissible, and a translation of an admissible set is admissible, $X_{\mathcal{A}_i}$ is a closed and shift-invariant subset of X_i , i.e. a subshift. We further have that μ^2 is an admissible sequence, and $X_{\mathcal{A}_3} = s^{-1}(X_{\mathcal{A}_2})$, where s is the square mapping. We denote by $\mathcal{B}(\mathcal{A}_i)$ the Borel σ -algebra generated by \mathcal{A}_i , $i = 2, 3$.

2.1 Some tools from spectral theory of dynamical systems

Let (X, \mathcal{A}, μ, T) be a measurable dynamical system, that is, (X, \mathcal{A}, μ) is a probability space and T is an invertible, bi-measurable map which preserve μ , i.e. $\mu(T^{-1}(A)) = \mu(A)$, for every $A \in \mathcal{A}$. The dynamical system is ergodic if the T -invariant set is trivial: $\mu(T^{-1}(A) \Delta A) = 0 \implies \mu(A) \in \{0, 1\}$. Transformation T induces an operator U_T in $L^p(X)$ via $f \mapsto U_T(f) = f \circ T$ called Koopman operator. For $p = 2$ this operator is unitary and its spectral resolution induces a spectral decomposition of $L^2(X)$ [24]:

$$L^2(X) = \bigoplus_{i=0}^{+\infty} C(f_i) \quad \text{and} \quad \sigma_{f_1} \gg \sigma_{f_2} \gg \dots$$

where

- $\{f_i\}_{i=1}^{+\infty}$ is a family of functions in $L^2(X)$;
- $C(f) \stackrel{\text{def}}{=} \overline{\text{span}}\{U_T^n(f) : n \in \mathbb{Z}\}$ is the cyclic space generated by $f \in L^2(X)$;
- σ_f is the *spectral measure* on the circle generated by f via the Bochner-Herglotz relation

$$\widehat{\sigma}_f(n) = \langle U_T^n f, f \rangle = \int_X f \circ T^n(x) \bar{f}(x) d\mu(x); \quad (2.4)$$

- for any two measures on the circle α and β , $\alpha \gg \beta$ means β is absolutely continuous with respect to α : for any Borel set A , $\alpha(A) = 0 \implies \beta(A) = 0$. The two measures α and β are equivalent if and only if $\alpha \gg \beta$ and $\beta \gg \alpha$. We will denote measure equivalence by $\alpha \sim \beta$.

The spectral theorem ensures this spectral decomposition is unique up to isomorphisms. The *maximal spectral type* of T is the equivalence class of the Borel measure σ_{f_1} . The multiplicity function $\mathcal{M}_T : \mathbb{T} \longrightarrow \{1, 2, \dots, \} \cup \{+\infty\}$ is defined σ_{f_1} a.e. and

$$\mathcal{M}_T(z) = \sum_{j=1}^{+\infty} \mathbb{1}_{Y_j}(z), \quad \text{where, } Y_1 = \mathbb{T} \text{ and } Y_j = \text{supp} \frac{d\sigma_{f_j}}{d\sigma_{f_1}} \quad \forall j \geq 2.$$

An integer $n \in \{1, 2, \dots, \} \cup \{+\infty\}$ is called an essential value of M_T if $\sigma_{f_1}\{z \in \mathbb{T} : M_T(z) = n\} > 0$. The multiplicity is uniform or homogeneous if there is only one essential value of M_T . The essential supremum of M_T is called the maximal spectral multiplicity of T . The map T

- has simple spectrum if $L^2(X)$ is a single cyclic space;
- has discrete spectrum if $L^2(X)$ has an orthonormal basis consisting of eigenfunctions of U_T , (in this case σ_{f_1} is a discrete measure);

- has Lebesgue spectrum if σ_{f_1} is equivalent to the Lebesgue measure. It has absolutely continuous (or singular) spectrum if σ_{f_1} is absolutely continuous (or singular) with respect to the Lebesgue measure.

Definition 2.2 *The reduced spectral type of the dynamical system is its spectral type on the $L^2_0(X)$ - the space of square integrable functions with zero mean. Two dynamical systems are called spectrally disjoint if their reduced spectral types are mutually singular.*

We will need also the following result, (Lemma 5, Chapter 3 in [24]).

Lemma 2.3 (Rokhlin [27],[24]) *Let (X, \mathcal{A}, μ, T) be a measure theoretic dynamical system. If \mathcal{C} is proper non-atomic sub- σ -algebra then orthocomplement of $L^2(X, \mathcal{C}, \mu)$ is infinite-dimensional.*

As a consequence, of above lemma and the proof of a theorem due to Rokhlin and Sinai, (see Theorem 15, Chapter 6 of [24]), we have the following.

Lemma 2.4 *Let (X, \mathcal{A}, μ, T) be a measure theoretic dynamical system on a Lebesgue space X . Suppose that the orthogonal complement $L^2(X, \Pi(T), \mu)^\perp$ to the subspace $L^2(X, \Pi(T), \mu)$ in $L^2(X, \mathcal{A}, \mu)$ is infinite dimensional, where $\Pi(T)$ is the Pinsker σ -algebra corresponding to T . Then U_T has countable Lebesgue spectrum on $L^2(X, \Pi(T), \mu)^\perp$.*

This lemma is the motivation behind one of the ingredient in our proof of Theorem 3.1.

2.2 Spectral measure of a sequence

The notion of spectral measure for sequences was introduced by Wiener in 1933, (see [31]). Therein, he define a space \mathcal{W} of complex bounded sequences $a = (a_n)_{n \in \mathbb{N}}$ such that

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=0}^{N-1} a_{n+k} \bar{a}_n = F(k) \quad (2.5)$$

exists for each integer $k \in \mathbb{N}$. The sequence $F(k)$ can be extended to negative integers by setting

$$F(-k) = \overline{F(k)}.$$

It can be further seen that F is positive definite on \mathbb{Z} and therefore by the Herglotz-Bochner theorem, there exists a unique positive finite measure σ_a on the circle \mathbb{T} such that the Fourier coefficients of σ_a are given by the sequence F , that is,

$$\widehat{\sigma}_a(k) \stackrel{\text{def}}{=} \int_{\mathbb{T}} e^{-ikx} d\sigma_a(x) = F(k).$$

The measure σ_a is called the *spectral measure of the sequence a* .

Remark 2.5

- (I) *We mention that according to Chowla conjecture, the spectral measure of the Möbius function is the Lebesgue measure.*

(II) Given a topological dynamical system (X, T) , $f \in C(X)$ and $x \in X$, there is a natural connection between the spectral measure of the dynamical sequence $n \mapsto f(T^n x)$ and the spectral type of the dynamical system. Indeed, for a uniquely ergodic system (X, T, μ) , this sequence belongs to the Wiener space \mathcal{W} and its spectral measure is exactly the spectral measure of the function f

$$\widehat{\sigma}_f(k) = \langle U_T^k(f), f \rangle = \int f \circ T^k(x) \cdot \overline{f(x)} d\mu(x).$$

In the general situation, the following notion of a ‘(set of) potential spectral measures’ of x captures the appropriate concept.

Definition 2.6 Given a compact metric dynamical system (X, T) and $x \in X$, let $\mathcal{I}_T(x)$ be the weak-star closer of the set $\{\frac{1}{N} \sum_{n=1}^N \delta_{T^n x} \mid N \in \mathbb{N}\}$. Note that any $\nu_x \in \mathcal{I}_T(x)$ is T -invariant. For each such ν_x , there is a subsequence (N_ℓ) such that for any $k \in \mathbb{Z}$ and for any $f \in C(X)$ we have,

$$\lim_{\ell \rightarrow \infty} \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} \delta_{T^n x}(f \circ T^k \cdot \overline{f}) = \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} f(T^{n+k} x) \overline{f(x)} = \int f \circ T^k(y) \overline{f(y)} d\nu_x = \widehat{\sigma_{f, \nu_x}}(k), \quad (2.6)$$

where σ_{f, ν_x} is the spectral measure associated to the vector f in the Hilbert space $L^2(X, \nu_x)$. Such a spectral measure σ_{f, ν_x} will be called a potential spectral measure of x corresponding to the function f .

2.3 Correlations and Affinity

The notion of ‘affinity’ was introduced and studied in a series of papers by Matusita [19],[20],[21] and it is also called Bahattacharyya coefficient [6]. It has been widely used in statistics literature to quantify the similarity between two probability distributions. The affinity between two finite measures is defined by the integral of the corresponding geometric mean. Let $\mathcal{M}_1(\mathbb{T})$ be a set of probability measures on the circle \mathbb{T} and $\eta, \nu \in \mathcal{M}_1(\mathbb{T})$ two measures in this space. There exists a probability measure λ such that η and ν are absolutely continuous with respect to λ , (take for example $\lambda = \frac{\eta + \nu}{2}$). Then the *affinity* between η and ν is defined by

$$G(\eta, \nu) = \int \sqrt{\frac{d\eta}{d\lambda} \cdot \frac{d\nu}{d\lambda}} d\lambda. \quad (2.7)$$

This definition does not depend on λ . Affinity is related to the Hellinger distance which is defined as

$$H(\eta, \nu) = \sqrt{2(1 - G(\eta, \nu))}.$$

As a consequence of Cauchy-Schwarz inequality we have the following,

$$0 \leq G(\eta, \nu) \leq 1.$$

Remark 2.7

- (I) The definition of affinity can be extended to any pair of positive finite measures by normalizing them.
- (II) It is an easy exercise to see that $G(\eta, \nu) = 0$ if and only if η and ν are mutually singular (denoted by $\eta \perp \nu$): this means that there exists a pair of disjoint Borel sets A and B such that η is concentrated on A and ν is concentrated on B (a measure ρ is concentrated on a Borel set E if $\rho(F) = 0$ if and only if $F \cap E = \emptyset$). Similarly, $G(\eta, \nu) = 1$ holds if and only if η and ν are equivalent. Affinity can be used to compare sequences of measures via the following theorem.

Theorem 2.8 (Coquet-Kamae-Mandès-France [9]) *Let (P_n) and (Q_n) be two sequences of probability measures on the circle, weakly converging to the probability measures P and Q respectively. Then*

$$\limsup_{n \rightarrow +\infty} G(P_n, Q_n) \leq G(P, Q). \quad (2.8)$$

Remark 2.9 *As in the case of the affinity, this result can be generalized to any sequence of positive non trivial finite measures P_n, Q_n converging weakly to two positive non trivial finite measures P and Q .*

We want to use the affinity and Theorem 2.8 above to estimate the orthogonality properties of pairs of sequences in the Wiener space \mathcal{W} (defined in subsection 2.2). To do that, we will need to replace the sequence of Fourier coefficients $\frac{1}{n} \sum_{j=0}^{n-1} g_{j+k} \overline{g_j}$ by a sequence of finite positive measures on the circle. For any $g \in \mathcal{W}$, we introduce the sequence of measures,

$$d\sigma_{g,n}(x) = \rho_{g,n}(x) \frac{dx}{2\pi}, \quad \text{where } \rho_{g,n}(x) = \left| \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} g_j e^{ijx} \right|^2. \quad (2.9)$$

With this definition $\sigma_{g,n}$ defines a finite positive measure on the circle. Moreover we have the relation

$$\frac{1}{n} \sum_{j=0}^{n-1} g_{j+k} \overline{g_j} = \int_0^{2\pi} e^{-ikx} d\sigma_{g,n}(x) + \Delta_{n,k} = \widehat{\sigma}_{g,n}(k) + \Delta_{n,k},$$

where

$$|\Delta_{n,k}| = \left| \frac{1}{n} \sum_{j=n-k}^{n-1} g_{j+k} \overline{g_j} \right| \leq \frac{k}{n} \sup_j |g_j|^2 \xrightarrow{n \rightarrow +\infty} 0.$$

Taking the limit we have

$$\widehat{\sigma}_g(k) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} g_{j+k} \overline{g_j} = \lim_{n \rightarrow \infty} \widehat{\sigma}_{g,n}(k),$$

so the sequence of measures $(\sigma_{g,n})_{n \in \mathbb{N}}$ converges weakly to σ_g .

It may be possible that the bounded sequence $(g_n)_{n \in \mathbb{N}}$ does not belong to the Wiener space \mathcal{W} (see eq. (2.5)) but we can always extract a subsequence (n_r) in (2.5) such that

$$\lim_{r \rightarrow \infty} \frac{1}{n_r} \sum_{j=0}^{n_r-1} g_{j+k} \overline{g_j}$$

exists for each $k \in \mathbb{N}$. In fact, consider the sequence of finite positive measures $(\sigma_{g,n})_{n \in \mathbb{N}}$ on the torus defined in (2.9). These measures are all finite and $\sigma_{g,n}(\mathbb{T}) \leq \|g\|_\infty^2 = \sup_j |g_j|^2$, for all n . Therefore they all belong the ball $B(0, \|g\|_\infty^2)$ centered at 0 with radius $\|g\|_\infty^2$ in the set of measures on the circle. This subset is compact so there exists a subsequence (n_r) such that the sequence of probability measures $(\sigma_{g,n_r})_{r \in \mathbb{N}}$ converges weakly to some probability measure $\sigma_{g,(n_r)}$. The measure $\sigma_{g,(n_r)}$ is called the spectral measure of the sequence g along the subsequence (n_r) .

We will also need the following result whose proof follows from Theorem 2.8.

Corollary 2.10 [5] *Let $g = (g_n)_{n \in \mathbb{N}}, h = (h_n)_{n \in \mathbb{N}} \in \mathcal{W}$ two non trivial sequences i.e. $\widehat{\sigma}_g(0) > 0$ and $\widehat{\sigma}_h(0) > 0$. Then*

$$\limsup_{n \rightarrow \infty} \left| \frac{1}{n} \sum_{j=1}^n g_j \bar{h}_j \right| \leq \sup \{G(\sigma_{g,(n_r)}, \sigma_{h,(m_r)})\}. \quad (2.10)$$

Where the supremum on the right-hand side is taken over all supsequence $(n_r), (m_r)$ for which the spectral measures exists.

3 Main result and its corollaries

In this section, we begin by stating our main result-Theorem 3.1 and prove some of its consequences. In the later half of this section we start building necessary tools, (Theorem 3.8 and Proposition 3.22). The proof of Theorem 3.1 will be completed towards the end of Section 4.

Theorem 3.1 *Let (X, T) be a dynamical system such that for any invariant measure the spectrum is singular. Then, the Möbius disjointness holds.*

Now we state a few corollaries, the first one follows rather immediately.

Corollary 3.2 *The Möbius disjointness conjecture holds for any rigid transformation.*

We recall that the transformation T is rigid if there is a sequence of integers (n_k) such that (T^{n_k}) converge strongly to the identity map. As a consequence, we obtain an improvement of the Theorem 5.1 in [32] which extended substantially the main result in [25]. In [32], the author proved that the Möbius orthogonality holds for the rigid map under an extra-condition.

In fact, our proofs allows us to obtain as a corollary the main ingredient needed in his proof. This follows from our Corollary 3.6.

The second corollary answer partially the question asked in [1] and it has a direct application to number theory. We state it as follows.

Corollary 3.3 *All the potential spectral measures of Möbius and Liouville function are absolutely continuous with respect to the Lebesgue measure.*

The proof of Corollary 3.3 will follows from our dissection of Möbius flow (see Section 4). As a consequence, we establish that Liouville flow is a factor of Möbius flow.

Remark 3.4

1. *As a corollary to our Theorem (3.1), Sarnak Möbius disjointness is fulfilled for a systems for which every invariant measure has discrete spectra. An earlier approach to this result was developed by applying Matomaki-Radziwiłł-Tao's Theorem on the average Chowla conjecture, (see [17]). More precisely, the crucial argument in that approach was based on the fact that the average of the correlations of order two of Möbius and Liouville functions converge to 0 with logarithmic speed (see Proposition 5.1 in [17]). Many authors used this approach to establish the Möbius disjointness conjecture for such systems, [2], [15] [16]. Our method bypasses all of this and obtains the validity of Sarnak conjecture for a much wider class of systems.*

2. We established that ‘Veech Systems’ have the property that all invariant measures have discrete spectrum, (Theorem 4.20 in [4]), and in particular the translation flow on the orbit closure of a ‘Veech function’ on \mathbb{Z} has the same property, (see [4] for the definitions). Hence the translation flow on the orbit closure of the Veech function satisfy the Möbius randomness law. This along with results in [4] allows us to obtain a dynamical proof of Motohashi-Ramachandra type theorem and Matomaki-Radzwil’s result [18] on the short interval for the case of Liouville and Möbius function. This later result can be stated as follows : For any $\epsilon > 0$, for any $H \leq X$ large, we have

$$\int_X^{2X} \left| \sum_{x < k < x+H} \mu(k) \right| dx = o(HX).$$

Corollary 3.3 allows us also to obtain as a consequence the main theorem in [17] and Corollary 1.2 from [18]. Precisely, we have

Corollary 3.5

(I) [Matomaki-Radzwil-Tao Theorem [17]] For any $\epsilon, \delta \in (0, 1)$, There is a large but fixed $H = H(\delta)$ such that, for all large enough X , the cardinality of the set of all h ’s with $|h| \leq H$ satisfying

$$\left| \sum_{1 \leq j \leq X} \lambda(j)\lambda(j+h) \right| \leq \delta X. \tag{3.1}$$

is greater than $(1 - \epsilon)H$.

(II) [Matomaki-Radzwil Corollary [18]] We further have, for any $h \geq 1$, there exists $\epsilon(h) > 0$ such that

$$\left| \frac{1}{X} \sum_{1 \leq j \leq X} \lambda(j)\lambda(j+h) \right| \leq 1 - \epsilon(h). \tag{3.2}$$

for all large enough X .

Proof. (I) Let (X_k) be a subsequence such that $X_k \rightarrow +\infty$. Then,

$$\left| \sum_{1 \leq j \leq X_k} \lambda(j)\lambda(j+h) \right| \xrightarrow{k \rightarrow +\infty} \int z^{-h} f(z) dz,$$

for some function $f \in L^1(dz)$. But we also have

$$\frac{1}{H} \sum_{h=1}^H |\widehat{f}(h)| \xrightarrow{H \rightarrow \infty} 0.$$

This proves the first part.

(II) For the second part, assume by contradiction, that there is $h \geq 1$ such that for any $\epsilon > 0$, we have

$$\limsup_{X \rightarrow +\infty} \left| \frac{1}{X} \sum_{1 \leq j \leq X} \lambda(j)\lambda(j+h) \right| \geq 1 - \epsilon.$$

Then, by taking a subsequence which may depend on h , we get

$$\lim_{k \rightarrow +\infty} \left| \frac{1}{X_k} \sum_{1 \leq j \leq X_k} \lambda(j)\lambda(j+h) \right| = |\widehat{f}(h)| = 1,$$

which is impossible by Cauchy-Schwarz inequality. ■

We further have the following corollary which generalizes the results in [32, Theorem 1.8] and [26, Theorem 6.1].

Corollary 3.6 *For any $k \in \mathbb{N}$ and for a large h , $\limsup_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=1}^N \left| \sum_{l=1}^h \mu(n+kl) \right|^2 = o(h^2)$.*

Proof. Let σ_μ be a potential spectral measure. Then

$$\frac{1}{N_j} \sum_{n=1}^{N_j} \left| \sum_{l=1}^h \mu(n+kl) \right|^2 \xrightarrow{j \rightarrow +\infty} \int_{\mathbb{T}} |D_h(k\theta)|^2 d\sigma_\mu,$$

where $D_h(\theta) = \sum_{l=1}^h e^{il\theta}$ is the classical Dirichlet kernel. We thus get, by Corollary 3.3 ,

$$\int_{\mathbb{T}} |D_h(k\theta)|^2 d\sigma_\mu = \int |D_h(k\theta)|^2 f(\theta) d\theta,$$

for some function $f \in L^1(dz)$. Furthermore, the functions $\phi_h(\theta) = \frac{1}{h} D_h(\theta)$ tend to zero uniformly outside any neighborhood of 0. The rest of the proof is left to the reader. ■

Remark 3.7 *Let us point out that our Corollary 3.3 assert much more than (3.1) and (3.2). Indeed, it states that any potential spectral measures of Möbius function σ_μ or Liouville function σ_λ is a Rajchman measure, that is, $|\widehat{\sigma}_\mu(n)|, |\widehat{\sigma}_\lambda(n)| \xrightarrow{|n| \rightarrow +\infty} 0$. However, (3.1) assert only that σ_λ is a continuous measure and (3.2) that σ_λ is in the orthogonal to Dirichlet measures.*

More precisely, let $\mathcal{M}(\mathbb{T})$ be the algebra of the regular Borel complex measures on the torus \mathbb{T} , equipped with the convolution product of measures, defined by $\mu * \nu$. This is the pushforward measure of $\mu \otimes \nu$ under the map $a : (x, y) \in T \times T \mapsto x + y \in \mathbb{T}$. We shall call a subset $L \subset \mathcal{M}(\mathbb{T})$ a L -subspace (resp. L -ideal or L -sub-algebra) of $\mathcal{M}(\mathbb{T})$ when L is a closed subspace (resp. ideal or sub-algebra) of $\mathcal{M}(\mathbb{T})$ that is invariant under absolute continuity of measures. This means

$$\text{if } \mu \in L \text{ and } \nu \ll \mu \text{ then } \nu \in L.$$

It is well know that the sets $\mathcal{M}_c(T) \stackrel{\text{def}}{=} \left\{ \mu \in \mathcal{M}(\mathbb{T}) \mid \mu \text{ is a continuous measure} \right\}$ and $\mathcal{M}_0(T) \stackrel{\text{def}}{=} \left\{ \mu \in \mathcal{M}(\mathbb{T}) \mid \mu \text{ is a Rajchman measure} \right\}$ are L -ideals and

$$\mu \in \mathcal{M}_c(T) \iff \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=0}^N |\widehat{\sigma}_\mu(n)|^2 \xrightarrow{N \rightarrow +\infty} 0.$$

Given an L -ideal, we define its orthogonal as follows

$$L^\perp = \left\{ \mu \in \mathcal{M}(\mathbb{T}) : |\mu| \perp |\nu|, \forall \nu \in L \right\}.$$

Furthermore the following decomposition is well know.

$$\mathcal{M}(T) = L \oplus L^\perp.$$

The probability measure μ on \mathbb{T} is said to be a Dirichlet measure if

$$\limsup_{\gamma \rightarrow +\infty} |\widehat{\mu}(\gamma)| = 1.$$

It is well known that the set $\mathcal{D}(\mathbb{T}) \stackrel{\text{def}}{=} \left\{ \mu \mid \mu \text{ is a Dirichlet measure} \right\}$ is an L -ideal. Moreover,

$$\mu \in \mathcal{D}(\mathbb{T})^\perp \iff \limsup_{k \rightarrow +\infty} |\widehat{\mu}(k)| < 1.$$

and

$$M_0(\mathbb{T}) \subset \mathcal{D}(\mathbb{T})^\perp \subset \mathcal{M}_c(\mathbb{T}).$$

For more details and proofs, we refer to [14, Chap. II].

We move now to the proof of our main results. We start by proving the following special case of the strong Sarnak's Möbius conjecture [28].

Theorem 3.8 *For any continuous function $f \in C(X_{A_2})$, for any $\theta \in \mathbb{R}$, we have*

$$\frac{1}{N} \sum_{n=1}^N \mu(n) f(\mu^2(n)) e^{in\theta} \xrightarrow{N \rightarrow +\infty} 0.$$

The following corollary follows immediately.

Corollary 3.9 *Let $k \geq 2$ and a_1, \dots, a_k be distinct non-negative integers and $\theta \in \mathbb{R}$. Then*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq x} \mu(n) \mu^2(n + a_1) \cdots \mu^2(n + a_k) e^{in\theta} = 0.$$

Proof. This follows by taking $f(x) = \text{pr}_{a_1}(x) \cdots \text{pr}_{a_k}(x)$, for $x \in X_{A_2}$. ■

Remark 3.10 *The following is a recent result due to R. Murty & A. Vatwani [23], (which can be viewed as a generalization of Davenport's estimate), was a motivation for our proof. We mention that our 'non-quantitative' version of their result can be obtained by only 'dynamical systems' arguments.*

Theorem 3.11 *Let $k \geq 2$ and a_1, \dots, a_k be distinct non-negative integers. Then we have for any $A > 0$,*

$$\left| \sum_{n \leq x} \mu(n) \mu^2(n + a_1) \cdots \mu^2(n + a_k) \right| \ll_{a_1, a_2, \dots, a_k, A} \frac{x}{\log(x)^A}.$$

3.1 Proof of Theorem 3.8

For the proof we need to recall the notion of a Besicovitch almost periodic function.

Definition 3.12

(1) A map $\phi : \mathbb{N} \rightarrow \mathbb{C}$ is Besicovitch almost periodic if given $\epsilon > 0$, there exists a trigonometric polynomial $P \equiv P_\epsilon$ given by $P(n) = \sum_{k=1}^M c_k e^{i\alpha_k n}$, $n \in \mathbb{N}$, where $\alpha_k \in \mathbb{R}$, for $1 \leq k \leq M$, such that

$$\|\psi - P\|_{B_1} \stackrel{\text{def}}{=} \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N |\psi(t) - P(t)| < \epsilon.$$

(2) Let (X, T) be a compact metric topological dynamical system. A point $x_0 \in X$ is a Besicovitch almost periodic point if the sequence $n \mapsto f(T^n(x_0))$ is a Besicovitch for each $f \in C(X)$.

Remark 3.13 Let $\phi : \mathbb{N} \rightarrow \mathbb{C}$ be a Besicovitch almost periodic function. Then for any $\theta \in \mathbb{R}$ the map ψ_θ is also a Besicovitch almost periodic function, where $\phi_\theta(n) = \psi(n)e^{in\theta}$.

To see this, given $\epsilon > 0$, let $P(n) = \sum_{k=1}^M c_k e^{i\alpha_k n}$ be a trigonometric polynomial such that $\|\psi - P\|_{B_1} < \epsilon$. Let $Q(n) = P(n)e^{in\theta} = \sum_{k=1}^M c_k e^{i(\alpha_k + \theta)n}$. Note that Q is also a trigonometric polynomial and

$$\begin{aligned} \|\psi_\theta - Q\|_{B_1} &= \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N |\psi_\theta(t) - Q(t)| \\ &= \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N |\psi(t)e^{it\theta} - P(t)e^{it\theta}| \\ &= \|\psi - P\|_{B_1} < \epsilon. \end{aligned}$$

Now the proof of Theorem 3.8 will follow from the following observations. The first is the following result of Mirsky, [22].

Proposition 3.14 The point μ^2 is a generic point for the ‘square-free’ dynamical system (X_{A_2}, S, ν_M) , where ν_M is the Mirsky measure.

In fact more is true, the following observation was made by P. Sarnak and a proof appears in [8].

Proposition 3.15 (Sarnak-Cellarosi-Sinai’s theorem) The dynamical system (X_{A_2}, S, ν_M) has discrete spectrum.

The next result is a sufficiency condition for a point in a dynamical system to be a Besicovitch point, (see [4] for a proof).

Lemma 3.16 Let (X, T, μ) be a compact metric dynamical ergodic system with discrete spectrum. Let $x_0 \in X$ be a μ -generic point. Then x_0 is a Besicovitch point.

As a consequence of the previous proposition and the lemma, we arrive at a generalization of an observation of Rauzy, (Rauzy’s result asserts that the sequence $n \rightarrow \mu^2(n)$ is Besicovitch almost periodic).

Proposition 3.17 Consider the dynamical system (X_{A_2}, S, ν_M) . Then, for any continuous map $f : X \rightarrow \mathbb{C}$, the map $n \mapsto f(S^n \mu^2)$ is Besicovitch almost periodic.

The next observation is a consequence of Davenport’s Möbius disjointness theorem.

Lemma 3.18 Let $\{b_n\}$ be a Besicovitch almost periodic sequence. Then

$$\frac{1}{N} \sum_{k=1}^N \mu(k) b_k \xrightarrow{N \rightarrow +\infty} 0.$$

Now, the proof of Theorem 3.8 follows by putting all of these observations together.

Proof of Theorem 3.8

Proof. Let $f \in C(X_{\mathcal{A}_2})$. Then the map $n \mapsto f(S^n(\boldsymbol{\mu}^2))$ is a Besicovitch almost periodic and so is the map $n \mapsto f(S^n(\boldsymbol{\mu}^2))e^{in\theta}$. Thus, by Lemma 3.18,

$$\frac{1}{N} \sum_{k=1}^N \boldsymbol{\mu}(k) f(S^k(\boldsymbol{\mu}^2)) e^{in\theta} \xrightarrow{N \rightarrow +\infty} 0.$$

■

Remark 3.19 Now we turn to the proof of our main result, Theorem 3.1. For this we need a theorem, (which we shall refer to as the Sarnak-Veech theorem) and the following crucial computation which is essentially due to W. Veech [29]. This Sarnak-Veech's theorem was announced in [28] and Veech gives an unpublished proof in [29], (see [3] where Veech's proof is presented). Sarnak-Veech's theorem states that there exist a unique ergodic admissible measure η_M on $X_{\mathcal{A}_3}$ such that the factor $(X_{\mathcal{A}_2}, S, \nu_M)$ is the 'Pinsker factor' of the dynamical system $(X_{\mathcal{A}_3}, S, \eta_M)$. This means that the Pinsker sigma algebra $\Pi(S)$ of this dynamical system is given by $\Pi(S) = s^{-1}(\mathcal{B}(\mathcal{A}_2))$. We denote by $L^2(X_{\mathcal{A}_3}, \Pi(S), \eta_M)^\perp$ the orthocomplement of $L^2(X_{\mathcal{A}_3}, \Pi(S), \eta_M)$ as a subspace of $L^2(X_{\mathcal{A}_3}, \mathcal{B}(\mathcal{A}_3), \eta_M)$. In the following, we recall the definition of an admissible probability measure.

Definition 3.20 A probability measure η_m on $X_{\mathcal{A}_3}$ is admissible if

- (i) $S\eta_m = \eta_m$, that is, $\eta_m(S^{-1}A) = \eta_m(A)$, for each Borel set $A \subset X_{\mathcal{A}_3}$.
- (ii) $s(\eta_m) = \nu_M$, and

$$\int_{X_{\mathcal{A}_3}} \prod_{a \in A} pr_a(x) \prod_{b \in B} pr_b^2(x) d\eta_m(x) = 0,$$

for any $A \neq \emptyset$. and B finite sets of \mathbb{N} .

Let us state precisely the Sarnak-Veech's theorem.

Theorem 3.21 (Sarnak-Veech's theorem on Möbius flow [28],[29],[3]) There exists a unique admissible measure η_M on $X_{\mathcal{A}_3}$ which is ergodic with the Pinsker algebra

$$\Pi_{\eta_M}(S) = s^{-1}(\mathcal{B}(\mathcal{A}_2)).$$

Moreover, $\mathbb{E}(pr_1 | \Pi_{\eta_M}(S)) = 0$.

We shall refer to this measure η_M as the 'Chowla-Sarnak-Veech measure', (or 'CSV measure'). W. Veech in [29], [3] addresses this measure as 'Chowla measure'. Obviously, under η_M , the spectral measure of pr_1 is the Lebesgue measure. For the proof of our main result, we need also the following crucial computation which is essentially due to W. Veech [29].

Proposition 3.22 For any invariant measure $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$, we have $pr_1 \in L^2(X_{\mathcal{A}_3}, \Pi(S), \eta)^\perp$.

Proof. Let $f \in C(X_{\mathcal{A}_2})$ and put $F = f \circ s$. Let $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$. Then,

$$\int \text{pr}_1(x) F(x) d\eta(x) = \lim_{k \rightarrow +\infty} \frac{1}{N_k} \sum_{n=1}^{N_k} \text{pr}_1(S^n \boldsymbol{\mu}) F(S^n \boldsymbol{\mu}) \quad (3.3)$$

$$= \lim_{k \rightarrow +\infty} \frac{1}{N_k} \sum_{n=1}^{N_k} \boldsymbol{\mu}(n) f(S^n \boldsymbol{\mu}^2) \quad (3.4)$$

$$= 0. \quad (3.5)$$

But the space $\mathcal{S} = \{f \circ s, f \in C(X_{\mathcal{A}_2})\}$ is dense in $L^2(X_{\mathcal{A}_3}, s^{-1}(\mathcal{B}(\mathcal{A}_3)), \eta)$. Therefore, $\text{pr}_1 \in L^2(X_{\mathcal{A}_3}, \Pi(S), \eta)^\perp$ and the proof is complete. \blacksquare

At this point, we state a conjecture of W. Veech, (which he announced in his unpublished notes [29]).

Conjecture 3.23 [W. Veech] *For any $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$, we have*

$$\text{pr}_1 \in L^2(X_{\mathcal{A}_3}, \Pi_\eta(S), \eta)^\perp,$$

In the same notes, W. Veech proved that his conjecture implies Sarnak Möbius disjointness. In the following, we state this precisely and give his proof in the next section.

Theorem 3.24 (Veech's Theorem [29]) *Suppose that for any $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$, we have*

$$\text{pr}_1 \in L^2(X_{\mathcal{A}_3}, \Pi_\eta(S), \eta)^\perp,$$

then Sarnak Möbius disjointness holds. Here $\Pi_\eta(S)$ denotes the Pinsker sigma algebra of the dynamical system $(X_{\mathcal{A}_3}, S, \eta)$.

4 Proof of Veech's theorem (Theorem 3.24).

We begin by denoting $X_{\mathcal{A}_i^*}$, $i = 2, 3$ the subset of sequence x for which the support- $\text{supp}(x)$ is infinite and let $\Omega_j = \{\pm 1\}^{\mathbb{Z}_j}$, with $\mathbb{Z}_0 = \mathbb{N} \cup \{0\}$ and $\mathbb{Z}_1 = \mathbb{Z}$. We introduce also the following skew product.

$$\begin{aligned} \psi_0 : X_{\mathcal{A}_2} \times \Omega_0 &\rightarrow X_{\mathcal{A}_2} \times \Omega_0 \\ (x, \omega) &\mapsto (Sx, S^{\text{pr}_1(x)}(\omega)). \end{aligned} \quad (4.1)$$

The natural extension of ψ_0 to $X_{\mathcal{A}_2} \times \Omega_0$ is denoted by ψ_1 . Let $\alpha : \mathbb{Z}_1 \rightarrow \mathbb{Z}_0$ be defined by putting $\alpha(\omega) = (\omega_k)_{k=0}^{+\infty}$. Obviously, α is onto and $\alpha \circ \psi_1 = \psi_0 \circ \alpha$. We further define 'co-ordinate wise' a map

$$\begin{aligned} \Phi : X_{\mathcal{A}_2} \times \Omega_0 &\rightarrow X_{\mathcal{A}_3} \\ (x, \omega) &\mapsto \Phi(x, \omega) : \text{pr}_n(\Phi(x, \omega)) = \begin{cases} 0 & \text{if } n \notin \text{supp}(x) = \{n_1 < n_2 < \dots < n_k < \dots\} \\ \text{pr}_k(\omega) & \text{if } n = n_k \in \text{supp}(x) \end{cases} \end{aligned} \quad (4.2)$$

Consequently, we have

$$\Phi(X_{\mathcal{A}_2} \times \Omega_0) = X_{\mathcal{A}_3}, \quad (4.3)$$

and

$$\Phi \circ \psi_0 = S \circ \Phi \quad (4.4)$$

Φ is also onto but not one to one. However, its restriction $\Phi : X_{\mathcal{A}_2^*} \times \Omega_0 \rightarrow X_{\mathcal{A}_3^*}$ is an onto homeomorphism. With a slight abuse of notation, denoting by $\Phi_*^{-1}\eta$ -the ‘push forward’ of measure η under the map Φ^{-1} , we see that the measure theoretic dynamical systems $(X_{\mathcal{A}_2^*} \times \Omega_0, \psi_0, \Phi_*^{-1}\eta)$ and $(X_{\mathcal{A}_3^*}, S, \eta)$ are measure theoretically isomorphic by the map Φ , where the underlying sigma algebras are respective Borel sigma algebras and this holds for any $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$. We also note that, since $\boldsymbol{\mu}^2$ is a generic point for the Mirsky measure, it follows that (i) $s_*\eta = \nu_M$ for any $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$ and (ii) $\nu_M(X_{\mathcal{A}_2} \setminus X_{\mathcal{A}_2^*}) = 0$. Next, notice that the Pinsker sigma algebra of the first of the above isomorphic dynamical systems can be written down by the following two equal expressions mod null sets,

$$\bigcap_{n=0}^{+\infty} \psi_0^{-n}(\mathcal{B}(\mathcal{A}_2^* \times \Omega_0)) = \Phi^{-1}\left(\bigcap_{n=0}^{+\infty} S^{-n}\mathcal{B}(\mathcal{A}_3^*)\right).$$

At this point, let us observe that the projection pr_1 on $X_{\mathcal{A}_3}$ may be represented in $X_{\mathcal{A}_2} \setminus \{0\} \times \Omega_0$ by a function Pr_1 given by

$$\text{Pr}_1(x, \omega) = \text{pr}_1(x)\text{pr}_1(\omega).$$

Moreover, we have

$$\text{Pr}_1(\Phi(x, \omega)) = \text{pr}_1(x, \omega).$$

Indeed, by definition, we have

$$\text{Pr}_1(\Phi(x, \omega)) = \begin{cases} 0 & \text{if } \text{pr}_1(x) = 0 \\ \text{pr}_1(\omega) & \text{if } \text{pr}_1(x) = 1 \end{cases}$$

and

$$\text{Pr}_1(x, \omega) = \begin{cases} 0 & \text{if } \text{pr}_1(x) = 0 \\ \text{pr}_1(\omega) & \text{if } \text{pr}_1(x) = 1. \end{cases}$$

We thus get

$$\forall (x, \omega) \in X_{\mathcal{A}_2} \times \Omega_0, \quad \text{Pr}_1 \circ \Phi(x, \omega) = \text{pr}_1(x, \omega). \quad (4.5)$$

Let $(\boldsymbol{\mu}^2, \omega_0)$ be the unique point such that

$$\Phi((\boldsymbol{\mu}^2, \omega_0)) = \boldsymbol{\mu}. \quad (4.6)$$

Therefore

$$\mathcal{I}_{\psi_0}(\boldsymbol{\mu}^2, \omega_0) = \Phi^{-1}(\mathcal{I}_S(\boldsymbol{\mu})).$$

At this point, notice that for the natural extension ψ_1 we have, by Cellarosi-Sinai Theorem [8], the system $(X_{\mathcal{A}}, S, \nu_M)$ is metrically isomorphic to the rotation on the compact group $\prod_{p \in \mathcal{P}} \mathbb{Z}/p^2\mathbb{Z}$. Whence, the entropy with respect to the Mirsky measure ν_M of S is 0 and hence the map S is ν_M -a.e invertible. Let $\mathcal{A}_{2,0}^* \subset \mathcal{A}_2^*$ be a Borel set such that $\nu_M(\mathcal{A}_{2,0}^*) = 1$. Then, ψ_1 is an homeomorphism of $\mathcal{A}_{2,0}^* \times \Omega_1$. We choose also $\omega_1 \in \Omega_1$ such that $\text{pr}_n(\omega_1) = \text{pr}_n(\omega_0)$, for all $n \in \mathbb{N} \cup \{0\}$. In the dynamical system $(X_{\mathcal{A}_2^*} \times \Omega_0, \psi_1)$, we keep the notation $\mathcal{I}_{\psi_1}(\boldsymbol{\mu}, \omega_1)$ for the set of invariant measure that arise from the forward orbit under ψ_1 of $(\boldsymbol{\mu}^2, \omega_1)$. For each $\lambda \in \mathcal{I}_{\psi_1}(\boldsymbol{\mu}, \omega_1)$, again, by Mirsky-Sarnak theorem, $\text{pr}_1\lambda = \nu_M$. We further have

Claim 4.1 *The map $\alpha : \lambda \in \mathcal{I}_{\psi_1}(\boldsymbol{\mu}^2, \omega_1) \longrightarrow \alpha(\lambda) \in \mathcal{I}_{\psi_0}(\boldsymbol{\mu}^2, \omega_0)$ is an isomorphism onto.*

Proof. Let $\lambda \in \mathcal{I}_{\psi_1}(\boldsymbol{\mu}, \omega_1)$. Then, there exists a sequence (N_k) such that, for any continuous function f on $\mathcal{A}_{2,0}^* \times \Omega_1$, we have

$$\frac{1}{N_k} \sum_{n=0}^{N_k} F(\psi^n(\boldsymbol{\mu}^2, \omega_1)) \xrightarrow[k \rightarrow +\infty]{} \int_{\mathcal{A}_{2,0}^* \times \Omega_0} f(x, \omega) d\lambda(x, \omega).$$

But, because only the forward orbit is involved, the corresponding limit for $(\boldsymbol{\mu}^2, \omega_1)$ not only exists but does not depend upon the choice of extension of the sequence ω_0 to be a bisequence ω_1 . Therefore, α is both onto and invertible. ■

Now, we start the proof of Theorem 3.24. Let T be a homeomorphism of a compact metric space Y and assume that $y \in Y$ is completely deterministic, that is, for any $\kappa \in \mathcal{I}_T(y)$, the measure entropy $h_\kappa(T) = 0$. Observe that T induces an affine homeomorphism of $\mathcal{P}(Y)$ the space of probability measures on Y equipped with the weak-star topology and $\mathcal{B}(\mathcal{P}(Y))$ the corresponding Borel field. Let us recall also the definition of quasi-factor needed in the proof.

Definition 4.2 *The dynamical system $(\mathcal{P}(Y), \mathcal{B}(\mathcal{P}(Y)), \Theta, T)$ is said to be a quasi-factor of $(Y, \mathcal{B}(Y), \kappa, T)$ if $\Theta \in \mathcal{P}(\mathcal{P}(Y))$ is*

(i) *T -invariant, and*

(ii) *for any continuous function $f \in C(Y)$, we have*

$$\int_{\mathcal{P}(Y)} \int_Y f(z) d\nu(z) d\Theta(\nu) = \int_Y f(z) d\kappa(z),$$

that is,

$$\int_{\mathcal{P}(Y)} \nu d\Theta(\nu) = \kappa,$$

we say that κ is the barycenter of Θ .

We need also the following theorem due to Glasner and Weiss [13, Theorem 18.17, p.326],

Lemma 4.3 *The quasi-factor of zero entropy system is a zero entropy system.*

Proof of Theorem 3.24. Let f be a continuous function on Y and write

$$\frac{1}{N} \sum_{n=1}^N \boldsymbol{\mu}(n) f(T^n y) = \frac{1}{N} \sum_{n=1}^N \text{pr}_1(S^n \boldsymbol{\mu}) f(T^n y). \quad (4.7)$$

Taking into account (4.5) combined with (4.6), we can rewrite (4.7) as follows

$$\frac{1}{N} \sum_{n=1}^N \boldsymbol{\mu}(n) f(T^n y) = \frac{1}{N} \sum_{n=1}^N \text{pr}_1(\Phi \circ \alpha \circ \psi_1^n)(\boldsymbol{\mu}^2, \omega_1) f(T^n y), \quad (4.8)$$

since $\Phi \circ \alpha \circ \psi_1 = S \circ \Phi \circ \alpha$, by the definition of α and (4.4). It follows, by (4.5), that

$$\frac{1}{N} \sum_{n=1}^N \mu(n) f(T^n y) = \frac{1}{N} \sum_{n=1}^N (\text{Pr}_1 \circ \alpha \circ \psi_1^n)(\mu^2, \omega_1) f(T^n y), \quad (4.9)$$

$$= \frac{1}{N} \sum_{n=1}^N \delta_{\psi_1 \times T((\mu^2, \omega_1), y)}((\text{Pr}_1 \circ \alpha) \otimes f) \quad (4.10)$$

Hence, by the compactness of $\mathcal{P}((X_{\mathcal{A}_2} \times \Omega_1) \times Y)$, we can extract a subsequence (N_k) such that

$$\frac{1}{N_k} \sum_{n=1}^{N_k} \delta_{\psi_1 \times T((\mu^2, \omega_1), y)} \xrightarrow[k \rightarrow +\infty]{} \xi,$$

in the weak-star topology. Denote by γ_1 and γ_2 the projections of $(X_{\mathcal{A}_2} \times \Omega_1) \times Y$ on $(X_{\mathcal{A}_2} \times \Omega_1)$ and Y , respectively. Then,

$$\gamma_1 \xi \stackrel{\text{def}}{=} \lambda \in \mathcal{I}_{\psi_1}(\mu^2, \omega_1),$$

and

$$\gamma_2 \xi \stackrel{\text{def}}{=} \kappa \in \mathcal{I}_T(y).$$

Disintegrate ξ over λ , we get, for any $A \in \mathcal{B}(X_{\mathcal{A}_2} \times \Omega_1)$ and $B \in \mathcal{B}(Y)$,

$$\xi(A \times B) = \int_A \kappa_{(x, \omega)}(B) d\lambda(x, \omega),$$

with $\kappa_{(x, \omega)} \in \mathcal{P}(Y)$. Moreover, $T\kappa_{(x, \omega)} = \kappa_{\psi_1(x, \omega)}$, since ψ_1 is λ a.e. invertible. We thus define λ a.e. a mapping

$$\begin{aligned} \sigma : (X_{\mathcal{A}_2} \times \Omega_1) &\rightarrow \mathcal{P}(Y) \\ (x, \omega) &\mapsto \sigma(x, \omega) = \kappa_{(x, \omega)}. \end{aligned} \quad (4.11)$$

We further have $\sigma \circ \psi_1 = T \circ \sigma$. Put $\Theta = \sigma_*(\lambda)$, that is, Θ is the pushforward measure of λ under σ . Whence, $\Theta \in \mathcal{P}(\mathcal{P}(Y))$ is an T -invariant probability measure on \mathcal{P} . We further have

Claim 4.4 *The barycenter of Θ is κ .*

Proof. We start by writing,

$$\int_{\mathcal{P}(Y)} \rho d\Theta(\rho) = \int_{X_{\mathcal{A}_2} \times \Omega_1} (\text{Id}_{\mathcal{P}(Y)} \circ \sigma)(x, \omega) d\lambda(x, \omega) \quad (4.12)$$

$$= \int_{X_{\mathcal{A}_2} \times \Omega_1} \sigma(x, \omega) d\lambda(x, \omega) \quad (4.13)$$

$$= \int_{X_{\mathcal{A}_2} \times \Omega_1} \kappa_{(x, \omega)} d\lambda(x, \omega) \quad (4.14)$$

Whence, for any $h \in C(Y)$, we have

$$\int_{\mathcal{P}(Y)} \int_Y h(z) \rho(z) d\Theta(\rho) = \int_{X_{\mathcal{A}_2} \times \Omega_1} \int_Y h(z) d(\kappa_{(x, \omega)})(z) d\lambda(x, \omega) \quad (4.15)$$

Put

$$H((x, \omega), z) = h(z), \quad \forall (x, \omega) \in X_{\mathcal{A}_2} \times \Omega_1, z \in Y.$$

Then

$$\int H((x, \omega), z) d\xi = \int_{X_{\mathcal{A}_2} \times \Omega_1} \int_Y H((x, \omega), z) d\kappa_{(x, \omega)}(z) d\lambda(x, \omega) \quad (4.16)$$

$$= \int_{X_{\mathcal{A}_2} \times \Omega_1} \int_Y h(z) d\kappa_{(x, \omega)}(z) d\lambda(x, \omega) \quad (4.17)$$

$$= \int_Y h(z) d\kappa(z) \quad (4.18)$$

The last equality follows from the fact that H depends only on the y variable. Summarizing, we have proved

$$\int_{\mathcal{P}(Y)} \int_Y h(z) \rho(z) d\Theta(\rho) = \int_Y h(z) d\kappa(z),$$

and the proof of the claim is complete. \blacksquare

Therefore $(\mathcal{P}(Y), T, \Theta)$ is a quasi-factor of (Y, T, κ) . But, since y is completely deterministic and $\kappa \in I_T(y)$, we get the entropy of the system $(\mathcal{P}(Y), T, \Theta)$ is zero, by Lemma 4.3. We thus deduce that the bounded function

$$F(x, \omega) = \int_Y f(z) d\kappa_{(x, \omega)}(z),$$

is measurable with respect to the Pinsker algebra $\Pi_\lambda(\psi_1) = \alpha^{-1} \Pi_{\alpha\lambda}(\psi_0)$. But, under our assumption, we have

Claim 4.5 $\mathbb{E}(Pr_1 \circ \alpha | \Pi_\lambda(\psi_1)) = 0$.

Proof. We have

$$\mathbb{E}(Pr_1 \circ \alpha | \Pi_\lambda(\psi_1)) = \mathbb{E}(Pr_1 \circ \alpha | \alpha^{-1} \Pi_{\alpha\lambda}(\psi_0)) \quad (4.19)$$

$$= \mathbb{E}(Pr_1 | \Pi_{\alpha\lambda}(\psi_0)) \circ \alpha \quad (4.20)$$

$$= \mathbb{E}(pr_1 \circ \Phi | \Phi^{-1} \Pi_{\Phi\alpha\lambda}(\psi_0)) \circ \alpha \quad (4.21)$$

$$= \mathbb{E}(pr_1 | \Pi_{\Phi\alpha\lambda}(\psi_0)) \circ \Phi \circ \alpha \quad (4.22)$$

$$= 0. \quad (4.23)$$

The last equality follows by our assumption. \blacksquare

We thus conclude that

$$\int Pr_1 \circ \alpha(x, \omega) F(x, \omega) d\lambda(x, \omega) = \int \mathbb{E}(Pr_1 \circ \alpha | \Pi_\lambda(\psi_1)) F(x, \omega) d\lambda(x, \omega) = 0,$$

that is, zero is the only adherent point of the sequence

$$\frac{1}{N} \sum_{n=1}^N \mu(n) f(T^n y).$$

This completes the proof of the theorem. \square

We now proceed to prove our first main result.

Proof of Theorem 3.1

Proof. By Proposition 3.22, for any invariant measure $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$, pr_1 is in the orthocomplement of the L^2 -function measurable with respect to the Pinsker algebra $\Pi(S) = s^{-1}(\mathcal{B}(\mathcal{A}_3))$ of η_M . Moreover, the conditional expectation $\mathbb{E}(\text{pr}_1 |_{\Pi_\eta(S)})$ is in $L^2(X_{\mathcal{A}_3}, \Pi(S), \eta)^\perp \cap L^2(X_{\mathcal{A}_3}, \Pi_\eta(S), \eta)$. Indeed, by the standard properties of the conditional expectation, along with Proposition 3.22, we have

$$\mathbb{E}(\text{pr}_1 |_{\Pi_\eta(S)} |_{\Pi(S)}) = \mathbb{E}(\text{pr}_1 |_{\Pi(S)}) = 0. \quad (4.24)$$

Now, observe that we can apply Lemma 2.3 to the orthocomplement of $L^2(X_{\mathcal{A}_3}, \Pi(S), \eta)$ in $L^2(X_{\mathcal{A}_3}, \mathcal{B}(\mathcal{A}_3), \eta)$, by taking according to Rokhlin-Sinai theorem (see [24, p.65] and [13, Theorem 18.9, p.323]) a sub-algebra \mathcal{D} such $S^n \mathcal{D} \nearrow \mathcal{B}(\mathcal{A}_3)$ and $S^{-n} \mathcal{D} \searrow \Pi_\eta(S) \supset \Pi(S)$, since $\Pi(S)$ cannot be atomic by the ergodicity of ν_M . It follows that the spectral type of pr_1 is absolutely continuous with respect to Lebesgue measure. Finally, by applying Lemma 2.10 combined with Remark 2.5, we get that for any $x \in X$,

$$\limsup \left| \frac{1}{N} \sum_{n=1}^N \boldsymbol{\mu}(n) f(T^n x) \right| \leq \sup G(\sigma_{\text{pr}_1, \eta}, \sigma_{f, \nu_x}),$$

where $\nu_x \in \mathcal{I}_T(x)$, and $\mathcal{I}_T(x)$ is the weak-star closure set of $\left\{ \frac{1}{N} \sum_{n=1}^N \delta_{T^n x} \right\}$. We thus get, by our assumption, that the right-side is zero, and this finishes the proof. \square ■

Remark 4.6 *The famous Chowla conjecture is equivalent to saying that $\boldsymbol{\mu}$ is quasi-generic, hence, generic for the Chowla-Sarnak-Veech measure η_M . Therefore, if Chowla conjecture holds then our assumption in Theorem 3.24 is satisfied and hence Chowla conjecture implies Sarnak's Möbius conjecture.*

We further notice that the point ω_0 in the equation (4.6) is the Liouville function $\boldsymbol{\lambda}$. Furthermore, $\boldsymbol{\mu}$ is generic for Chowla-Sarnak-Veech measure is equivalent to $(\boldsymbol{\mu}^2, \boldsymbol{\lambda})$ is generic for $\nu_M \times b(\frac{1}{2}, \frac{1}{2})$. We thus deduce that Chowla conjecture is equivalent to saying that all the systems $(X_{\mathcal{A}_3}, \mathcal{B}(\mathcal{A}_3), \eta)$, $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$ are reduced to just one dynamical system.

One might think of formulating a weaker form of Chowla conjecture by demanding that for $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$, the dynamical systems $(X_{\mathcal{A}_3}, \mathcal{B}(\mathcal{A}_3), \eta)$ are measure theoretically isomorphic. But, this weak form, by Sarnak-Veech theorem (Theorem 3.21), is actually equivalent to Chowla conjecture.

The above proof actually allows us to view our main theorem differently. In the following corollary we formulate Theorem 3.1 in terms of ‘spectral isomorphism’ and this is where the notation and ideas developed in this section come into play.

Corollary 4.7 *For each $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$, the dynamical system $(X_{\mathcal{A}_3}, \mathcal{B}(\mathcal{A}_3), \eta, S)$ is spectrally isomorphic to $(X_{\mathcal{A}_2} \times \Omega_0, \nu_M \times b(\frac{1}{2}, \frac{1}{2}), \psi_0)$.*

Proof. Our proof of Theorem 3.1 shows that for any $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$ the unitary operator on the closed invariant subspace $L^2(X_{\mathcal{A}_3}, \Pi(S), \eta)^\perp$ has countable Lebesgue spectrum and on $L^2(X_{\mathcal{A}_3}, \Pi(S), \eta)$ has discrete spectrum, independent of η . Thus for any $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$ the system $(X_{\mathcal{A}_3}, \mathcal{B}(\mathcal{A}_3), \eta, S)$ is spectrally isomorphic to $(X_{\mathcal{A}_2} \times \Omega_0, \nu_M \times b(\frac{1}{2}, \frac{1}{2}), \psi_0)$. ■

Remark 4.8

- (1) *W. Veech's theorem proves that the system $(X_{\mathcal{A}_2} \times \Omega_0, \nu_M \times b(\frac{1}{2}, \frac{1}{2}), \psi_0)$ is measure theoretically isomorphic to $(X_{\mathcal{A}_3}, \mathcal{B}(\mathcal{A}_3), \eta_M, S)$. Let us point out also that the spectral type of $(X_{\mathcal{A}_2} \times \Omega_0, \nu_M \times b(\frac{1}{2}, \frac{1}{2}), \psi_0)$ is given by the sum of discrete measure and a countable Lebesgue component. The discrete measure is exactly the spectral measure of the rotation on the group $G = \prod_{p \in \mathcal{P}} \mathbb{Z}/p^2\mathbb{Z}$.*
- (2) *We further notice that if $(X_{\mathcal{A}_3}, \Pi(S), \eta_M, S)$ and $(X_{\mathcal{A}_3}, \Pi_\eta(S), \eta, S)$ are spectrally isomorphic, then again the hypothesis of Veech conjecture holds and hence by the Veech Theorem 3.24, Sarnak conjecture holds. We thus make the following conjecture between Chowla conjecture and Veech's conjecture.*

Conjecture 4.9 *For any $\eta \in \mathcal{I}_S(\boldsymbol{\mu})$, dynamical systems $(X_{\mathcal{A}_3}, \Pi(S), \eta_M, S)$ and $(X_{\mathcal{A}_3}, \Pi_\eta(S), \eta, S)$ are spectrally isomorphic.*

One may ask if this later conjecture is implied by the Veech's conjecture. The answer is yes, since Chowla and Sarnak Möbius conjectures are equivalent by the main result in [3].

5 Other consequences of Theorem 3.8

The following corollary of Theorem 3.8 follows from the spectral theorem.

Corollary 5.1 *Let (X, T, μ) be a compact metric, topological dynamical system. Then for any $F \in L^2(X, \mu)$, for any $f \in C(X_{\mathcal{A}_2})$, we have*

$$\frac{1}{N} \sum_{n=1}^N \boldsymbol{\mu}(n) f(\boldsymbol{\mu}^2(n)) F(T^n x) \xrightarrow{N \rightarrow +\infty} 0.$$

Proof. By the spectral theorem, we can write

$$\left\| \frac{1}{N} \sum_{n=1}^N \boldsymbol{\mu}(n) f(\boldsymbol{\mu}^2(n)) F(T^n x) \right\|_{L^2(X, \mu)} = \left\| \frac{1}{N} \sum_{n=1}^N \boldsymbol{\mu}(n) f(\boldsymbol{\mu}^2(n)) e^{in\theta} \right\|_{L^2(\sigma_F)},$$

where σ_F is the spectral measure of F for the Koopmann operator $F \mapsto F \circ T$. Now, by Theorem 3.8, we have

$$\frac{1}{N} \sum_{n=1}^N \boldsymbol{\mu}(n) f(\boldsymbol{\mu}^2(n)) e^{in\theta} \xrightarrow{N \rightarrow +\infty} 0.$$

Moreover, the sequence $\left(\frac{1}{N} \sum_{n=1}^N \boldsymbol{\mu}(n) f(\boldsymbol{\mu}^2(n)) e^{in\theta} \right)_{N \in \mathbb{N}}$ is bounded. Hence, by the Lebesgue dominated theorem, we obtain

$$\left\| \frac{1}{N} \sum_{n=1}^N \boldsymbol{\mu}(n) f(\boldsymbol{\mu}^2(n)) F(T^n x) \right\|_{L^2(X, \mu)} \xrightarrow{N \rightarrow +\infty} 0.$$

The proof of the corollary is complete. ■

At this point we ask whether in the above corollary the convergence can be almost sure convergence?

In the class of dynamical systems for which every invariant measure has discrete spectrum or the Lebesgue spectrum. this can be done.

Proposition 5.2 *Let (X, T, μ) be a dynamical system for which every invariant measure has discrete spectrum or the Lebesgue spectrum.. Then for any $f \in C(X_{A_2})$ and $F \in C(X)$,*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(n) f(\mu^2(n)) F(T^n x) = 0, \quad \mu \text{ almost surely.}$$

In particular, for $k \geq 2$ and a_1, \dots, a_k be distinct non-negative integers, for any $F \in C(X)$, on a set of points x of full measure we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \mu(n) \mu^2(n + a_1) \cdots \mu^2(n + a_k) F(T^n x) = 0.$$

Proof. We present only the proof for the case in which the spectrum is discrete. For the case of Lebesgue spectrum, we refer to [5, Theorem 4.2.]. Recall that the set of points x which are generic for some ergodic invariant measure have full measure, (see Proposition 5.12) [12]). By our hypothesis, such x is generic for an ergodic measure with discrete spectrum. Hence the map $n \mapsto F(T^n x)$ is Besicovitch almost periodic. ■

For certain very special class of dynamical systems this almost sure convergence can be extended to convergence everywhere.

Proposition 5.3 *Let (X, T, μ) be mean equicontinuous compact metric dynamical system. Then the above convergence holds for any $x \in X$.*

Proof. Since (X, T, μ) is mean equicontinuous, the orbit closure of each $x \in X$ is uniquely ergodic with discrete spectrum, Hence the map $n \rightarrow F(T^n x)$ is Besicovitch almost periodic for any $F \in C(X)$. ■

We thus ask the following question:

Question 5.4 *Do we have for any dynamical system (X, T, μ) with topological entropy zero, for any $f \in C(X_{A_2})$ and $F \in C(X)$, for any $x \in X$,*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \mu(n) f(\mu^2(n)) F(T^n x) = 0?$$

We ask also on the convergence almost everywhere for any dynamical system. Let us further notice that Question 5.4 is a weak form of strong Sarnak Möbius conjecture.

Acknowledgment. The first author would like to thanks Michael Lin for his invitation and for a discussion on the subject. He gratefully thanks Ben Gurion University of the Negev and the Institute of Mathematical Sciences at Chennai for their hospitality, where part of this work was done.

References

- [1] e. H. el Abdalaoui & M. Disertori, Spectral properties of the Möbius function and a random Möbius model. *Stoch. Dyn.* 16 (2016), no. 1, 1650005, 25 pp.
- [2] e. H. el Abdalaoui , M. Lemanczyk & T. de la Rue, Automorphisms with quasi-discrete spectrum, multiplicative functions and average orthogonality along short intervals, *IRMN*, 14 (2017), 43504368.
- [3] e. H. el Abdalaoui, *On Veech's proof of Sarnak's theorem on the Möbius flow*, Preprint, 2017, arXiv:1711.06326 [math.DS].
- [4] e. H. el Abdalaoui and M. Nerurkar, *Note on the spectral characterization of tameness for group action.*, Preprint in preparation, 2019.
- [5] A. Bellow and V. Losert, The weighted pointwise ergodic theorem and the individual ergodic theorem along subsequences, *TAMS* v. 288, no. 1 (1985), 307-345.
- [6] A. Bhattacharyya, *On a measure of divergence between two statistical populations defined by their probability distributions*, *Bulletin of the Calcutta Mathematical Society* 35 (1943), 99–109.
- [7] J. Bourgain, P. Sarnak, and T. Ziegler. Disjointness of Möbius from horocycle flows. In *From Fourier analysis and number theory to radon transforms and geometry*, volume 28 of *Dev. Math.*, p. 6783. Springer, New York, 2013.
- [8] F. Cellarosi & G. Ya. Sinai, Ergodic properties of square-free numbers. *J. Eur. Math. Soc. (JEMS)* 15 (2013), no. 4, 1343-1374.
- [9] J. Coquet, T. Kamae, M. Mendès-France, *Sur la mesure spectrale de certaines suites arithmétiques*, *Bull. Soc. Math. France*, 105 (1977), 369–384.
- [10] S. Chowla, *The Riemann hypothesis and Hilbert's tenth problem*, *Mathematics and Its Applications*, Vol 4, Gordon and Breach Science Publishers, New York, 1965.
- [11] H. Davenport, *On some infinite series involving arithmetical functions. II*, *Quart. J. Math. Oxf.* 8 (1937), 313–320.
- [12] M. Denker, C. Grillenberger, K. Sigmund, *Ergodic Theory on compact spaces*, *Lecture Notes in Mathematics* 527, Springer Verlag Pbl.
- [13] E. Glasner, *Ergodic theory via joinings*. *Mathematical Surveys and Monographs*, 101. American Mathematical Society, Providence, RI, 2003. xii+384 pp.
- [14] B. Host, J.-F. Méla and F. Parreau, *Analyse harmonique des mesures* [Harmonic analysis of measures], *Astérisque*, No. 135-136 (1986), 261 pp.
- [15] W. Huang, Z. Wang; G. Zhang, Möbius disjointness for topological models of ergodic systems with discrete spectrum. *J. Mod. Dyn.* 14 (2019), 277-290.
- [16] W. Huang, Z. Wang and X. Ye, Measure complexity and Möbius disjointness, *Adv. Math.* 347 (2019), 827858.

- [17] K. Matomäki, M. Radziwiłł and T. Tao, An averaged form of Chowla’s conjecture, *Algebra Number Theory*, Vol (9), 2015, 2167-2196.
- [18] K. Matomäki, M. Radziwiłł, Multiplicative functions in short intervals. *Annals of Math.* 183 (2016) 1015-1056.
- [19] K. Matusita, *Decision rules, based on the distance for problems of fit, two samples, and estimation*, *Ann. Math. Statist.* 26 (1955), 631–640.
- [20] K. Matusita, *A distance and related statistics in multivariate analysis*, *Multivariate Analysis (Proc. Internat. Sympos., Dayton, Ohio, 1965)*, 1966, pp. 187–200.
- [21] K. Matusita, *On the notion of affinity of several distributions and some of its applications*, *Ann. Inst. Statist. Math.* 19 1967 181-192.
- [22] L. Mirsky, Arithmetical pattern problems relating to divisibility by r -th powers, *Proc. London Math. Soc.* (2) 50, (1949). 497-508.
- [23] R. Murty and A. Vatwani, A remark on a conjecture of Chowla. *J. Ramanujan Math. Soc.* 33 (2018), no. 2, 111-123.
- [24] W. Parry, *Topics in ergodic theory*, Reprint of the 1981 original. Cambridge Tracts in Mathematics, 75. Cambridge University Press, Cambridge, 2004.
- [25] A. Kanigowski, M. Lemańczyk and M. Radziwiłł, Rigidity in dynamics and Möbius disjointness, Preprint (2019), arXiv:1905.13256.
- [26] L. Ge, Topology of natural numbers and entropy of arithmetic functions, in *Operator Algebras and Their Applications: A Tribute to Richard V. Kadison*, *Contemporary Mathematics*, vol. 671 (American Mathematical Society, Providence, RI, 2016), 127-144.
- [27] V. A. Rohlin, Lectures on the entropy theory of transformations with invariant measure. *Usp. Mat. Nauk.* 22 (1967), 3-56 (Russian)-*Russian Math. Surveys* 22 (1967),1-52.
- [28] P. Sarnak, Möbius randomness and dynamics. *Not. S. Afr. Math. Soc.* 43 (2012), no. 2, 89-97.
- [29] W. Veech, Möbius dynamics, *Lectures Notes*, Spring Semester 2016, +164 pp, private.
- [30] P. Walters, *An introduction to ergodic theory*. Graduate Texts in Mathematics, 79. Springer-Verlag, New York-Berlin, 1982.
- [31] N. Wiener, *The Fourier integral and certain of its applications*, Dover publications, (1958).
- [32] F. Wei, Disjointness of Möbius from asymptotically periodic functions, arXiv:1810.07360v5 [math.NT].