Pure and Applied Mathematics Quarterly
Volume 18, Number 3, 863–922, 2022

Disjointness of Möbius from asymptotically periodic functions

Fei Wei

Abstract: We investigate Sarnak's Möbius Disjointness Conjecture through asymptotically periodic functions. It is shown that Sarnak's conjecture for rigid dynamical systems is equivalent to the disjointness of Möbius from asymptotically periodic functions. We give sufficient conditions and a partial answer to the later one. As an application, we show that Sarnak's conjecture holds for a class of rigid dynamical systems, which improves an earlier result of Kanigowski-Lemańczyk-Radziwiłł.

Keywords: Asymptotically periodic function, mean state, Möbius function, Sarnak's Möbius Disjointness Conjecture.

1. Introduction

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$ denote the set of natural numbers and $\mathbb{N}^* = \{1, 2, \ldots\}$. Functions from $\mathbb{N}$ (or $\mathbb{N}^*$) into $\mathbb{C}$ are called *arithmetic functions*. Many problems in number theory can often be reformulated in terms of properties of arithmetic functions. For example, the Möbius function $\mu(n)$ is defined by 0 if n is not square free (i.e., divisible by a nontrivial square), and $(-1)^r$ if n is the product of r distinct primes. It is well known that the Prime Number Theorem is equivalent to that $\sum_{n\leq x} \mu(n) = o(x)$; the Riemann Hypothesis holds if and only if $\sum_{n\leq x} \mu(n) = o(x^{\frac{1}{2}+\epsilon})$, for any $\epsilon > 0$.

An arithmetic function f is said to be *disjoint* from another one g if $\sum_{n=1}^{N} f(n)\overline{g}(n) = o(N)$. Disjointness is a commonly concerned relation between arithmetic functions. The disjointness of Möbius from arithmetic functions plays an important role in number theory since they reflect certain random distribution among the values of the Möbius function and are closely related to the distribution of primes. For example, the disjointness of $\mu(n)$

arXiv: 1810.07360
Received August 18, 2021.
2010 Mathematics Subject Classification: Primary 37A55; secondary 11N37.

from periodic functions is equivalent to the prime number theorem in arithmetic progressions. Sarnak ([35]) conjectured that the Möbius function is disjoint from all arithmetic functions arising from any topological dynamical systems with zero topological entropy. More specifically,

Conjecture 1 (Sarnak's Möbius Disjointness Conjecture (SMDC))**.** *Let X be a compact Hausdorff space and T a continuous map on X with zero topological entropy, then*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu(n)f(T^n x_0)=0$$

for any $x_0\in X$ and $f\in C(X)$.

In recent years, a lot of progress have been made on Conjecture 1. See [2, 3, 11, 12, 18, 19, 20, 23, 26, 27, 33, 37, 39, 42, 43], to list a few. In the following, we shall discuss only the results that are more related to this paper. Sarnak proved that SMDC is implied by Chowla's conjecture which is stated as follows [6].

Conjecture 2 (Chowla's conjecture)**.** *Let $a_0, a_1, a_2, \dots, a_m$ be distinct natural numbers, and $i_s\in\{1,2\}$ for $s=0,1,2,\dots,m$, not all i_s are even numbers. Then*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu^{i_0}(n+a_0)\mu^{i_1}(n+a_1)\cdots\mu^{i_m}(n+a_m)=0.$$

Chowla's conjecture is a longstanding open problem in number theory. It is open even in one of its simplest forms: $\sum_{n=1}^{N}\mu(n)\mu(n+2)=o(N)$. This estimate should be closely related to the twin prime conjecture.

1.1. Asymptotically periodic functions

In order to use tools from operator algebra to study Sarnak's conjecture, Ge introduced the following notion of asymptotically periodic function in the survey paper [13] and proved that the Möbius function is disjoint from certain asymptotically periodic functions.

Definition 1.1. A function $f\in l^\infty(\mathbb{N})$ is called **asymptotically periodic**[1] if for any mean state E, there is a sequence $\{n_j\}_{j=1}^{\infty}$ of positive integers such that $f-A^{n_j}f$ has limit zero in $\mathcal{H}_E$.

[1] This definition is a little weaker than [13, Definition 5.7], in which the sequence $\{n_j\}_{j=1}^{\infty}$ is independent of E.

In the above definition, the action A on $l^\infty(\mathbb{N})$, the algebra of all bounded arithmetic functions endowed with the pointwise addition and multiplication, is defined as

$$Af(n) = f(n+1), \tag{1}$$

for all $f \in l^\infty(\mathbb{N})$ and $n \in \mathbb{N}$. The mean states E on $l^\infty(\mathbb{N})$ are given by certain limits of $\frac{1}{N}\sum_{n=0}^{N-1} f(n)$ along "ultrafilters" and $\mathcal{H}_E$ is the Hilbert space obtained by the GNS construction on $l^\infty(\mathbb{N})$ with respect to E. We refer readers to Section 4 for more details.

In this paper, we further study properties of asymptotically periodic functions, the Möbius disjointness of asymptotically periodic functions and give some applications of these results to Sarnak's conjecture. We first introduce the following subclass of asymptotically periodic functions.

Definition 1.2. A function $f \in l^\infty(\mathbb{N})$ is called **strongly asymptotically periodic** if for any mean state E, there is a sequence $\{n_j\}_{j=1}^\infty$ of positive integers such that when j goes to infinity, $f - A^{ln_j}f$ converges to zero in $\mathcal{H}_E$ uniformly with respect to all $l \in \mathbb{N}$.

Here are some examples. The function $e^{2\pi i\sqrt{n}}$ is a strongly asymptotically periodic function. For any strictly increasing sequence $\{N_j\}_{j=0}^\infty$ ($N_0 = 0$) and any bounded sequence $\{a_j\}_{j=0}^\infty$ of complex numbers, define $f(n) = a_j$ when $N_j \le n < N_{j+1}$. Then f is strongly asymptotically periodic. If θ is an irrational number, then $e^{2\pi in\theta}$ is an asymptotically periodic function but not in the strong sense. The function $e^{2\pi in^2\theta}$ with θ irrational is disjoint from all asymptotically periodic functions. These results and more examples of strongly asymptotically periodic functions are shown in Section 4.

In [8], Eberlein introduced the notion of weakly almost periodic (WAP) functions. These functions have been studied in dynamical systems (see e.g., [15, 36]). Moreover, all these functions can be realized in topological dynamical systems with zero topological entropy (see [37, Theorem 9.1]). We shall show that WAP functions belong to the class of asymptotically periodic functions satisfying conditions (4) and (5) below, see Proposition 5.5.

Interestingly, there are strongly asymptotically periodic functions that cannot be realized in any topological dynamical system with zero topological entropy. In the following we give an example to illustrate it.

Proposition 1.3. *Suppose $s \ge 1$ and $m_1, \ldots, m_s \in \mathbb{N}$ with at least one $m_i \ge 1$. Let $f = A^{m_1}(\mu^2)\cdots A^{m_s}(\mu^2)$. Then $f(n)$ is a strongly asymptotically periodic functions. Moreover, $f(n)$ cannot be realized in any topological dynamical system with zero topological entropy, i.e., there does not exist a* *topological dynamical system* (X, T) *such that the topological entropy of* T *is zero and* $f(n) = F(T^n x_0)$ *for some* $F \in C(X)$ *and* $x_0 \in X$.

1.2. The Möbius disjointness of asymptotically periodic functions

We are interested in the following problem.

Problem 1. *Is* μ *disjoint from all asymptotically periodic functions?*

We now explain a motivation for us to investigate the above problem. The positive answer to Chowla's conjecture implies that the set $\{A^n\mu : n \geq 0\}$ is an orthogonal set (with respect to a given mean state) of vectors of the same norm. Denote the norm of μ as c. By Bessel's inequality, we have

$$\langle f, f\rangle \geq \frac{1}{c^2}\sum_{n\in\mathbb{N}} |\langle A^n\mu, f\rangle|^2 \tag{2}$$

for any $f \in l^\infty(\mathbb{N})$. Assume that f is an asymptotically periodic function, then there exists a sequence $\{n_j\}_{j=1}^\infty$ of distinct positive integers such that $\lim_{j\to\infty} |\langle A^{n_j}\mu, f - A^{n_j} f\rangle| = 0$. This implies that $\lim_{j\to\infty} |\langle A^{n_j}\mu, f\rangle| = |\langle \mu, f\rangle|$, then $\langle \mu, f\rangle = 0$ by the inequality (2).

The process of exploring Problem 1 motivates us to study the average value of the Möbius function in short arithmetic progressions. Precisely, we should estimate the second moment of this average: $\sum_{n=1}^N |\sum_{l=1}^h \mu(n + kl)|^2$. About this, we show the following result.

Theorem 1.4. *Let* k *be a positive integer. Then for any* $h \geq 3$,

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \left|\sum_{l=1}^h \mu(n + kl)\right|^2 \ll \frac{k}{\varphi(k)}\frac{\log\log h}{\log h}h^2. \tag{3}$$

Throughout this paper, $f \ll g$ means that there is an absolute constant c, such that $|f| \leq c|g|$; $f = g + O(h)$ means $f - g \ll h$. We use $\varphi(k)$ to denote the Euler totient function.

We are more concerned about whether the left hand side of formula (3) is still $o(h^2)$ when k is far large than h. The estimate presented in formula (3) implies that this holds for k as large as $\exp(h^{o(1)})$ since $k/\varphi(k) \ll \log\log k$. We expect that the right hand side of formula (3) is $o(h^2)$ independent of $k \geq 1$. This is likely to be true because the positive answer to Chowla's conjecture implies that the left hand side of formula (3) should be $\frac{6}{\pi^2}h$.

In Theorem 1.4, we can replace μ by non-pretentious 1-bounded multiplicative functions such as the Liouville functions and $\mu(n)\chi(n)$, where χ is a

Dirichlet character, see Proposition 6.1. Moreover, we recently extended Theorem 1.4 to the case that $\mu(n)$ is replaced by $\mu(n)e(P(n))$ for any $P(x) \in \mathbb{R}[x]$ ([42]), and this is possible to be generalized to $\mu(n)$ twisted by any nilsequence.

Using Theorem 1.4, we give a partial answer to Problem 1, which states that $\mu(n)$ is disjoint from a class of asymptotically periodic functions. Precisely,

Theorem 1.5. *Let $f \in l^\infty(\mathbb{N})$ satisfying that for any mean state E, there are sequences $\{h_j\}_{j=1}^\infty$ and $\{n_j\}_{j=1}^\infty$ of positive integers with*

$$\lim_{j\to\infty} \frac{\log\log h_j}{\log h_j}\frac{n_j}{\varphi(n_j)} = 0 \tag{4}$$

such that

$$\lim_{j\to\infty} \frac{1}{h_j}\sum_{l=1}^{h_j} E(|f - A^{ln_j}f|^2) = 0. \tag{5}$$

Then we have

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

By the definition of strongly asymptotically periodicity, it is not hard to check the following result as an application of Theorem 1.5.

Corollary 1.6. *Problem 1 holds for all strongly asymptotically periodic functions.*

For solving Problem 1 completely, we provide a sufficient condition as follows.

Proposition 1.7. *Assume that*

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N}\left|\sum_{l=1}^{h}\mu(n+kl)\right|^2 = o(h^2), \tag{6}$$

where the little "o" term is independent of $k \geq 1$. Then Problem 1 holds.

It is unknown if the converse of the above proposition is true. It is proved in Proposition 4.9 that the disjointness of Möbius from all strongly asymptotically periodic functions is equivalent to that for any given k,

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N}\left|\sum_{l=1}^{h}\mu(n+kl)\right|^2 = o(h^2).$$

Although many asymptotically periodic functions cannot be realized in topological dynamical systems with zero topological entropy, they can be approximated measure-theoretically by realizable functions (see Theorem 7.2). This leads to the following result.

Theorem 1.8. *Assume that Sarnak's Möbius Disjointness Conjecture is true, then Problem 1 holds.*

1.3. Applications to Sarnak's conjecture for rigid dynamical systems

Before introducing more results, we first recall the definition of rigid dynamical system. Let $(X, \mathcal{B}, \nu, T)$ be a measure-preserving dynamical system, i.e., X is a compact metric space, T a continuous map on X, $\mathcal{B}$ the Borel σ-algebra of subsets of X and ν a T-invariant Borel probability measure. Such a dynamical system is called *rigid* if there is a sequence $\{n_j\}_{j=1}^{\infty}$ of positive integers such that for any $f \in L^2(X, \nu)$,

$$\lim_{j\to\infty} \|f \circ T^{n_j} - f\|_{L^2(\nu)}^2 = 0.$$

Rigid dynamical systems contain dynamical systems with discrete spectrum and a large class of skew products on the torus over a rotation of the circle [25]. In the following for simplicity, we use (X, ν, T) to denote a measure-preserving dynamical system.

From the viewpoint of dynamical systems, asymptotically periodic functions correspond to rigid measure-preserving dynamical systems (see Theorem 5.3). The major tool we use to build this connection between arithmetics and dynamics is anqie (of natural numbers), which was introduced by Ge in [13]. We refer readers to Section 3 for knowledge on anqie. Based on this connection, corresponding to Problem 1, it is natural to consider the following problem.

Problem 2 (Sarnak's conjecture for rigid dynamical systems)**.** Let X be a compact metric space and T a continuous map on X. Suppose $x_0 \in X$ satisfies the following condition: for any ν in the weak* closure of $\{\frac{1}{N}\sum_{n=0}^{N-1} \delta_{T^n x_0} : N = 1, 2, \ldots\}$ in the space of Borel probability measures on X, there is a dense set $\mathcal{F} \subseteq C(X)$, such that for each $g(x) \in \mathcal{F}$ we can find a sequence $\{n_j\}_{j=1}^{\infty}$ (may depend on ν, g) of positive integers satisfying

$$\lim_{j\to\infty} \|g \circ T^{n_j} - g\|_{L^2(\nu)}^2 = 0. \tag{7}$$

Is it true that for any $f \in C(X)$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(n) f(T^n x_0) = 0?$$

Proposition 1.9. *Problem 1 holds if and only if Problem 2 holds.*

When ν satisfies the condition in Problem 2, (X, T, ν) is rigid and then has zero measure-theoretic entropy (see e.g., [34, Example 5.3.3]), while (X, T) may not have zero topological entropy, see the paragraphs below Proposition 7.1 for an example. Recently, in [23], Kanigowski, Lemańczyk and Radziwiłł gave a partial answer to Problem 2.

Theorem 1.10. *[23, Theorem 2.1] With the same assumptions as Problem 2, if T is a homeomorphism and for each $g(x) \in \mathcal{F}$ we can find a sequence $\{n_j\}_{j=1}^{\infty}$ (may depend on ν, g) of positive integers satisfying either*

(BPV rigidity)*: there is a constant $c > 0$ such that $\sum_{p|n_j} \frac{1}{p} < c$ for any $j = 1, 2, \ldots$, and*

$$\lim_{j\to\infty} \|g \circ T^{n_j} - g\|_{L^2(\nu)}^2 = 0.$$

or

(PR rigidity)*: for some $\delta > 0$, the following holds:*

$$\lim_{j\to\infty} \sum_{l=-n_j^\delta}^{n_j^\delta} \|g \circ T^{l n_j} - g\|_{L^2(\nu)}^2 = 0.$$

Then for any $f \in C(X)$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(n) f(T^n x_0) = 0.$$

Employing the estimate we obtained in Theorem 1.4, we improve the above theorem to the following.

Theorem 1.11. *With the same assumptions as Problem 2, if T is a continuous map and for each $g(x) \in \mathcal{F}$, there are sequences $\{h_j\}_{j=1}^{\infty}$ and $\{n_j\}_{j=1}^{\infty}$ of positive integers with*

$$\lim_{j\to\infty} \frac{\log\log h_j}{\log h_j} \frac{n_j}{\varphi(n_j)} = 0 \tag{8}$$

satisfying

$$\lim_{j\to\infty}\frac{1}{h_j}\sum_{l=1}^{h_j}\|g\circ T^{ln_j}-g\|^2_{L^2(\nu)}=0. \tag{9}$$

Then for any $f\in C(X)$,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu(n)f(T^nx_0)=0.$$

Moreover, the above disjointness also holds over short intervals in average, that is

$$\lim_{h\to\infty}\limsup_{N\to\infty}\frac{1}{Nh}\sum_{n=1}^{N}\Big|\sum_{l=1}^{h}\mu(n+l)f(T^{n+l}x_0)\Big|=0.$$

In comparison with Theorem 1.10, we relax T to a continuous map. Also both BPV rigidity and PR rigidity are included in the scenario described by conditions (8), (9). See Remark 8.1 for details. There are examples that satisfy conditions (8), (9), but not BPV rigidity and PR rigidity (see Remark 8.3). Indeed, we show that conditions (8), (9) hold for any (X,ν,T) with discrete spectrum in Proposition 5.4, while the set of these dynamical systems are not strictly contained in the set of rigid dynamical systems satisfying BPV rigidity or PR rigidity (see Remark 8.4).

Related to the above result, we recently proved that Sarnak's conjecture holds for product flows between rigid dynamical systems satisfying conditions in Theorem 1.11 and affine linear flows on compact abelian groups of zero topological entropy [42].

Our paper is organized as follows. In Section 2, we list some frequently used notation, and prove some preliminary results. In Section 3, we study properties of anqies and describe the topological characterizations of anqies in terms of the generating arithmetic functions. We perform the GNS constructions on anqies, and show some examples of asymptotically and strongly asymptotically periodic functions in Section 4, where Proposition 1.3 is proved. In Section 5, we study the connections between arithmetic functions and measure-preserving dynamical systems. In Section 6, we show the estimate about the self-correlations of the Möbius stated in Theorem 1.4. In Section 7, we prove Theorems 1.5, 1.8, and Proposition 1.7. As applications, we prove Proposition 1.9 and Theorem 1.11 in Section 8.

This work arose as part of my Ph.D. thesis at the Chinese Academy of Sciences [40] under the supervision of Professor Liming Ge. We refer to [7, 22]

for basics and preliminary results in operator algebra, to [14, 21] for that on topological dynamics and number theory.

2. Preliminaries

In this section, we prove some preliminary results. First, we list some notation that will be used.

Let $\mathcal{H}$ be a Hilbert space. Denote by $\mathcal{B}(\mathcal{H})$ the algebra consists of all bounded linear operators on $\mathcal{H}$. By Riesz representation theorem, for any $T \in \mathcal{B}(\mathcal{H})$, there is a unique bounded linear operator T^* satisfying $\langle Tx, y\rangle = \langle x, T^*y\rangle$ for any $x, y \in \mathcal{H}$. Such a T^* is called the *adjoint of* T. We call a norm-closed *-subalgebra of $\mathcal{B}(\mathcal{H})$ a *C*-algebra*. In this paper, we always assume that all C*-algebras are unital.

Suppose that $\mathcal{A}$ is a C*-algebra. We use $\mathcal{A}^\sharp$ to denote the set of all bounded linear functionals on $\mathcal{A}$. Denote by $(\mathcal{A}^\sharp)_1$ the unit ball in $\mathcal{A}^\sharp$, i.e., $(\mathcal{A}^\sharp)_1 = \{\rho \in \mathcal{A}^\sharp : \|\rho\| \leq 1\}$. In general, the space $\mathcal{A}^\sharp$ can be equipped with many topological structures. Among them, the norm topology and weak* topology are used most frequently. For $\rho \in \mathcal{A}^\sharp$, its norm is given by $\|\rho\| = \sup_{x\in\mathcal{A},\|x\|\leq 1} |\rho(x)|$. When $x \in \mathcal{A}$, the equation $\sigma_x(\rho) = |\rho(x)|$ defines a semi-norm on $\mathcal{A}^\sharp$. The family $\{\sigma_x : x \in \mathcal{A}\}$ of semi-norms determines the *weak* topology* on $\mathcal{A}^\sharp$. Note that each $\rho_0 \in \mathcal{A}^\sharp$ has a base of neighborhoods consisting of sets of the form $\{\rho \in \mathcal{A}^\sharp : |\rho(x_j) - \rho_0(x_j)| < \epsilon\}$ $(j = 1, \ldots, m)$, where $\epsilon > 0$ and $x_1, \ldots, x_m \in \mathcal{A}$.

A non-zero linear functional ρ on an abelian C*-algebra $\mathcal{A}$ is called *a multiplicative state* if for any $A, B \in \mathcal{A}$, $\rho(AB) = \rho(A)\rho(B)$.

Suppose now that $\mathcal{A}$ is an abelian C*-algebra and X is its maximal ideal space. We define the map $\gamma : \mathcal{A} \to C(X)$ by

$$\gamma(f)(\rho) = \rho(f), \quad f \in \mathcal{A},\ \rho \in X. \tag{10}$$

Here we use the fact that X is also the space of all multiplicative states of $\mathcal{A}$. The map γ is known as the *Gelfand transform* from $\mathcal{A}$ onto $C(X)$, which is a *-isomorphism (see, e.g., [7, Theorem 2.1]).

It is known that the above Hausdorff space X is weak* compact. Next we show that $\mathcal{A}$ is countably generated as an abelian C*-algebra if and only if X is metrizable and the topology induced by the metric coincides with the weak* topology. The sufficient part directly follows from [22, Remark 3.4.15]. The necessary part is shown in the following proposition.

Proposition 2.1. *Let $\mathcal{A}$ be an abelian C*-algebra. If $\mathcal{A}$ is countably generated, then $(\mathcal{A}^\sharp)_1$ is metrizable and the topology induced by the metric is* *equivalent to the weak* topology on* $(\mathcal{A}^\sharp)_1$. *In particular, the maximal ideal space of* $\mathcal{A}$ *is a compact metrizable space.*

Proof. Since $\mathcal{A}$ is countably generated, there is a countable dense subset in $\mathcal{A}$. Let $\{g_1, g_2, \ldots\}$ be a dense subset of $(\mathcal{A})_1$, the unit ball in $\mathcal{A}$. For any $\rho_1, \rho_2 \in (\mathcal{A}^\sharp)_1$, we define $d(\rho_1, \rho_2) = \sum_{i=1}^{\infty} \frac{|(\rho_1 - \rho_2)(g_i)|}{2^i}$. It is not hard to check that d is a metric on $(\mathcal{A}^\sharp)_1$. Moreover, for any net $\{\rho_\alpha\}$ of elements of $(\mathcal{A}^\sharp)_1$, the net $\{d(\rho_\alpha, \rho)\}$ converges to 0 is equivalent to the condition that, for any $i \geq 1$, the net $\{\rho_\alpha(g_i)\}$ converges to $\rho(g_i)$.

Next, we show that the weak* topology is equivalent to the topology induced by the metric d on $(\mathcal{A}^\sharp)_1$. Suppose that the net $\{\rho_\alpha\}$ of elements in $(\mathcal{A}^\sharp)_1$, weak* converges to ρ. Then, for any $i \geq 1$, the net $\{\rho_\alpha(g_i)\}$ converges to $\rho(g_i)$. Thus the net $\{d(\rho_\alpha, \rho)\}$ converges to 0. Conversely, if the net $\{d(\rho_\alpha, \rho)\}$ converges to 0, where $\rho_\alpha \in (\mathcal{A}^\sharp)_1$, then $\{\rho_\alpha(g_i)\}$ converges to $\rho(g_i)$ for any $i \geq 1$. Note that, for any α, $\|\rho_\alpha\| \leq 1$. Then for any $g \in \mathcal{A}$, the net $\{\rho_\alpha(g)\}$ converges to $\rho(g)$. So the net $\{\rho_\alpha\}$ is weak* convergent to ρ in $(\mathcal{A}^\sharp)_1$.

By Alaoglu-Bourbaki theorem $(\mathcal{A}^\sharp)_1$ is weak* compact. Let X be the maximal ideal space of $\mathcal{A}$. Then, relative to the weak* topology, X is a closed subset of $(\mathcal{A}^\sharp)_1$. From the above analysis, we see that the weak* topology on $(\mathcal{A}^\sharp)_1$ coincides with the topology induced by the metric d on it. Thus X is a compact metrizable space. $\square$

Proposition 2.2. *Suppose that* $\mathcal{A}$ *is a C*-subalgebra of* $l^\infty(\mathbb{N})$ *and* X *the maximal ideal space of* $\mathcal{A}$. *Let* $\iota : \mathbb{N} \to X$ *be the map given by*

$$\iota(n) : f \mapsto f(n), \tag{11}$$

for any $f \in \mathcal{A}$. *Then the weak* closure of* $\iota(\mathbb{N})$ *is* X *(write* $\overline{\iota(\mathbb{N})} = X$*).*

Proof. Assume on the contrary that $\overline{\iota(\mathbb{N})} \neq X$. Choose $y \in X \setminus \overline{\iota(\mathbb{N})}$. By Urysohn's lemma, there is a $G \in C(X)$ such that $G(y) = 1$ and $G(x) = 0$ for any $x \in \overline{\iota(\mathbb{N})}$. By equation (10), for any $n \in \mathbb{N}$, $0 = G(\iota(n)) = \iota(n)(\gamma^{-1}G) = (\gamma^{-1}G)(n)$. Then $\gamma^{-1}(G) = 0$ and $G = 0$ correspondingly. This contradicts $G(y) = 1$. Hence $\overline{\iota(\mathbb{N})} = X$. $\square$

If ι is injective, then we can view $\mathbb{N}$ as a subset of X. For $\mathcal{A} = l^\infty(\mathbb{N})$, ι is injective. We shall use $\beta\mathbb{N}$ to denote the maximal ideal space of $l^\infty(\mathbb{N})$, which is also known as the Stone-Čech compactification of $\mathbb{N}$ [4]. Since $l^\infty(\mathbb{N})$ is not a separable C*-algebra, it is not countably generated as a C*-algebra. Correspondingly, the maximal ideal space $\beta\mathbb{N}$ is not metrizable.

3. Anqie of $\mathbb{N}$

In this section, we briefly introduce the concept of anqie and list some results that will be used later. For more about anqie, we refer to [13] and [41].

Definition 3.1. Let X be a compact Hausdorff space and ι a map from $\mathbb{N}$ to X with dense range. We call X **an anqie** (of $\mathbb{N}$) if $\iota(n) \mapsto \iota(n+1)$ is a well-defined map on $\iota(\mathbb{N})$ and it can be extended to a continuous map from X into itself. If we denote this extended map by σ_A, we also call (X, σ_A) **an anqie** (of $\mathbb{N}$).

We now explain a little about the above notion. For a general map $\iota : \mathbb{N} \to X$, $\iota(n) \mapsto \iota(n+1)$ may not be well-defined, such as $\iota : \mathbb{N} \to S^1$ (the unit circle) defined as $\iota(n) = e^{2\pi i \sqrt{n}}$. Even though $\iota(n) \mapsto \iota(n+1)$ is well defined, it may not induce a continuous map on X, such as $\iota : \mathbb{N} \to S^1$ defined as $\iota(n) = e^{2\pi i n^2 \theta}$ with θ irrational. So an anqie of $\mathbb{N}$ preserves the addition structure of natural numbers when $\mathbb{N}$ is mapped to X. Here is a simple example of anqie.

Example 3.2. Let θ be an irrational number with $0 < \theta < 1$. Define $\iota : n \mapsto e^{2\pi i n\theta}$, a map from $\mathbb{N}$ into S^1. It is easy to see that ι has a dense range in S^1 and $e^{2\pi i n\theta} \mapsto e^{2\pi i(n+1)\theta} = e^{2\pi i\theta} e^{2\pi i n\theta}$ induces a continuous map $z \mapsto e^{2\pi i\theta} z$, denoted by σ_A on S^1. Thus (S^1, σ_A) is an anqie of $\mathbb{N}$.

Next we consider how to construct anqies of $\mathbb{N}$. One way to obtain anqies of $\mathbb{N}$ is to construct point transitive topological dynamical systems. Recall that a topological dynamical system (or, equivalently an $\mathbb{N}$-dynamics) is a pair (X, T), where X is a compact Hausdorff space and T a continuous map on X. Suppose that (X, T, x_0) is a point transitive topological dynamical system, i.e., the set $\{T^n x_0 : n \in \mathbb{N}\}$ is dense in X. Then $\iota : n \mapsto T^n x_0$ is a map from $\mathbb{N}$ to X with dense range. It is easy to see that $\iota(n) \mapsto \iota(n+1)$ can be extended to the continuous map T on X. Then (X, T) is an anqie of $\mathbb{N}$. Summarize the above analysis, we conclude that an $\mathbb{N}$-dynamics (X, T) is an anqie of $\mathbb{N}$ if it is point transitive. In this construction, the structure of anqie depends on the choice of the transitive point.

Another method to construct anqies is through C*-algebras.

Proposition 3.3. *Suppose that $\mathcal{A}$ is a C*-subalgebra of $l^\infty(\mathbb{N})$ and X the maximal ideal space of $\mathcal{A}$. Then X is an anqie of $\mathbb{N}$ with the map ι given by equation (11) if and only if $\mathcal{A}$ is closed under the action A defined in (1), i.e., $Af \in \mathcal{A}$ for any $f \in \mathcal{A}$.*

Proof. Suppose that X is an anqie of $\mathbb{N}$. Then the map $\iota(n) \mapsto \iota(n+1)$ is extended to a continuous map on X, denoted by σ_A. Given $f \in \mathcal{A}$, assume that $F = \gamma(f) \in C(X)$ (see equation (10)). Note that $F \circ \sigma_A \in C(X)$. Let $g = \gamma^{-1}(F \circ \sigma_A)$ in $\mathcal{A}$. Then $g(n) = F \circ \sigma_A(\iota(n)) = F(\iota(n+1)) = f(n+1) = Af(n)$. Thus $g = Af$ in $\mathcal{A}$. This shows that $\mathcal{A}$ is A-invariant.

On the other hand, suppose that $\mathcal{A}$ is closed under A. Let σ_A be the map from X to itself given by $\sigma_A\rho(f) = \rho(Af)$ for any $\rho \in X$ and $f \in \mathcal{A}$. It is easy to see that $\sigma_A(\iota(n)) = \iota(n+1)$. Now we show that σ_A is a continuous map on X. If $\{\rho_\alpha\}$ is a weak* convergent net of elements of X, with limit ρ, then for any $f \in \mathcal{A}$, $\rho_\alpha(Af) = (\sigma_A\rho_\alpha)(f)$ converges to $\rho(Af) = (\sigma_A\rho)(f)$. Thus the net $\{\sigma_A\rho_\alpha\}$ weak* converges to $\sigma_A\rho$ in X. Hence σ_A is the continuous map on X extended by $\iota(n) \mapsto \iota(n+1)$ and X is an anqie of $\mathbb{N}$. $\square$

From the above proposition, we can obtain anqies of $\mathbb{N}$ through constructing A-invariant C*-subalgebras of $l^\infty(\mathbb{N})$. In particular, we often consider the anqie generated by a single arithmetic function f, i.e., the C*-algebra generated by $\{1, A^j f : j \in \mathbb{N}\}$. Denote it by $\mathcal{A}_f$. We use X_f to denote the maximal ideal space of $\mathcal{A}_f$. From Proposition 3.3, we know that $\iota(n) \mapsto \iota(n+1)$ can be extended to a continuous map on X_f, denoted by σ_A. We also call (X_f, σ_A) or $\mathcal{A}_f$ *the anqie generated by* f. Let $\overline{f(\mathbb{N})}$ denote the closure of $f(\mathbb{N})$ in the complex plane $\mathbb{C}$. Since $\mathcal{A}_f$ contains f, there is a continuous map from X_f onto $\overline{f(\mathbb{N})}$. But these two spaces may not be the same.

The following theorem describes X_f in terms of $\overline{f(\mathbb{N})}$ and gives a representation of σ_A (corresponding to the Bernoulli shift on a product space).

Theorem 3.4. *[13, Theorem 2.3]Suppose that f is a function in $l^\infty(\mathbb{N})$ and that X_f is the maximal ideal space of the anqie $\mathcal{A}_f$ generated by f. Denote by $\prod_{\mathbb{N}} \overline{f(\mathbb{N})}$ the Cartesian product of $\overline{f(\mathbb{N})}$ indexed by $\mathbb{N}$, endowed with the product topology. Assume that B is the Bernoulli shift on $\prod_{\mathbb{N}} \overline{f(\mathbb{N})}$ defined by*

$$B : (a_0, a_1, a_2, \ldots) \mapsto (a_1, a_2, a_3, \ldots).$$

Let F be the map from X_f into $\prod_{\mathbb{N}} \overline{f(\mathbb{N})}$, such that for any $\rho \in X_f$,

$$F(\rho) = (\rho(f), \rho(Af), \ldots). \tag{12}$$

The following statements hold.

(i) The space X_f is homeomorphic to $F(X_f)$.

(ii) $F(X_f)$ is the closure of $\{(f(n), f(n+1), \ldots) : n \in \mathbb{N}\}$ in $\prod_{\mathbb{N}} \overline{f(\mathbb{N})}$.

(iii) The restriction of the Bernoulli shift B on $F(X_f)$ is identified with σ_A on X_f.

Remark 3.5. It follows from the above theorem that for $f \in l^\infty(\mathbb{N})$, X_f can be identified as the set of all pointwise limits of sequences $\{A^n f,\ n = 0, 1, 2, \ldots\}$ in $l^\infty(\mathbb{N})$. This still holds when the semigroup $\mathbb{N}$ is replaced by the group $\mathbb{Z}$ or a general abelian topological group G. While for the case of $\mathbb{Z}$ or G, X_f has been extensively studied in dynamical systems (see e.g., [8], [36] and [15]). In [8], Eberlein introduced the concept of weakly almost periodic function, i.e., $f \in l^\infty(G)$ with X_f weak compact in $l^\infty(G)$. Let $W(G)$ denote the set of these functions. In [36], Veech introduced a *-subalgebra $K(G)$ of $l^\infty(G)$ consisting of all $f \in l^\infty(G)$ with X_f norm separable in $l^\infty(G)$, which contains $W(G)$. Recently, the Möbius disjointness of $W(\mathbb{Z})$ has been proved in [37] and that of $K(\mathbb{Z})$ has been proved in [20].

Applying Theorem 3.4, we can obtain many interesting examples of anqies (X_f, σ_A). The following two examples are given in [13, 41].

Example 3.6. Let $f(n) = e^{2\pi i\sqrt{n}}$, for $n \in \mathbb{N}$. Then X_f is homeomorphic to $\{e^{-\frac{1}{n}} f(n) : n \in \mathbb{N}\} \cup S^1$, a subset of $\mathbb{C}$, denoted by X. And σ_A is the identity map on S^1, while, on the set $\{e^{-\frac{1}{n}} f(n) : n \in \mathbb{N}\}$, σ_A maps $e^{-\frac{1}{n}} f(n)$ to $e^{-\frac{1}{n+1}} f(n+1)$.

Example 3.7. Suppose that θ is irrational and $f(n) = e^{2\pi i n^2\theta}$, then X_f is homeomorphic to $S^1 \times S^1$. Moreover, if we identify $S^1 \times S^1$ with $\mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}$, then we can rewrite the map σ_A as

$$\sigma_A(\alpha_1, \alpha_2) = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 2\theta \end{pmatrix}.$$

At the end of this section, we state the following result which will be used in later parts of this paper. Recall that for two topological dynamical systems (X_1, T_1) and (X_2, T_2), if there is a continuous surjective map φ from X_1 onto X_2 such that $\varphi \circ T_1 = T_2 \circ \varphi$, we call φ *a factor map* and (X_2, T_2) a *factor* of (X_1, T_1). Moreover, if φ is a homeomorphism, we say that (X_1, T_1) and (X_2, T_2) are *(topologically) conjugate (to each other)*.

Proposition 3.8. *Let f be an arithmetic function realized in (X, T), i.e., there is a continuous function $g \in C(X)$ and $x_0 \in X$ such that $f(n) = g(T^n x_0)$. Suppose that X_f is the maximal ideal space of the anqie generated by f. Let Y be the closure of the set $\{T^n x_0 : n \in \mathbb{N}\}$ in X. Then (X_f, σ_A) is a factor of (Y, T).*

Proof. Since $\varrho : n \mapsto T^n x_0$ is a map from $\mathbb{N}$ to Y with dense range, it induces an embedding from $C(Y)$ into $l^\infty(\mathbb{N})$ (denoted by ϱ again), i.e., for any $h \in C(Y)$, $\varrho(h)(n) = h(T^n x_0)$. Then $\varrho(C(Y))$ is a C*-subalgebra of $l^\infty(\mathbb{N})$. Denote the maximal ideal space of $\varrho(C(Y))$ by $\widetilde{Y}$. By Proposition 2.2, the map $T^n x_0 \mapsto \iota(n)$ can be extended to a homeomorphism from Y onto $\widetilde{Y}$ (denoted by ϱ again). Note that $\varrho(g) = f$. So $\mathcal{A}_f$, the anqie generated by f, is a *-subalgebra of $\varrho(C(Y))$.

Since each multiplicative state on $\varrho(C(Y))$ (an element in $\widetilde{Y}$) is also a multiplicative state on $\mathcal{A}_f$ (an element in X_f) and that every maximal ideal in $\mathcal{A}_f$ extends to a maximal ideal (may not be unique) in $\varrho(C(Y))$, the induced map π from $\widetilde{Y}$ onto X_f given by

$$\pi(\rho)(h) = \rho(h), \quad \rho \in \widetilde{Y},\, h \in \mathcal{A}_f.$$

is continuous and surjective. It is not hard to check that $\pi \circ \varrho : Y \to X_f$ is a factor map. Then (X_f, σ_A) is a factor of (Y, T). $\square$

4. Mean states and asymptotically periodic functions

In number theory, we are often more concerned about the estimates of the form $\frac{1}{x}\sum_{n\le x} f(n)$. For this purpose, we shall consider states on $l^\infty(\mathbb{N})$ given by certain limits of $\frac{1}{N}\sum_{n=0}^{N-1} f(n)$ along “ultrafilters”. Then the inner product of two functions f and g given by the states is exactly certain limits of sums like $\frac{1}{N}\sum_{n=0}^{N-1} f(n)\overline{g(n)}$.

Recall that $\beta\mathbb{N}$ is the maximal ideal space of $l^\infty(\mathbb{N})$. Elements in $\beta\mathbb{N}\setminus\mathbb{N}$ are called *free ultrafilters*. By Proposition 2.2, $\mathbb{N}$ is dense in $\beta\mathbb{N}$. Given a free ultrafilter ω, for any $f \in l^\infty(\mathbb{N})$, there is a subsequence $\{m_j\}_{j=1}^\infty$ of $\mathbb{N}$ (depending on f) such that $\omega(f) = \lim_{j\to\infty} f(m_j)$. We usually write $\omega(f) = \lim_{n\to\omega} f(n)$, called *the limit of f at ω*.

For a C*-subalgebra $\mathcal{A}$ of $l^\infty(\mathbb{N})$, the linear functional ρ is called a *state* on $\mathcal{A}$ if $\rho(1) = 1$ and $\rho(f) \ge 0$ for any $f \in \mathcal{A}$ with $f \ge 0$. We shall study the A-invariant states on anqies defined below.

Definition 4.1. Suppose that $\mathcal{A}$ is an A-invariant C*-subalgebra of $l^\infty(\mathbb{N})$, i.e., $Af \in \mathcal{A}$ for any $f \in \mathcal{A}$. A state ρ on $\mathcal{A}$ is called **A-invariant**, or “invariant” for short, if $\rho(Af) = \rho(f)$ for any $f \in \mathcal{A}$.

Invariant states may or may not be related to average values of functions. Here we give an example to explain this phenomena.

Example 4.2. Let $G = \cup_{n=1}^\infty\{n^2 - n, n^2 - n + 1, \ldots, n^2 - 1\}$ be a subset of $\mathbb{N}$, and $G_n = \{i \in G : 0 \le i \le n-1\}$. Define $F_n(f) = \frac{1}{|G_n|}\sum_{i\in G_n} f(i)$ for $f \in l^\infty(\mathbb{N})$. Then for each given f the function $n \mapsto F_n(f)$ gives rise to a function in $l^\infty(\mathbb{N})$. Choose $\omega \in \beta\mathbb{N} \setminus \mathbb{N}$, and define $F_\omega(f) = \lim_{n\to\omega} F_n(f)$. Then F_ω is an A-invariant state on $l^\infty(\mathbb{N})$. If χ_G is the characteristic function supported on G, then $F_\omega(\chi_G) = 1$. But the relative density of G in $\mathbb{N}$ is zero. Thus $F_\omega(f)$ does not depend on the average sum $\frac{1}{n}\sum_{i=0}^{n-1} f(i)$.

On the other hand, there are A-invariant states depending on average values of functions, which are called "mean states" in [13, Definition 5.3].

Definition 4.3. Suppose $\omega \in \beta\mathbb{N}\setminus\mathbb{N}$ is a given free ultrafilter. For any $n \in \mathbb{N}$ and any f in $l^\infty(\mathbb{N})$, we define $E_n(f) = \frac{1}{n}\sum_{j=0}^{n-1} f(j)$. Then, for each given f, the function $n \to E_n(f)$ gives rise to another function in $l^\infty(\mathbb{N})$. The limit of $E_n(f)$ at ω is denoted by $E_\omega(f)$. Then E_ω is an A-invariant state defined on $l^\infty(\mathbb{N})$ or called "a **mean state**" (or, "a **mean**" for short).

From now on, we shall use E to denote a given mean state on $l^\infty(\mathbb{N})$ (depending on a free ultrafilter). For a real-valued function $f \in l^\infty(\mathbb{N})$, we always have:

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f(j) \le E(f) \le \limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f(j).$$

Suppose $\mathcal{A}$ is a countably generated anqie of $\mathbb{N}$. For each $N \in \mathbb{N}$, define the state ρ_N on $\mathcal{A}$ by $\rho_N(f) = \frac{1}{N}\sum_{j=0}^{N-1} f(j)$. So $\{\rho_N\}_{N=1}^\infty$ is a sequence in $(\mathcal{A}^\sharp)_1$. By Proposition 2.1, $(\mathcal{A}^\sharp)_1$ is metrizable and compact, so there is a subsequence $\{\rho_{N_m}\}_{m=1}^\infty$ that converges to some $\rho \in (\mathcal{A}^\sharp)_1$. We call this ρ *the limit of ρ_{N_m} or the state given (uniquely) by the sequence $\{N_m\}_{m=1}^\infty$*. It is not hard to check that ρ is an A-invariant state, and for any free ultrafilter ω in the closure of $\{N_m : m = 1, 2, 3, \ldots\}$ in $\beta\mathbb{N}$, the restriction of E_ω on $\mathcal{A}$ is ρ.

Now we perform the GNS construction on $l^\infty(\mathbb{N})$ with respect to E. Define $\langle f, g\rangle_E = E(\bar{g}f)$, the semi-inner product on $l^\infty(\mathbb{N})$ and $\|f\|_E = (\langle f, f\rangle_E)^{\frac{1}{2}}$, the semi-norm on $l^\infty(\mathbb{N})$ (see [22, Proposition 4.3.1]). We use $\mathcal{K}$ to denote the subalgebra of $l^\infty(\mathbb{N})$ containing all f so that $E(|f|^2) = \langle f, f\rangle_E = 0$. Then $\mathcal{K}$ is a closed two-sided ideal in $l^\infty(\mathbb{N})$. Thus $\mathcal{B} := l^\infty(\mathbb{N})/\mathcal{K}$ is a C*-algebra, and $\langle\ ,\ \rangle_E$ induces an inner product on $\mathcal{B}$. For $f \in l^\infty(\mathbb{N})$, we may use $\tilde{f}$ (or simply f if there is no ambiguity) to denote the coset $f + \mathcal{K}$ in $\mathcal{B}$. When $\tilde{f}, \tilde{g} \in \mathcal{B}$, we still use

$$\langle \tilde{f}, \tilde{g}\rangle_E = E(f\overline{g}) = \lim_{n\to\omega} \frac{1}{n}\sum_{j=0}^{n-1} f(j)\overline{g}(j) \tag{13}$$

to denote the inner product on $\mathcal{B}$ and

$$\|\tilde{f}\|_E = (\langle \tilde{f}, \tilde{f}\rangle_E)^{\frac{1}{2}} = \Big(\lim_{n\to\omega} \frac{1}{n}\sum_{j=0}^{n-1} |f(j)|^2\Big)^{\frac{1}{2}} \tag{14}$$

for the (Hilbert space) vector norm on $\mathcal{B}$. The completion of $\mathcal{B}$ under this norm is denoted by $\mathcal{H}_E$.

Remark 4.4. Our later results will depend on E but not on a specific one. Therefore, our definitions or properties stated later are for any mean state E. For example, if f and g are *orthogonal*, $E(f\overline{g}) = 0$ holds for any mean state E. The orthogonality of arithmetic functions may be viewed as disjointness between two functions in number theory.

Next, we study some properties of (strongly) asymptotically periodic functions (Definitions 1.3 and 1.4). We start from the following generalized notion of periodicity that introduced in [13, Section 5]. An arithmetic function $f \in l^\infty(\mathbb{N})$ is said to be *essentially periodic* (or "e-periodic") if there is an integer $n_0 \geq 1$ such that $f = A^{n_0} f$ in $\mathcal{H}_E$. The smallest such n_0 (≥ 1) is called the e-period of f. From the definition of strongly asymptotically periodic functions, it is easy to see that e-periodic functions belong exactly to this class.

In the following, we use $e(x)$ to denote $e^{2\pi i x}$ for simplicity, and 1_S to denote the indicator of a predicate S, that is $1_S = 1$ when S is true and $1_S = 0$ when S is false. It is not hard to check that $e(\sqrt{n})$ is an e-periodic function of e-period 1. Note that arithmetic functions satisfying $f(n) = f(n+1)$ for all n must be constant ones. Thus e-periodic functions are far from periodic ones. In the following, we shall construct e-periodic functions with e-period k, for any $k \geq 1$.

Example 4.5. Let $\{m_j\}_{j=1}^\infty$ and $\{n_j\}_{j=1}^\infty$ be two sequences of positive integers with $\lim_{j\to\infty} m_j = \lim_{j\to\infty} n_j = \infty$. For any given q, choose $\alpha = \{0, 1, \dots, 1\}, \beta = \{1, 0, \dots, 0\}$ as two vectors of length q. We construct the function f (written as $\{f(n)\}_{n=1}^\infty$) successively:

$$\underbrace{\alpha\alpha\cdots\alpha}_{m_1}\underbrace{\beta\beta\cdots\beta}_{n_1}\underbrace{\alpha\alpha\cdots\alpha}_{m_2}\underbrace{\beta\beta\cdots\beta}_{n_2}\cdots.$$

Then f is an e-periodic function with e-period q, and so a strongly asymptotically periodic function.

The proof of the above fact is more involved. Here are some details.

By $\lim_{j\to\infty} m_j = \lim_{j\to\infty} n_j = \infty$, for any $\epsilon > 0$, there is an j_0 such that $qm_j, qn_j > \frac{1}{\epsilon}+1$ when $j > j_0$. Then the number of n between 1 and N satisfying $f(n+q) \neq f(n)$ is less than $2qj_0+qN\epsilon$. Moreover, $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} |f(n+q) - f(n)|^2 \leq \lim_{N\to\infty}(\frac{2qj_0}{N} + q\varepsilon) \leq q\varepsilon$. Since ϵ is arbitrarily small, it follows that $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} |f(n+q) - f(n)|^2 = 0$. It is easy to see that for any positive integer $l \leq q-1$, $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} |f(n+l) - f(n)|^2 \neq 0$. Hence f is an e-periodic function with e-period q.

In Example 4.5, if the two sequences $\{m_j\}_{j=1}^{\infty}$ and $\{n_j\}_{j=1}^{\infty}$ further satisfy $\lim_{j\to\infty} \frac{m_j}{n_j} = a \neq 0$, then we can show that f is not the weak limit of periodic functions. That is for any mean state E, there does not exist a sequence $\{f_n\}_{n=1}^{\infty}$ of periodic functions, such that the limit of $f - f_n$ is zero in $\mathcal{H}_E$.

Using a similar argument to the proof of Example 4.5, we have the following result.

Example 4.6. *Let $\{N_j\}_{j=0}^{\infty}$ be a sequence of natural numbers with $N_0 = 0$ and $\lim_{j\to\infty}(N_{j+1} - N_j) = \infty$. Let $q \geq 1$ and $a \geq 0$ be given integers. Define $f(n)$ to be $a_j 1_{n\equiv a (mod\ q)}$ when $N_j \leq n < N_{j+1}$ for $j = 0, 1, \ldots$, where $\{a_j\}_{j=0}^{\infty}$ is a sequence of complex numbers with $\sup_j |a_j| < \infty$. Then f is an e-periodic function with e-period q, and so a strongly asymptotically periodic function.*

By Definitions 1.1 and 1.2, e-periodic and strongly asymptotically periodic functions are asymptotically periodic. There are many asymptotically periodic functions that are far from e-periodic ones. For example, if θ is irrational then $f(n) = e(n\theta)$ is asymptotically periodic. But it is not the weak (or l^2-) limit of e-periodic functions. In fact, we have the following result.

Proposition 4.7. *Let θ be an irrational number and $f(n) = e(n\theta)$. Then f is orthogonal to all e-periodic functions, that is for any e-periodic function g, $E(f\overline{g}) = 0$ holds for all mean states E.*

Proof. Suppose that the e-period of g is k. Given a mean state E, by the A-invariance of E, $E(f\overline{g}) = \langle f, g\rangle_E = \langle A^{lk}f, A^{lk}g\rangle_E = \langle A^{lk}f, g\rangle_E$ for any $l \geq 1$. Thus $\langle f, g\rangle_E = \langle \frac{1}{m}\sum_{l=1}^{m} A^{lk}f, g\rangle_E$. For any $\epsilon > 0$, we can choose a sufficiently large integer m such that $|\frac{1}{m}\sum_{l=1}^{m} e((n+lk)\theta)| = |\frac{1}{m}\sum_{l=1}^{m} e(lk\theta)| < \epsilon$ for any $n \in \mathbb{N}$. Hence $\|\frac{1}{m}\sum_{l=1}^{m} A^{lk}f\|_E < \epsilon$. It follows from the Cauchy-Schwarz inequality that $|\langle f, g\rangle_E| < \epsilon\|g\|_{l^\infty}$. Letting $\epsilon \to 0$. Then $\langle f, g\rangle_E = 0$. $\square$

Next, we provide another example of strongly asymptotically periodic function.

Example 4.8. *Let $f \in l^\infty(\mathbb{N})$. Suppose that the closure of $\{(f(n), f(n+1), \ldots) : n \in \mathbb{N}\}$ in $\prod_{\mathbb{N}} \overline{f(\mathbb{N})}$, endowed with the product topology, is countable. Then f is a strongly asymptotically periodic function.*

In the following, we give some detailed argument for the above nontrivial fact. Denote $\mathcal{A}_f$ as the anqie generated by f and X_f as the maximal ideal space of $\mathcal{A}_f$. By Theorem 3.4, X_f is homeomorphic to the closure of $\{(f(n), f(n+1), \ldots) : n \in \mathbb{N}\}$ in $\prod_{\mathbb{N}} \overline{f(\mathbb{N})}$ and so is a countable space. Let E be a mean state on $l^\infty(\mathbb{N})$. Then the restriction E on $\mathcal{A}_f$ is an invariant state on $\mathcal{A}_f$. By Theorem 5.1, there is a σ_A-invariant probability measure ν on X_f such that $E(g) = \int_{X_f} g(x)\, d\nu$ for any $g \in \mathcal{A}_f$. Since X_f is a countable and compact metric space, ν must be an atomic measure. Assume that ν is supported at $x_1, x_2, \ldots$ in X_f. For each x_i, there are two nature numbers s_i and t_i such that $A^{-s_i}\{x_i\} \cap A^{-t_i}\{x_i\} \neq \emptyset$. Thus there is a k_i such that $A^{k_i} x_i = x_i$. Set $n_j = \prod_{i=1}^{j} k_i$. So for each $l \geq 1$,

$$\begin{aligned}
\|f - A^{ln_j} f\|_E^2 &= E(|f - A^{ln_j} f|^2) \\
&= \int_{X_f} |(f - f \circ A^{ln_j})(x)|^2 \, d\nu \\
&= \sum_{m=1}^{\infty} |f(x_m) - f \circ A^{ln_j}(x_m)|^2 \, \nu(\{x_m\}) \\
&= \sum_{m=j+1}^{\infty} |f(x_m) - f \circ A^{ln_j}(x_m)|^2 \, \nu(\{x_m\}).
\end{aligned}$$

Then

$$\|f - A^{ln_j} f\|_E^2 \leq 4\|f\|_{l^\infty}^2 \sum_{m=j+1}^{\infty} \nu(\{x_m\}) \to 0$$

as j goes to ∞. Hence f is a strongly asymptotically periodic function.

Proposition 4.9. *The Möbius function is disjoint from all strongly asymptotically periodic functions if and only if for any given integer $q \geq 1$,*

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| \sum_{l=1}^{h} \mu(n + ql) \right|^2 = o(h^2). \tag{15}$$

Proof. We first prove " $\Rightarrow$ " part. Given integers $q \geq 1$ and $a \geq 0$. Take

$a_j = e(\theta_j)$ such that

$$\sum_{N_j \le n < N_{j+1}} \mu(n) e(\theta_j) 1_{n \equiv a (mod\ q)} = \left| \sum_{\substack{N_j \le n < N_{j+1} \\ n \equiv a (mod\ q)}} \mu(n) \right|.$$

Then the Möbius disjointness of $f(n)$ for any $f(n)$ defined as in Example 4.6 is equivalent to

$$\lim_{m \to \infty} \frac{1}{N_m} \sum_{j=0}^{m-1} \left| \sum_{\substack{N_j \le n < N_{j+1} \\ n \equiv a (mod\ q)}} \mu(n) \right| = 0$$

for any sequence $\{N_j\}_{j=0}^{\infty}$ with $N_0 = 0$ and $\lim_{j \to \infty}(N_{j+1} - N_j) = \infty$. This is further equivalent to (see e.g., [19, Lemma 5.2])

$$\lim_{h \to \infty} \limsup_{N \to \infty} \frac{1}{Nh} \sum_{m=1}^{N} \left| \sum_{\substack{n=m \\ n \equiv a (mod\ q)}}^{m+h} \mu(n) \right| = 0. \tag{16}$$

It is not hard to check that for N large enough,

$$\sum_{n=1}^{N} \left| \sum_{l=1}^{h} \mu(n + ql) \right| = \frac{1}{q} \sum_{a=1}^{q} \sum_{m=1}^{N} \left| \sum_{\substack{n=m \\ n \equiv a (mod\ q)}}^{m+hq} \mu(n) \right| + O(N). \tag{17}$$

Note that q is given and by the trivial estimate,

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| \frac{1}{h} \sum_{l=1}^{h} \mu(n + ql) \right|^2 \le \limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| \frac{1}{h} \sum_{l=1}^{h} \mu(n + ql) \right|.$$

By equations (16) and (17), we obtain equation (15).

We then prove " $\Leftarrow$ " part. Let $f(n)$ be a strongly asymptotically periodic function. It suffices to show that for any mean state E, $\langle f, \mu \rangle_E = E(f\mu) = 0$. By the strongly asymptotical periodicity of f, for any $\epsilon > 0$, there is a positive integer n_0 such that

$$\|f - A^{ln_0} f\|_E < \epsilon \tag{18}$$

for any $l \in \mathbb{N}$. By equation (15), there is a sufficiently large l_0 such that

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| \frac{1}{l_0} \sum_{l=1}^{l_0} \mu(n + ln_0) \right|^2 < \epsilon.$$

This implies

$$\|\frac{1}{l_0}\sum_{l=1}^{l_0} A^{ln_0}\mu\|_E < \epsilon. \tag{19}$$

Note that for any $l \in \mathbb{N}$, $\langle f, \mu\rangle_E = \langle A^{ln_0} f, A^{ln_0}\mu\rangle_E = \langle A^{ln_0} f - f, A^{ln_0}\mu\rangle_E + \langle f, A^{ln_0}\mu\rangle_E$. Then

$$\langle f, \mu\rangle_E = \frac{1}{l_0}\sum_{l=1}^{l_0}\langle A^{ln_0} f - f, A^{ln_0}\mu\rangle_E + \langle f, \frac{1}{l_0}\sum_{l=1}^{l_0} A^{ln_0}\mu\rangle_E. \tag{20}$$

By the Cauchy-Schwarz inequality and equations (18), (19), (20), we conclude that $|\langle f, \mu\rangle_E| < \epsilon(\|f\|_{l^\infty} + 1)$ for any $\epsilon > 0$. Letting $\epsilon \to 0$, $\langle f, \mu\rangle_E = 0$. ☐

There are many arithmetic functions that are not asymptotically periodic, such as $f(n) = e(n^2\theta)$ with θ irrational. This follows from the fact $\|f - A^m f\|_E = \sqrt{2}$ for any $m \geq 1$ and any mean state E. Moreover, this function is orthogonal to all asymptotically periodic functions, which is claimed in [13, Theorem 5.9] without proof. Here we give the proof in the following theorem.

Theorem 4.10. *Let $f(n) = e(n^2\theta)$ with θ irrational. We have*

(i) For any $l \neq m$, $\langle A^l f, A^m f\rangle_E = 0$ for any mean state E.

(ii) The function f is orthogonal to all asymptotically periodic functions in $l^\infty(\mathbb{N})$.

Proof. (i) This result is true as

$$\langle A^l f, A^m f\rangle_E = e((l^2 - m^2)\theta) \cdot \lim_{N\to\infty}\frac{1}{N}\sum_{j=1}^{N} e((2l - 2m)j\theta) = 0.$$

(ii) Let g be an asymptotically periodic function. Assume on the contrary that $|\langle f, g\rangle_E| > \delta$ for some $\delta > 0$. By definition, there is a sequence $\{n_j\}_{j=1}^\infty$ such that $\lim_{j\to\infty}\|A^{n_j} g - g\|_E = 0$. So when j is large enough, $\|A^{n_j} g - g\|_E < \delta/2$. By the Cauchy-Schwarz inequality and the fact that $\|f\|_E = 1$, we have

$$\begin{aligned}
|\langle g, A^{n_j} f\rangle_E| &= |\langle g - A^{n_j} g, A^{n_j} f\rangle_E + \langle A^{n_j} g, A^{n_j} f\rangle_E| \\
&\geq |\langle A^{n_j} g, A^{n_j} f\rangle_E| - |\langle g - A^{n_j} g, A^{n_j} f\rangle_E| \\
&> \delta/2.
\end{aligned}$$

It follows from (i) that the set $\{A^{n_j} f : j = 1, 2, \ldots\}$ is an orthogonal set in $\mathcal{H}_E$. By Bessel's inequality, $\|g\|_E^2 \geq \sum_{j=1}^\infty |\langle g, A^{n_j} f\rangle_E|^2 = \infty$. This contradicts the fact that $g \in l^\infty(\mathbb{N})$. Hence $\langle f, g\rangle_E = 0$ for any mean state E. ☐

Remark 4.11. Based on the above proof, an arithmetic function is orthogonal to all asymptotically periodic functions if it satisfies condition (i) of Theorem 4.10.

Theorem 4.12. *Let $r \geq 2$ and $\mu_r(n) = 1$ if n is r-th power-free and zero otherwise. For any $s \geq 1$ and $m_1, \ldots, m_s \in \mathbb{N}$, $\prod_{i=1}^{s} A^{m_i}(\mu_r)$ is strongly asymptotically periodic.*

Proof. It suffices to prove that μ_r is strongly asymptotically periodic. Let p_j be the j-th prime, define the sequence $\{n_j\}_{j=1}^{\infty}$ by $n_j = p_1^r p_2^r \cdots p_j^r$. By [30], for any positive integer m, we have for any mean state E,

$$\langle \mu_r, A^m \mu_r \rangle_E = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu_r(n)\mu_r(n+m) = \prod_p (1 - \frac{2}{p^r}) \prod_{p^r|m} (1 + \frac{1}{p^r - 2}).$$

So for any positive integer l, $\langle \mu_r, A^{l n_j} \mu_r \rangle_E \geq \langle \mu_r, A^{n_j} \mu_r \rangle_E$. Moreover,

$$\begin{aligned}
\|\mu_r - A^{l n_j} \mu_r\|_E^2 \leq & \|\mu_r - A^{n_j}\mu_r\|_E^2 = 2\langle \mu_r, \mu_r\rangle_E - 2\langle \mu_r, A^{n_j}\mu_r\rangle_E \\
= & 2\sum_{n=1}^{\infty} \frac{\mu(n)}{n^r} - 2\prod_p (1 - \frac{1}{p^r}) \prod_{p>p_j} (1 + \frac{1}{p^r - 2})^{-1} \\
= & 2\prod_p (1 - \frac{1}{p^r})(1 - \prod_{p>p_j} (1 + \frac{1}{p^r - 2})^{-1})
\end{aligned}$$

tends to 0 when j goes to infinity. Then $\{\mu_r - A^{l n_j}\mu_r\}_{j=1}^{\infty}$ converges to zero in $\mathcal{H}_E$ uniformly with respect to all $l \in \mathbb{N}$. $\square$

Now we prove Proposition 1.3.

Proof of Proposition 1.3. We know that $f = A^{m_1}(\mu^2) \cdots A^{m_s}(\mu^2)$ is strongly asymptotically periodic by Theorem 4.12. Let (X_f, σ_A) be the anqie generated by f. Then by Theorem 3.4, we describe X_f as a closed subspace $\widetilde{X_f}$ of $\{0,1\}^{\mathbb{N}}$ and represent σ_f as the Bernoulli shift B on the space. It follows from [35] that $(\widetilde{X_f}, B)$ has positive topological entropy and then (X_f, σ_A) has positive entropy. Assume on the contrary that there is a dynamical system (X, T) with the topological entropy of T zero, such that $f(n) = F(T^n x_0)$ for some $F \in C(X)$ and $x_0 \in X$. By Proposition 3.8, the topological entropy of T is greater than or equal to that of σ_A. This contradicts the assumption that the topological entropy of T is zero. Hence $f(n)$ cannot be realized in any dynamical systems with zero topological entropy. $\square$

In the next section, we shall see that the anqie of $\mathbb{N}$ generated by any asymptotically periodic function is closely related to a rigid dynamical system.

5. σ_A-invariant measures

Suppose that (X, σ_A) (or $\mathcal{A}$) is an anqie of $\mathbb{N}$. It is a basic fact that a continuous map on X is always (Borel) measurable. Then σ_A is a measurable transformation. We call a (Borel) measure ν on X *σ_A-invariant* if for any Borel set F of X, $\nu(F) = \nu((\sigma_A)^{-1}F)$. In the following, we show that for any given invariant state on $\mathcal{A}$, there is an induced σ_A-invariant Borel probability measure on X.

Theorem 5.1. *Let (X, σ_A) (or $\mathcal{A}$) be an anqie of $\mathbb{N}$. Suppose ρ is an invariant state on $\mathcal{A}$. Then there is a unique σ_A-invariant Borel probability measure ν on X such that for any $g \in \mathcal{A}$,*

$$\rho(g) = \int_X g(x)\ d\nu, \tag{21}$$

where $g(x)$ is the image of g under the Gelfand transform (see equation (10)).

Proof. Since ρ is an invariant state on $\mathcal{A}$ and $\mathcal{A} \cong C(X)$, ρ can be viewed as a state on $C(X)$ satisfying $\rho(f \circ \sigma_A) = \rho(f)$ for any $f \in C(X)$. By the Riesz representation theorem, there is a unique Borel measure ν on X, such that for any $f \in C(X)$,

$$\rho(f) = \int_X f(x)\ d\nu. \tag{22}$$

Moreover, ν is regular and it has the property that for any compact subset $K \subset X$,

$$\nu(K) = \inf\{\rho(h) : h \in C(X), h|_K = 1\}. \tag{23}$$

In the following, we show that $\nu(\sigma_A^{-1}(F)) = \nu(F)$ for any Borel set F. First we prove that if $\nu(F) = 0$, then $\nu(\sigma_A^{-1}(F)) = 0$. In fact, by the regularity of ν, for any $\epsilon > 0$, there is a compact set $K \subset \sigma_A^{-1}(F)$ such that

$$\nu(\sigma_A^{-1}(F)) < \nu(K) + \epsilon. \tag{24}$$

Note that $\sigma_A(K) \subset F$. So $\nu(\sigma_A(K)) = 0$. By equation (23), there is an $h \in C(X)$ such that $h|_{\sigma_A(K)} = 1$ and $\rho(h) < \epsilon$. By equation (23) again, $\nu(K) \leq \rho(h \circ \sigma_A) = \rho(h) < \epsilon$. Then $\nu(\sigma_A^{-1}(F)) < 2\epsilon$ by equation (24). Since ϵ can be arbitrarily small, $\nu(\sigma_A^{-1}(F)) = 0$.

Now we assume that $\nu(F) \neq 0$. By Lusin's Theorem, there is a sequence $\{f_n\}_{n=1}^{\infty}$ in $C(X)$ and a Borel set G with $\nu(G) = 0$, such that $\|f_n\| \leq 1$ and

$\lim_{n\to\infty} f_n(x) = \chi_F(x)$ for any $x \in X \setminus G$, where χ_F is the characteristic function supported on F. Thus $\lim_{n\to\infty} f_n \circ \sigma_A(x) = \chi_F \circ \sigma_A(x)$ for $x \in X \setminus \sigma_A^{-1} G$. By the analysis in the above paragraph, $\nu(\sigma_A^{-1} G) = 0$. By equation (22) and the Lebesgue Dominated Convergence Theorem,

$$\lim_{n\to\infty} \rho(f_n) = \lim_{n\to\infty} \int_X f_n(x)\, d\nu = \nu(F)$$

and

$$\lim_{n\to\infty} \rho(f_n \circ \sigma_A) = \lim_{n\to\infty} \int_X f_n \circ \sigma_A(x)\, d\nu = \nu(\sigma_A^{-1}(F)).$$

Since $\rho(f_n) = \rho(f_n \circ A)$, we obtain $\nu(F) = \nu(\sigma_A^{-1}(F))$ as claimed.

Finally, it follows from $\rho(1) = 1$ that $\nu(X) = 1$. Thus ν is a σ_A-invariant Borel probability measure on X. □

We call the σ_A-invariant (Borel) probability measure ν given by equation (21) *the measure induced by* ρ. Suppose $\mathcal{A}$ is a countably generated anqie of $\mathbb{N}$, then by Proposition 2.1, $(\mathcal{A}^\sharp)_1$ is a compact metrizable space. Hence for any mean state E on $l^\infty(\mathbb{N})$, there is a sequence $\{N_m\}_{m=1}^\infty$ of positive integers such that for any $g \in \mathcal{A}$, $E(g) = \lim_{m\to\infty} \frac{1}{N_m} \sum_{n=0}^{N_m-1} g(n)$. By Theorem 5.1, there is a σ_A-invariant probability measure ν on X such that for any $g \in \mathcal{A}$,

$$\lim_{m\to\infty} \frac{1}{N_m} \sum_{n=0}^{N_m-1} g(n) = \int_X g(x) d\nu. \tag{25}$$

On the other hand, for any $x \in X$ define a Borel probability measure δ_x on X such that for any Borel set B in $\mathcal{B}$, $\delta_x(B) = 1$ if $x \in B$, and 0 otherwise. For each $N \geq 1$, define

$$\delta_{N,x} = \frac{1}{N} \sum_{n=0}^{N-1} \delta_{(\sigma_A)^n x}. \tag{26}$$

It is easy to check that $\delta_{N,x}$ is a Borel probability measure on X. Now fix $x = \iota(0)$, i.e., the multiplicative state on $\mathcal{A}$ given by $\iota(0) : f \mapsto f(0)$ for any $f \in \mathcal{A}$. Then

$$\int_X g(x) d\delta_{N_m,\iota(0)} = \frac{1}{N_m} \sum_{n=0}^{N_m-1} \int_X g(x) d\delta_{(\sigma_A)^n(\iota(0))} = \frac{1}{N_m} \sum_{n=0}^{N_m-1} g(n)$$

holds for any $g \in \mathcal{A}$, correspondingly $g(x) \in C(X)$. By equation (25),

$$\lim_{m\to\infty} \int_X g(x) d\delta_{N_m,\iota(0)} = \int_X g(x) d\nu.$$

We call ν the *(weak*) limit* of $\delta_{N_m,\iota(0)}$. In [9], $\iota(0)$ is also called the *quasi-generic for ν along $\{N_m\}_{m=1}^{\infty}$*.

Remark 5.2. Let $\mathcal{A}_f$ be the anqie generated by f and X_f the maximal ideal space of $\mathcal{A}_f$. By Theorem 3.4, X_f is the closure of $\{(f(n), f(n+1), \ldots) : n \in \mathbb{N}\}$ in $\prod_{\mathbb{N}} \overline{f(\mathbb{N})}$. Suppose that ρ is an invariant state on $\mathcal{A}_f$, and ν the measure induced by ρ. Naturally, ν can be extended to a probability measure (denote by $\widetilde{\nu}$) on $\prod_{\mathbb{N}} \overline{f(\mathbb{N})}$, which is defined by $\widetilde{\nu}(F) = \nu(F \cap X_f)$ for any Borel set F of $\prod_{\mathbb{N}} \overline{f(\mathbb{N})}$. It is easy to see that $\widetilde{\nu}$ is B-invariant, where B is the Bernoulli shift on $\prod_{\mathbb{N}} \overline{f(\mathbb{N})}$. In the following, we still use ν to denote $\widetilde{\nu}$ if it makes no ambiguity.

Next, we discuss the connection between asymptotically periodic functions and rigid dynamical systems.

Theorem 5.3. *Suppose that f is an asymptotically periodic function. Let (X_f, σ_A) (or, $\mathcal{A}_f$) be the anqie generated by f. Let E be a mean state and ν the measure induced by E on X_f. Then (X_f, ν, σ_A) is rigid.*

Proof. By the definition of asymptotically periodic function, there is a sequence of positive integers $\{n_j\}_{j=1}^{\infty}$ with $\lim_{j\to\infty} E(|A^{n_j} f - f|^2) = 0$. It is not hard to check that for any $g \in \mathcal{A}_f$, $\lim_{j\to\infty} E(|A^{n_j} g - g|^2) = 0$. Thus for any $g(x) \in C(X_f)$, by equation (21), we have

$$\lim_{j\to\infty} \int_{X_f} |g \circ (\sigma_A)^{n_j}(x) - g(x)|^2 d\nu = 0. \tag{27}$$

By Lusin's Theorem, for any Borel set F of X_f, there is a sequence $\{g_n\}_{n=1}^{\infty}$ of continuous functions on X_f with $\|g_n\| \leq 1$ such that $\lim_{n\to\infty} g_n(x) = \chi_F(x)$ for almost all x. Then by Lebesgue's Dominated Convergence Theorem, for any $\epsilon > 0$, there is a g_{n_0} such that $\int_{X_f} |\chi_F(x) - g_{n_0}(x)| d\nu < \epsilon/3$. By equation (27), there is a sufficiently large K such that $\int_{X_f} |g_{n_0} \circ (\sigma_A)^{n_j}(x) - g_{n_0}(x)|^2 d\nu < \epsilon/3$ when $j > K$. Thus

$$\begin{aligned}
\nu((\sigma_A)^{-n_j} F \triangle F) &= \int_{X_f} |\chi_F \circ (\sigma_A)^{n_j}(x) - \chi_F(x)| d\nu \\
&\leq \int_{X_f} |\chi_F \circ (\sigma_A)^{n_j}(x) - g_{n_0} \circ (\sigma_A)^{n_j}(x)| d\nu \\
&\quad + \int_{X_f} |g_{n_0} \circ (\sigma_A)^{n_j}(x) - g_{n_0}(x)| d\nu \\
&\quad + \int_{X_f} |g_{n_0}(x) - \chi_F(x)| d\nu < \epsilon.
\end{aligned}$$

Hence $\lim_{j\to\infty} \nu((\sigma_A)^{-n_j}F\triangle F) = 0$. $\square$

It is known that for a measure-preserving dynamical system (X, ν, T) with T a homeomorphism, if ν has discrete spectrum, then T is rigid (see e.g., [34])). In the following, we show that this rigidity satisfies conditions (8), (9) in Theorem 1.11.

Proposition 5.4. *Let X be a compact metric space and $T : X \to X$ be a homeomorphism. Let ν be a T-invariant probability measure on X. Suppose that ν has discrete spectrum. Then (X, ν, T) satisfies conditions (8), (9) in Theorem 1.11. That is, for any $g(x) \in L^2(X, \nu)$, there are sequences $\{h_j\}_{j=1}^{\infty}$ and $\{n_j\}_{j=1}^{\infty}$ of positive integers with*

$$\lim_{j\to\infty} \frac{\log\log h_j}{\log h_j} \frac{n_j}{\varphi(n_j)} = 0 \tag{28}$$

such that

$$\lim_{j\to\infty} \frac{1}{h_j} \sum_{l=1}^{h_j} \|g \circ T^{ln_j} - g\|^2_{L^2(\nu)} = 0. \tag{29}$$

Proof. Since ν has discrete spectrum, by definition there is a standard orthogonal basis $\{g_s(x)\}_{s=1}^{\infty}$ in $L^2(X, \nu)$ with $g_s(Tx) = e^{2\pi i\lambda_s} g_s(x)$ for some real number λ_s, where $s = 1, 2, \ldots$. Let $g(x) \neq 0 \in L^2(X, \nu)$, write $g(x) = \sum_{s=1}^{\infty} a_s g_s(x)$. Then $\|g\|^2_{L^2(\nu)} = \sum_{s=1}^{\infty} |a_s|^2 < \infty$. For $j = 1, 2, \ldots$, choose $\epsilon_j = \|g\|^2_{L^2(\nu)}/j$ and $N_j \geq 1$ with $\sum_{s=N_j+1}^{\infty} |a_s|^2 < \frac{\epsilon_j}{8}$. Let $t_j = 2\|g\|^2_{L^2(\nu)} e^{N_j}/\epsilon_j$. Choose n_j such that $|e^{2\pi i n_j \lambda_s} - 1| \leq 1/t_j$ for $s = 1, \ldots, N_j$, where $1 \leq n_j \leq t_j^{N_j}$. Let $h_j = t_j^{1/2}$. By the choice of n_j and h_j, as well as the estimate $\frac{n_j}{\varphi(n_j)} \ll \log\log n_j$, it is not hard to check that they satisfy condition (28). Then

$$\begin{aligned}
&\frac{1}{h_j} \sum_{l=1}^{h_j} \|g \circ T^{ln_j} - g\|^2_{L^2(\nu)} = \|\sum_{s=1}^{\infty} a_s g_s \circ T^{ln_j}(x) - \sum_{s=1}^{\infty} a_s g_s(x)\|^2_{L^2(\nu)} \\
= \;& \frac{1}{h_j} \sum_{l=1}^{h_j} \sum_{s=1}^{\infty} |a_s|^2 |e^{2\pi i l n_j \lambda_s} - 1|^2 \\
= \;& \frac{1}{h_j} \sum_{l=1}^{h_j} \sum_{s=1}^{N_j} |a_s|^2 |e^{2\pi i l n_j \lambda_s} - 1|^2 + \frac{1}{h_j} \sum_{l=1}^{h_j} \sum_{s=N_j+1}^{\infty} |a_s|^2 |e^{2\pi i l n_j \lambda_s} - 1|^2 \\
\leq \;& \frac{1}{h_j} \sum_{l=1}^{h_j} N_j \|g\|^2_{L^2(\nu)} \frac{l^2}{t_j^2} + \frac{\epsilon_j}{2} \leq \frac{N_j \|g\|^2_{L^2(\nu)}}{t_j} + \frac{\epsilon_j}{2} < \epsilon_j \to 0, \qquad \text{as } j \to \infty.
\end{aligned}$$

$\square$

Proposition 5.5. *Let $f \in l^\infty(\mathbb{Z})$ be a weakly almost periodic function (see Remark 3.5 for definition). Then f_1 (=f restricted to $\mathbb{N}$) belongs to the class of asymptotically periodic functions described by conditions (4) and (5). That is, for any mean state E, there are sequences $\{h_j\}_{j=1}^\infty$ and $\{n_j\}_{j=1}^\infty$ of positive integers with*

$$\lim_{j\to\infty} \frac{\log\log h_j}{\log h_j}\frac{n_j}{\varphi(n_j)} = 0 \tag{30}$$

such that

$$\lim_{j\to\infty} \frac{1}{h_j}\sum_{l=1}^{h_j} E(|f - A^{ln_j} f|^2) = 0. \tag{31}$$

Proof. Let $\mathcal{A}_f$ be the C*-algebra of $l^\infty(\mathbb{Z})$ generated by 1 and $\{A^n f : n \geq 0\}$. We use X_f to denote the maximal ideal space of $\mathcal{A}_f$ and σ_A the homeomorphism on X_f induced by $n \mapsto n+1$ on $\mathbb{Z}$. Let E be a mean state on $l^\infty(\mathbb{N})$ depending on ω (see Definition 4.3), where ω is in the weak* closure of the sequence $\{N_m\}_{m=1}^\infty$ of positive integers in $\beta\mathbb{N}$. Then E can be naturally treated as a state on $\mathcal{A}_f$ in the way that

$$g \mapsto E(g_1)$$

for any $g(n) \in \mathcal{A}_f$, where $g_1(n)$ is the restriction of $g(n)$ to $\mathbb{N}$. Define the state ρ_{N_m} on $\mathcal{A}_f$ by $\rho_{N_m}(g) = \frac{1}{N_m}\sum_{n=0}^{N_m-1} g(n)$ for any $g \in \mathcal{A}_f$. So E is in the weak* closure of $\{\rho_{N_m}\}_{m=1}^\infty$ in $(\mathcal{A}_f^\sharp)_1$. Since $(\mathcal{A}_f^\sharp)_1$ is compact and metrizable, there is a subsequence $\{N_{m_s}\}_{s=1}^\infty$ of positive integers satisfying for any $g \in \mathcal{A}_f$,

$$E(g) = \lim_{s\to\infty} \frac{1}{N_{m_s}} \sum_{n=0}^{N_{m_s}-1} g(n). \tag{32}$$

By Theorem 5.1, there is a σ_A-invariant probability measure ν on X_f such that for any $g \in \mathcal{A}_f$,

$$\lim_{s\to\infty} \frac{1}{N_{m_s}} \sum_{n=0}^{N_{m_s}-1} g(n) = \int_{X_f} \widetilde{g}(x) d\nu, \tag{33}$$

where $\widetilde{g}(x)$ is the image of $g(n)$ under the Gelfand transform. Thanks to [20, Proposition 5.1], ν has discrete spectrum. By Proposition 5.4, there are sequences $\{n_j\}_{j=1}^{\infty}$ and $\{h_j\}_{j=1}^{\infty}$ of positive integers satisfying condition (30) such that

$$\lim_{j\to\infty}\frac{1}{h_j}\sum_{l=1}^{h_j}\|\widetilde{f}((\sigma_A)^{ln_j}x)-\widetilde{f}(x)\|_{L^2(\nu)}^2=0.$$

By equations (32) and (33),

$$\lim_{j\to\infty}\frac{1}{h_j}\sum_{l=1}^{h_j}E(|f-A^{ln_j}f|^2)=\lim_{j\to\infty}\frac{1}{h_j}\sum_{l=1}^{h_j}\int_{X_f}|\widetilde{f}((\sigma_A)^{ln_j}x)-\widetilde{f}(x)|^2d\nu=0.$$

□

To summarize, through invariant states we establish a connection between arithmetics and measurable dynamics. Specifically, for any given arithmetic function f in $l^{\infty}(\mathbb{N})$, it corresponds to a measure-preserving dynamical system (X_f,ν,σ_A). So we can apply tools in ergodic theory to study the system (X_f,ν,σ_A) and further study properties of f.

6. Proof of Theorem 1.4

Theorem 1.4 comes from a general result on the average of bounded multiplicative functions in short arithmetic progressions (see Proposition 6.1 below). In the statement of the next result, we shall use the following distance function of Granville and Soundararajan,

$$\mathbb{D}_k(f(n),g(n);x):=\Big(\sum_{\substack{p\leq x\\ p\nmid k}}\frac{1-\mathrm{Re}(f(p)\overline{g(p)})}{p}\Big)^{\frac{1}{2}}$$

for two multiplicative functions $f(n)$ and $g(n)$ with $|f(n)|,|g(n)|\leq 1$ for all $n\geq 1$. This distance function was used in [1] to measure the pretentiousness between any multiplicative function $f(n)$ and some function for which exceptional modulus k does exist. Throughout define

$$\begin{aligned}M_k(f;x;T):&=\inf_{|t|\leq T}\mathbb{D}_k(f,n\mapsto n^{it};x)^2,\\ M_k(f;k;x;T):&=\inf_{\chi(mod\ k)}\inf_{|t|\leq T}\mathbb{D}_k(f\chi,n\mapsto n^{it};x)^2.\end{aligned}$$

Proposition 6.1. *Let X be large enough with $1 \leq k \leq (\log X)^{1/32}$. Let $3 \leq h \leq X/k$. Let $f(n)$ be a multiplicative function with $|f(n)| \leq 1$ for all $n \geq 1$. Then*

$$
\begin{aligned}
&\sum_{\substack{a=1\\(a,k)=1}}^{k} \sum_{x=X}^{2X} \Bigg| \sum_{\substack{n=x\\ n\equiv a(mod\ k)}}^{x+hk} f(n) \Bigg|^2 \\
&\ll h^2 X\varphi(k)\Big(\frac{k}{\varphi(k)}\frac{\log\log h}{\log h} + \frac{1}{(\log X)^{1/300}} + \frac{M_k(f;k;X;2X)+1}{\exp(M_k(f;k;X;2X))}\Big).
\end{aligned} \tag{34}
$$

The major ingredient of our proof of the above result is the estimate [28] on averages of multiplicative functions in short intervals and the large sieve. We defer the proof to Appendix C. Here we list some recent results on averages of multiplicative functions in short arithmetic progressions: the method used in [29] can give that the coefficient before $\frac{\log\log h}{\log h}$ is k in formula (34); [23, Theorem 3.1] gave the result that when $f = \mu(n)$, then for any $\epsilon > 0$, the left hand side of formula (34)$\leq \epsilon h^2 X\varphi(k)$, whenever $\sum_{p|k} 1/p \leq (1-\epsilon)\sum_{p\leq h} 1/p$; For general multiplicative function, [24, Theorem 1.6, Corollary 1.7] gave the result that for any $\epsilon > 0$, the left hand side of formula (34)$\leq \epsilon h^2 X\varphi(k)$ when k is h^{ϵ^2}-typical (i.e., there are not many prime factors of k less than h^{ϵ^2}).

The reason that we give the estimate in form of formula (34) is that our main interest is to concern about when k is far larger than h, whether the first term of the right hand side of formula (34) is still $h^2X\varphi(k)o_h(1)$. According to our result, this is true if k is as large as $\exp(h^{o(1)})$ since $k/\varphi(k) \ll \log\log k$. We believe that the coefficient before $(\log\log h)/\log h (= o_h(1))$ in formula (34) can be as small as $O(1)$, an absolute constant independent of k. Actually, note that

$$
\frac{1}{X}\sum_{n=1}^{X}\Bigg|\sum_{l=1}^{h}\mu(n+lk)\Bigg|^2 = \frac{1}{Xk}\sum_{a=1}^{k}\sum_{x=1}^{X}\Bigg|\sum_{\substack{n=x\\ n\equiv a(mod\ k)}}^{x+hk}\mu(n)\Bigg|^2 + O(1).
$$

It is likely to believe that

$$
\limsup_{X\to\infty}\frac{1}{Xk}\sum_{a=1}^{k}\sum_{x=1}^{X}\Bigg|\sum_{\substack{n=x\\ n\equiv a(mod\ k)}}^{x+hk}\mu(n)\Bigg|^2 = o(h^2), \tag{35}
$$

where the little "o" term is independent of $k \geq 1$. This is implied by a positive answer to the Chowla conjecture. For a general non-pretentious multiplicative function $f(n)$ with $|f(n)| \leq 1$ for any $n \in \mathbb{N}$, equation (35) in which $\mu(n)$ is replaced by $f(n)$ would be implied by a positive answer to Elliott's conjecture (see [10, Conjecture II], [29]).

As an application of Proposition 6.1, we shall prove certain self correlations of the Möbius function which is stated in Theorem 1.4. In this proof, we need the following known result about the non-pretentious nature of $\mu(n)1_{(n,k)=1}$ (see e.g., [29, Lemma C.1]).

Lemma 6.2. *Let X be large enough with $k \le \log X$. Let $f(n) = \mu(n)1_{(n,k)=1}$. Then*

$$\inf_{1\le d\le k} M_k(f;d;X;2X) \ge (1/3-\epsilon)\log\log X + O(1),$$

where $\epsilon > 0$ is sufficiently small.

Proof of Theorem 1.4. Given $k \ge 1$ and $h \ge 2$. For X large enough with $\log X > h^2k$,

$$\sum_{n=X}^{2X}|\sum_{l=1}^{h}\mu(n+kl)|^2 = \sum_{a=1}^{k}\sum_{\substack{n=X\\ n\equiv a(mod\ k)}}^{2X}|\sum_{l=1}^{h}\mu(n+kl)|^2$$

$$= \sum_{a=1}^{k}\sum_{m=X/k}^{2X/k}|\sum_{l=1}^{h}\mu(km+kl+a)|^2 + O(h^2k)$$

$$= \sum_{a=1}^{k}\sum_{m=X/k}^{2X/k}|\sum_{\substack{n=(m+1)k\\ n\equiv a(mod\ k)}}^{(m+h+1)k}\mu(n)|^2 + O(h^2k)$$

$$= \sum_{a=1}^{k}\sum_{\substack{x=X\\ k|x}}^{2X}|\sum_{\substack{n=x\\ n\equiv a(mod\ k)}}^{x+hk}\mu(n)|^2 + O(h^2k)$$

$$= \frac{1}{k}\sum_{a=1}^{k}\sum_{x=X}^{2X}|\sum_{\substack{n=x\\ n\equiv a(mod\ k)}}^{x+hk}\mu(n)|^2 + O(X) \tag{36}$$

$$= \frac{1}{k}\sum_{d|k}\sum_{\substack{a=1\\ (a,k/d)=1}}^{k/d}\sum_{x=X}^{2X}|\sum_{\substack{n=x/d\\ n\equiv a(mod\ k/d)}}^{x/d+hk/d}\mu(dn)|^2 + O(X)$$

$$= \frac{1}{k}\sum_{d|k}d\sum_{\substack{a=1\\ (a,k/d)=1}}^{k/d}\sum_{x=X/d}^{2X/d}\Big|\sum_{\substack{n=x\\ n\equiv a(mod\ k/d)}}^{x+hk/d}\mu(n)1_{(n,k)=1}(n)\Big|^2$$

$$+ O(\frac{X}{k}\sum_{d|k}\varphi(k/d)) + O(X)$$

$$= \frac{1}{k}\sum_{d|k} d \sum_{\substack{a=1\\(a,k/d)=1}}^{k/d} \sum_{x=X/d}^{2X/d} \Big| \sum_{\substack{n=x\\ n\equiv a (mod\ k/d)}}^{x+hk/d} \mu(n)1_{(n,k)=1}(n)\Big|^2 + O(X).$$

Summarize the above, we have

$$\begin{aligned}\sum_{n=X}^{2X} |\sum_{l=1}^{h}\mu(n+kl)|^2 \\ =&\frac{1}{k}\sum_{d|k} d \sum_{\substack{a=1\\(a,k/d)=1}}^{k/d} \sum_{x=X/d}^{2X/d} \Big| \sum_{\substack{n=x\\ n\equiv a (mod\ k/d)}}^{x+hk/d} \mu(n)1_{(n,k)=1}(n)\Big|^2 + O(X).\end{aligned} \tag{37}$$

Then for X large enough, by Proposition 6.1 and Lemma 6.2,

$$\begin{aligned}&\sum_{n=X}^{2X} |\sum_{l=1}^{h}\mu(n+kl)|^2 \\ \ll&\frac{1}{k}(\sum_{d|k} d)h^2X/d\varphi(k/d)\Big(\frac{k/d}{\varphi(k/d)}\frac{\log\log h}{\log h} + \frac{1}{(\log X)^{1/400}}\Big) \\ =&h^2X(\sum_{d|k}\frac{1}{d})\frac{\log\log h}{\log h} + \frac{h^2X}{(\log X)^{1/400}} \\ \le&h^2X\prod_{p|k}(1-1/p)^{-1}\frac{\log\log h}{\log h} + \frac{h^2X}{(\log X)^{1/400}} \\ \le&h^2X\frac{k}{\varphi(k)}\frac{\log\log h}{\log h} + \frac{h^2X}{(\log X)^{1/400}}.\end{aligned}$$

Hence

$$\limsup_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}|\sum_{l=1}^{h}\mu(n+kl)|^2 \ll h^2\frac{k}{\varphi(k)}\frac{\log\log h}{\log h}, \tag{38}$$

as claimed. □

7. Proofs of Theorems 1.5, 1.8, and Proposition 1.7

As an application of Theorem 1.4, at the beginning of this section, we prove that the Möbius function is disjoint from certain asymptotically periodic functions (i.e., Theorem 1.5).

Proof of Theorem 1.5. Assume on the contrary, there is an $f \in l^\infty(\mathbb{N})$ with conditions (4) and (5) such that $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu(n)f(n) \neq 0$, then there is a constant $c_0 > 0$ and a mean state E such that

$$|\langle f, \mu\rangle_E| \geq c_0 \tag{39}$$

By conditions (4) and (5), there are correspondingly sequences $\{h_j\}_{j=0}^\infty$ and $\{n_j\}_{j=0}^\infty$ of positive integers with

$$\lim_{j\to\infty} \frac{\log\log h_j}{\log h_j}\frac{n_j}{\varphi(n_j)} = 0 \tag{40}$$

and

$$\lim_{j\to\infty} \frac{1}{h_j}\sum_{l=1}^{h_j} E(|f - A^{ln_j}f|^2) = 0. \tag{41}$$

Let $\delta = \frac{c_0}{2(\|f\|_{l^\infty}+1)}$. By Theorem 1.4, formulas (40) and (41), there is a k_0 such that

$$\frac{1}{h_{k_0}}\sum_{l=1}^{h_{k_0}} \|A^{ln_{k_0}}f - f\|_E^2 < \delta^2$$

and

$$\|\frac{1}{h_{k_0}}\sum_{l=1}^{h_{k_0}} A^{ln_{k_0}}\mu\|_E^2 < \delta^2.$$

For any $l \in \mathbb{N}$,

$$\langle f, \mu\rangle_E = \langle A^{ln_{k_0}}f, A^{ln_{k_0}}\mu\rangle_E = \langle A^{ln_{k_0}}f - f, A^{ln_{k_0}}\mu\rangle_E + \langle f, A^{ln_{k_0}}\mu\rangle_E.$$

Then

$$\begin{aligned}
|\langle f, \mu\rangle_E| =& \Big|\frac{1}{h_{k_0}}\sum_{l=1}^{h_{k_0}}\langle A^{ln_{k_0}}f - f, A^{ln_{k_0}}\mu\rangle_E + \langle f, \frac{1}{h_{k_0}}\sum_{l=1}^{h_{k_0}} A^{ln_{k_0}}\mu\rangle_E\Big| \\
\leq& \frac{1}{h_{k_0}}\sum_{l=1}^{h_{k_0}}\|A^{ln_{k_0}}f - f\|_E \cdot \|\mu\|_E + \|\frac{1}{h_{k_0}}\sum_{l=1}^{h_{k_0}} A^{ln_{k_0}}\mu\|_E \cdot \|f\|_E \\
\leq& \Big(\frac{1}{h_{k_0}}\sum_{l=1}^{h_{k_0}}\|A^{ln_{k_0}}f - f\|_E^2\Big)^{\frac{1}{2}} + \|\frac{1}{h_{k_0}}\sum_{l=1}^{h_{k_0}} A^{ln_{k_0}}\mu\|_E \cdot \|f\|_E \\
\leq& \delta(\|f\|_{l^\infty} + 1) = c_0/2.
\end{aligned}$$

Here we applied the Cauchy-Schwarz inequality to the first and second inequalities in the above, and the fact that $\|\mu\|_E \leq \|\mu\|_{l^\infty} = 1$ and $\|f\|_E \leq \|f\|_{l^\infty}$. This contradicts formula (39). Hence the claim in this theorem holds. □

Now we prove Proposition 1.7.

Proof of Proposition 1.7. Assume on the contrary that Problem 1 does not hold, that is, there is an asymptotically periodic function $f(n)$ such that $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu(n)f(n) \neq 0$. Then there is a $c_0 > 0$, a mean state E and a sequence $\{n_j\}_{j=1}^{\infty}$ of positive numbers such that

$$|\langle \mu, f\rangle_E| \geq c_0 \tag{42}$$

and

$$\lim_{j\to\infty} \|A^{n_j} f - f\|_E^2 = 0. \tag{43}$$

Let $\delta = \frac{c_0}{2(\|f\|_{l^\infty}+1)}$. By formula (6), choose a sufficiently large l_0 with

$$\|\frac{1}{l_0}\sum_{l=1}^{l_0} A^{lk}\mu\|_E < \delta, \tag{44}$$

for any $k \geq 1$. By equation (43), there is an n_0 such that

$$\|A^{n_0} f - f\|_E < \frac{2\delta}{l_0 + 1}.$$

Then by the triangle inequality,

$$\begin{aligned}\frac{1}{l_0}\sum_{l=1}^{l_0} \|f - A^{ln_0} f\|_E &\leq \frac{1}{l_0}\sum_{l=1}^{l_0}\sum_{j=1}^{l} \|A^{(j-1)n_0} f - A^{jn_0} f\|_E \\ &= \frac{1}{l_0}\sum_{l=1}^{l_0}\sum_{j=1}^{l} \|f - A^{n_0} f\|_E < \delta.\end{aligned}$$

By the A-invariance of E and the Cauchy-Schwarz inequality,

$$\langle f, \mu\rangle_E = \frac{1}{l_0}\sum_{l=1}^{l_0} \langle A^{ln_0} f - f, A^{ln_0}\mu\rangle_E + \langle f, \frac{1}{l_0}\sum_{l=1}^{l_0} A^{ln_0}\mu\rangle_E$$

$$\leq \frac{1}{l_0}\sum_{l=1}^{l_0}\|f - A^{ln_0}f\|_E \cdot \|\mu\|_{l^\infty} + \|\frac{1}{l_0}\sum_{l=1}^{l_0} A^{ln_0}\mu\|_E \cdot \|f\|_{l^\infty}$$
$$< \delta(1 + \|f\|_{l^\infty}) = c_0/2.$$

This contradicts formula (42). Hence formula (6) implies Problem 1. □

In the rest of this section, we shall prove Theorem 1.8, which states that if SMDC holds, then μ is disjoint from all asymptotically periodic functions. Before proving it, we need some preparations. We first provide a property of asymptotically periodic functions.

Proposition 7.1. *Let f be an asymptotically periodic function and ρ an invariant state on $\mathcal{A}_f$. Then for the measure-preserving dynamical system (X_f, ν, σ_A) with ν the probability measure induced by ρ on X_f, the measure-theoretic entropy of σ_A is zero.*

The above proposition follows immediately from Theorem 5.3 and [34, Example 5.3.3]. The basic connection between topological entropy (denoted by $h(T)$) and measure-theoretic entropy (denoted by $h_\nu(T)$) is the variational principle (see, e.g., [38, Theorem 8.6]). It states that for any topological dynamical system (X, T), $h(T) = \sup\{h_\nu(T) : \nu$ is a T-invariant Borel probability measure on $X\}$. By this principle, it is easy to see that if $h(T) = 0$, then $h_\nu(T) = 0$ for any T-invariant probability measure ν.

Here is an interesting example about topological entropy and measure-theoretic entropy. By Theorem 4.12, μ^2 is an asymptotically periodic function. So by Proposition 7.1, for any measure induced by a mean state ρ on X_{μ^2}, the measure-theoretic entropy of σ_A is zero. While Peckner proved in [32] that there is a σ_A-invariant measure on X_{μ^2} such that the measure-theoretic entropy of σ_A is equal to $\frac{6}{\pi^2}\log 2$, which equals the topological entropy of σ_A. So the measure-theoretic entropy varies with respect to different measures.

The following lemma is a consequence of Proposition 7.1 and [9, Lemmas 4.28, 4.29], which are used to prove the equivalence between SMDC and the Möbius disjointness of completely deterministic sequences.

Lemma 7.2. *Let f be an asymptotically periodic function and $\mathcal{A}_f$ be the anqie generated by f. Suppose $\{N_m\}_{m=1}^\infty$ is a strictly increasing sequence of positive integers such that $N_m|N_{m+1}$. Further suppose the sequence $\{N_m\}_{m=1}^\infty$ satisfies the condition that there is an A-invariant state ρ on $\mathcal{A}_f$, such that for any $h \in \mathcal{A}_f$, $\rho(h) = \lim_{m\to\infty} \frac{1}{N_m}\sum_{n=1}^{N_m} h(n)$. Then for any $\epsilon > 0$, there is an arithmetic function g with finite range, and a subsequence $\{N_{m(l)}\}_{l=1}^\infty$ such that* *(i) for (X_g, σ_A) the anqie generated by g, the topological entropy of σ_A is zero.*

(ii) $\frac{1}{N_{m(l)}}\sum_{n=1}^{N_{m(l)}}|f(n)-g(n)|<\epsilon$.

Based on such connections between asymptotically periodic functions and arithmetic functions with associated anqies having zero entropy, we are ready to prove Theorem 1.8.

Proof of Theorem 1.8. Assume on the contrary that there is some asymptotically periodic function f such that $\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu(n)f(n)\neq 0$, then there is a constant $c_0>0$ and an increasing sequence $\{N_m\}_{m=1}^{\infty}$ of positive integers with $N_m|N_{m+1}$ such that

$$\frac{1}{N_m}\left|\sum_{n=1}^{N_m}\mu(n)f(n)\right|\geq c_0. \tag{45}$$

For each N_m, define a state ρ_{N_m} on $\mathcal{A}_f$ by $\rho_{N_m}(h)=\frac{1}{N_m}\sum_{n=1}^{N_m}h(n)$ for any $h\in\mathcal{A}_f$. It follows from Proposition 2.1 that there is a subsequence $\{\rho_{N_{m(l)}}\}_{l=1}^{\infty}$ and a state ρ on $\mathcal{A}_f$, such that $\rho(h)=\lim_{l\to\infty}\frac{1}{N_{m(l)}}\sum_{n=1}^{N_{m(l)}}h(n)$ for any $h\in\mathcal{A}_f$. Then ρ is A-invariant. By Lemma 7.2, there is a $g(n)$ with the topological entropy of (X_g,σ_A) zero, and a subsequence of $\{N_{m(l)}\}_{l=1}^{\infty}$ (denoted by $\{N_{m(l)}\}_{l=1}^{\infty}$ again), such that

$$\frac{1}{N_{m(l)}}\sum_{n=1}^{N_{m(l)}}|f(n)-g(n)|<\frac{c_0}{2}. \tag{46}$$

Applying Sarnak's Möbius Disjointness Conjecture to (X_g,σ_A),

$$\lim_{l\to\infty}\frac{1}{N_{m(l)}}\sum_{n=1}^{N_{m(l)}}\mu(n)\widetilde{g}(A^n(\iota(0)))=\lim_{l\to\infty}\frac{1}{N_{m(l)}}\sum_{n=1}^{N_{m(l)}}\mu(n)g(n)=0, \tag{47}$$

where $\widetilde{g}(x)$ is the image of $g(n)$ in $C(X_g)$ under the Gelfand transform (see equation (10)). By equations (46) and (47), we obtain a result which contradicts formula (45). Then μ is disjoint from all asymptotically periodic functions. $\square$

8. Disjointness of Möbius from rigid dynamical systems

In this section, we shall prove Theorem 1.11, Corollary 8.2 and Proposition 1.9.

Proof of Theorem 1.11. Assume on the contrary, there is an $f \in C(X)$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(n) f(T^n x_0) \neq 0,$$

then there is a constant $c_0 > 0$ and an increasing sequence $\{N_m\}_{m=1}^{\infty}$ of positive integers such that

$$\frac{1}{N_m} \left| \sum_{n=1}^{N_m} \mu(n) f(T^n x_0) \right| \geq 2c_0. \tag{48}$$

Since X is a compact metric space, $C(X)$ is countably generated as an abelian C*-algebra. By Proposition 2.1, there is a subsequence of $\{N_m\}_{m=1}^{\infty}$ (denoted by $\{N_m\}_{m=1}^{\infty}$ again for convenience) and a T-invariant measure ν on X, such that $\nu_{N_m} = \frac{1}{N_m} \sum_{n=0}^{N_m-1} \delta_{T^n x_0}$ weak* converges to ν as $m \to \infty$, i.e., for any $f \in C(X)$,

$$\lim_{m\to\infty} \int_X f(x) d\nu_{N_m} = \lim_{m\to\infty} \frac{1}{N_m} \sum_{n=0}^{N_m-1} f(T^n x_0) = \int_X f(x) d\nu.$$

By formula (48) and the condition stated in this theorem, there is a $g \in C(X)$ and sequences $\{h_j\}_{j=1}^{\infty}$ and $\{n_j\}_{j=1}^{\infty}$ of positive integers with

$$\lim_{j\to\infty} \frac{\log\log h_j}{\log h_j} \frac{n_j}{\varphi(n_j)} = 0,$$

such that

$$\lim_{j\to\infty} \frac{1}{h_j} \sum_{l=0}^{h_j-1} \|g \circ T^{l n_j} - g\|_{L^2(\nu)}^2 = 0, \tag{49}$$

and

$$\frac{1}{N_m} \left| \sum_{n=1}^{N_m} \mu(n) g(T^n x_0) \right| \geq c_0. \tag{50}$$

Choose a free ultrafilter ω in the closure of $\{N_m : m = 1, 2, 3, \ldots\}$ in $\beta\mathbb{N}$. Then the mean state E on $l^\infty(\mathbb{N})$ defined by $E(h) = \lim_{N_m\to\omega} \frac{1}{N_m} \sum_{n=0}^{N_m-1} h(n)$ for any $h \in l^\infty(\mathbb{N})$ is A-invariant. Recall the GNS construction in Section 4, we use $\langle\ ,\ \rangle_E$ and $\|\cdot\|_E$ to denote the inner product and norm induced by E on

$\mathcal{H}_E$, respectively (see equations (13) and (14)). Let $\widetilde{g}(n) = g(T^n x_0)$. Then by equation (50), we have

$$|\langle \widetilde{g}, \mu \rangle_E| \geq c_0. \tag{51}$$

For any $l = 1, 2, \ldots$, note that

$$\|g \circ T^{ln_j} - g\|^2_{L^2(\nu)} = \lim_{m\to\infty} \frac{1}{N_m} \sum_{n=0}^{N_m-1} |g(T^{ln_j+n}x_0) - g(T^n x_0)|^2.$$

So by equation (49),

$$\lim_{j\to\infty} \frac{1}{h_j} \sum_{l=0}^{h_j-1} \|A^{ln_j}\widetilde{g} - \widetilde{g}\|^2_E = 0. \tag{52}$$

By an argument similar to the proof in Theorem 1.5, we have $|\langle \widetilde{g}, \mu \rangle_E| \leq c_0/2$. This contradicts formula (51). Hence we obtain

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(n) f(T^n x_0) = 0. \tag{53}$$

This completes the proof of the first part of this theorem.

In the rest, we show the second part of the claim in this theorem, which states the above disjointness holds over short intervals in average, that is

$$\lim_{h\to\infty} \limsup_{N\to\infty} \frac{1}{Nh} \sum_{n=1}^{N} \Big| \sum_{l=1}^{h} \mu(n+l) f(T^{n+l}x_0) \Big| = 0.$$

It is not hard to check that the above is equivalent to for any increasing sequence $\{N_j\}_{j=0}^{\infty}$ of natural numbers with $N_0 = 0$ and $\lim_{j\to\infty}(N_{j+1} - N_j) = \infty$,

$$\lim_{m\to\infty} \frac{1}{N_m} \sum_{j=0}^{m-1} \left| \sum_{N_j \leq n < N_{j+1}} \mu(n) f(T^n x_0) \right| = 0,$$

(see e.g., [19, Lemma 5.2]). Take $\{\theta_j\}_{j=0}^{\infty}$ such that

$$\sum_{N_j \leq n < N_{j+1}} \mu(n) f(T^n x_0) e(\theta_j) = \left| \sum_{N_j \leq n < N_{j+1}} \mu(n) f(T^n x_0) \right|.$$

Define $s(n) = e(\theta_j)$ when $N_j \le n < N_{j+1}$, $j = 0, 1, \ldots$. According to the above analysis, it suffices to prove that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu(n) f(T^n x_0) s(n) = 0. \tag{54}$$

Then $s(n)$ is an e-periodic function with e-period 1. Namely, for any mean state E and $l \in \mathbb{N}$,

$$E(|s(n+l) - s(n)|^2) = 0.$$

Let (X_s, σ_A) be the anqie generated by $s(n)$ and $\widetilde{s}(x)$ be the image of $s(n)$ in $C(X_s)$ under the Gelfand transform. Let $\mathcal{G}$ be the algebra generated by $\{1, \widetilde{s} \circ (\sigma_A)^n(x) : n = 0, 1, \ldots\}$. Then $\mathcal{G}$ is dense in $C(X_s)$. By Theorem 5.3, for any mean sate E, it induces a measure κ in the weak* closure of $\{\frac{1}{N}\sum_{n=0}^{N-1} \delta_{(\sigma_A)^n \iota(0)} : N = 1, 2, \ldots\}$ in the space of Borel probability measures on X_s satisfying

$$E(|s(n+l) - s(n)|^2) = \int_{X_s} |\widetilde{s} \circ (\sigma_A)^l(x) - \widetilde{s}(x)|^2 d\kappa = 0$$

for any $l \in \mathbb{N}$. By the above equation and the triangle inequality, it is not hard to check that conditions (8) and (9) in Theorem 1.11 hold for $(X \times X_s, T \times \sigma_A, (x_0, \iota(0))$ with $\mathcal{F} \times \mathcal{G}$ a dense set in $C(X \times X_s)$. By a similar argument to prove (53), we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu(n) f(T^n x_0) \widetilde{s}((\sigma_A)^n \iota(0)) = 0.$$

Note that $\widetilde{s}((\sigma_A)^n \iota(0)) = s(n)$. We obtain equation (54). Now we complete the proof of this theorem. ☐

Remark 8.1. Both BPV rigidity and PR rigidity in Theorem 1.10 are included in conditions (8), (9) in Theorem 1.11. Firstly, $\frac{n_j}{\varphi(n_j)} = \prod_{p|n_j} \frac{p}{p-1} = \prod_{p|n_j}(1 - \frac{1}{p})^{-1} \ll \exp(\sum_{p|n_j} \frac{1}{p}) = O(1)$ by the BPV rigidity, so (8) holds for any sequence $\{h_j\}_{j=1}^{\infty}$ with $\lim_{j\to\infty} h_j = \infty$. By BPV rigidity, there is a subsequence of $\{n_j\}_{j=1}^{\infty}$ (denoted by $\{n_j\}_{j=1}^{\infty}$ again for convenience) such that $\|g \circ T^{n_j} - g\|_{L^2(\nu)} \le \frac{1}{2^j}$. Choose $h_j = j$. Then by the triangle inequality and T-invariance of ν, $\|g \circ T^{l n_j} - g\|_{L^2(\nu)} \le l \|g \circ T^{n_j} - g\|_{L^2(\nu)}$. So

$$\frac{1}{h_j}\sum_{l=1}^{h_j} \|g \circ T^{l n_j} - g\|^2_{L^2(\nu)} \le \frac{j^2}{4^j} \to 0, \ as\ j \to \infty,$$

as claimed in formula (9). Secondly, we explain that PR rigidity is a special case of (8) and (9). Let $h_j = n_j^\delta$. Then $\lim_{j\to\infty} \frac{\log\log h_j}{\log h_j} \frac{n_j}{\varphi(n_j)} = 0$ since $\frac{n_j}{\varphi(n_j)} \ll \log\log n_j$.

Next, we give an example that satisfies conditions (8), (9), but not BPV rigidity and PR rigidity. Let $\eta = (\mu^2(0), \mu^2(1), \ldots)$ and B the Bernoulli shift on $\{0,1\}^{\mathbb{N}}$. Let X_η be the closure of $\{B^n\eta : n = 0, 1, \ldots\}$ in $\{0,1\}^{\mathbb{N}}$. We call (X_η, B) the *square-free flow*. The study of dynamical properties of the square-free flow have received much attention (see, e.g., [5, 32, 35]). In [35], Sarnak proved that (X_η, B) is proximal (i.e., for any $x, y \in X_\eta$, $\inf_{n\geq 1} d(T^n x, T^n y) = 0$) and it is topologically ergodic having topological entropy $\frac{6}{\pi^2}\log 2$. As a result of Theorem 1.11, we obtain the following Möbius disjointness for the square-free flow[2].

Corollary 8.2. *Let (X_η, B) be the square-free flow. Then for any $f \in C(X_\eta)$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(n) f(B^n\eta) = 0.$$

Proof. For $i = 0, 1, \ldots$, let $\pi_i : X_\eta \to \{0,1\}$ be the projection map from X_η onto its i-th coordinate. Let $\mathcal{F}$ be the *-subalgebra of $C(X_\eta)$ generated by $\{\pi_0, \pi_1, \ldots\}$. By the Stone-Weierstrass theorem (see, e.g., [22, Theorem 3.4.14]), $\mathcal{F}$ is dense in $C(X_\eta)$. By [35], there is a B-invariant measure ν such that $\frac{1}{N}\sum_{n=0}^{N-1} \delta_{B^n\eta}$ weak* converges to ν as $N \to \infty$. Let p_l be the l-th prime and $n_j = p_1^2 p_2^2 \cdots p_j^2$. By an argument similar to the proof in Theorem 4.12, for $i = 0, 1, \ldots$,

$$\begin{aligned}
\|\pi_i \circ B^{ln_j} - \pi_i\|_{L^2(\nu)}^2 &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\pi_i(B^{ln_j+n}\eta) - \pi_i(B^n\eta)|^2 \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mu^2(i + ln_j + n) - \mu^2(i+n)|^2 \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mu^2(ln_j + n) - \mu^2(n)|^2 \\
&\leq \frac{12}{\pi^2}(1 - \prod_{p>p_j} (1 + \frac{1}{p^2-2})^{-1}).
\end{aligned}$$

[2]There are some other methods to prove Corollary 8.2. Our primary interest here is to provide an example that distinguish Theorem 1.11 we obtained from [23, Theorem 2.1] (presented in Theorem 1.10 in this paper).

Then, for any increasing sequence $\{h_j\}_{j=1}^{\infty}$ of positive integers,

$$\lim_{j\to\infty}\frac{1}{h_j}\sum_{l=0}^{h_j-1}\|\pi_i\circ B^{ln_j}-\pi_i\|_{L^2(\nu)}^2\le\lim_{j\to\infty}\frac{1}{h_j}\sum_{l=0}^{h_j-1}\frac{12}{\pi^2}(1-\prod_{p>p_j}(1+\frac{1}{p^2-2})^{-1})=0.$$

It is not hard to check that for any $g\in\mathcal{F}$,

$$\lim_{j\to\infty}\frac{1}{h_j}\sum_{l=0}^{h_j-1}\|g\circ B^{ln_j}-g\|_{L^2(\nu)}^2=0.$$

Hence by Theorem 1.11, we obtain the claim in this corollary. □

Remark 8.3. In the following, we explain that for any π_i, $i=0,1,\ldots$, in the above dense set $\mathcal{F}$ of $C(X)$, there is no sequence $\{n_j\}_{j=1}^{\infty}$ satisfying BPV and PR rigidity in Theorem 1.10.

On one hand, by the argument in Corollary 8.2,

$$\|\pi_i\circ B^{n_j}-\pi_i\|_{L^2(\nu)}^2=\frac{12}{\pi^2}\Big(1-\prod_{p^2\nmid n_j}(1+\frac{1}{p^2-2})^{-1}\Big).$$

If $\lim_{j\to\infty}\|\pi_i\circ B^{n_j}-\pi_i\|_{L^2(\nu)}^2=0$, it is not hard to check that there is a subsequence $\{n_{j_s}\}_{s=1}^{\infty}$ with $p_1^2\cdots p_s^2|n_{j_s}$, where p_s is the s-th prime. Then $\sum_{p|n_{j_s}}\frac{1}{p}\ge\sum_{l\le s}\frac{1}{p_l}\to\infty$ as $s\to\infty$ by Mertens' Theorem (see e.g., [21]). So $\{n_j\}_{j=1}^{\infty}$ does not satisfy BPR rigidity in Theorem 1.10.

On the other hand, for a given $\delta>0$ and $(n_j)^{\frac{\delta}{2}}\le l\le h_j=n_j^{\delta}$ with j sufficiently large, note that the number of distinct prime factors of ln_j is $O_\delta(\log n_j)$, we have

$$\begin{aligned}
&\|\pi_i\circ B^{ln_j}-\pi_i\|_{L^2(\nu)}^2\\
=&\frac{12}{\pi^2}\Big(1-\prod_{p^2\nmid ln_j}(1+\frac{1}{p^2-2})^{-1}\Big)\\
\ge&\frac{12}{\pi^2}\Big(1-\prod_{p}(1+\frac{1}{p^2-2})^{-1}\prod_{p^2\le h_j}(1+\frac{1}{p^2-2})\prod_{\substack{p^2|ln_j\\ p^2>h_j}}(1+\frac{1}{p^2-2})\Big)\\
=&\frac{12}{\pi^2}\Big(1-\prod_{p^2>h_j}(1+\frac{1}{p^2-2})^{-1}(1+O_\delta(\frac{\log n_j}{h_j}))\Big)\gg\frac{1}{\sqrt{h_j}\log h_j}.
\end{aligned}$$

Hence, $\lim_{j\to\infty}\sum_{l=1}^{h_j}\|\pi_i\circ B^{ln_j}-\pi_i\|_{L^2(\nu)}^2\ne 0$. So the sequence $\{n_j\}_{j=1}^{\infty}$ does not satisfy PR rigidity in Theorem 1.10.

Remark 8.4. For the square-free flow (X_η, B), let ν be the B-invariant measure such that η is generic for ν. Then ν has discrete spectrum by [5]. From Corollary 8.2 and Remark 8.3, we know that (X_η, B, ν) satisfies conditions (8), (9) in Theorem 1.11, but not BPV rigidity and PR rigidity in Theorem 1.10.

By a similar argument to the proof of Corollary 8.2, the conclusion also holds for η replaced by $(\prod_{i=1}^{w}\mu_r(m_i), \prod_{i=1}^{w}\mu_r(m_i+1), \cdots, \prod_{i=1}^{w}\mu_r(m_i+n), \cdots)$, where $r \geq 2$, $w \geq 1$ and $m_1, \ldots, m_w \in \mathbb{N}$ are given, $\mu_r(n) = 1$ if n is r-th power-free and zero otherwise.

At the end, let us prove Proposition 1.9.

Proof of Proposition 1.9. We first show that Problem 1 implies Problem 2. Let $f \in C(X)$. Then for any ν in the weak* closure of $\{\frac{1}{N}\sum_{n=0}^{N-1}\delta_{T^n x_0} : N = 0, 1, 2, \ldots\}$ in the space of Borel probability measures on X, there is a sequence $\{n_j\}_{j=1}^{\infty}$ (may depend on ν) of positive integers satisfying

$$\lim_{j\to\infty} \|f \circ T^{n_j} - f\|_{L^2(\nu)}^2 = 0. \tag{55}$$

Let $g(n) = f(T^n x_0)$. In the following, we want to show that $g(n)$ is an asymptotically periodic function. Let $\mathcal{A}_g$ be the anqie generated by $g(n)$ and E be a mean state. Then there is a sequence $\{N_m\}_{m=1}^{\infty}$ of positive integers such that for any $h \in \mathcal{A}_g$, $E(h) = \lim_{m\to\infty}\frac{1}{N_m}\sum_{n=0}^{N_m-1}h(n)$. By Theorem 5.1, there is a probability measure ν_1 on X_f, such that

$$E(h) = \lim_{m\to\infty}\frac{1}{N_m}\sum_{n=0}^{N_m-1}h(n) = \int_{X_f} h(x)d\nu_1(x), \tag{56}$$

where $h(x)$ is the image of $h(n)$ under the Gelfand transform in $C(X_f)$. This implies that $\frac{1}{N_m}\sum_{n=0}^{N_m-1}\delta_{T^n x}$ weak* converges to ν_1 in the space of Borel probability measures on X_f. Choose a ν in the weak* closure of the sequence $\{\frac{1}{N_m}\sum_{n=0}^{N_m-1}\delta_{T^n x_0}\}_{m=1}^{\infty}$ in the space of Borel probability measures on X. When restricted to X_f, ν is identified as ν_1 by Proposition 3.8. Then by equation (55), there is a sequence $\{n_j\}_{j=1}^{\infty}$ of positive integers such that

$$\lim_{j\to\infty}\int_X |f \circ T^{n_j}(x) - f(x)|^2 d\nu(x) = 0.$$

Note that the image of $A^{n_j}g(n)$ under the Gelfand transform is $f \circ T^{n_j}(x)$ in $C(X_f)$. Then by equation (56),

$$\lim_{j\to\infty} E(|A^{n_j}g - g|^2) = 0.$$

So g is an asymptotically periodic function. Assume that Problem 1 holds, then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(n)g(n)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(n)f(T^n x_0)=0.$$

In the remaining part, we prove that Problem 2 implies the disjointness of μ from all asymptotically periodic function. Suppose that $h(n)$ is an asymptotically periodic function, i.e., for any mean state E, there is a sequence $\{n_j\}_{j=1}^{\infty}$ of positive integers such that $\lim_{j\to\infty}\|h-A^{n_j}h\|_E=0$. Let (X_h,σ_A) (or $\mathcal{A}_h$) be the anqie generated by h. Let $x_0=\iota(0)$ (corresponding to $(h(0),h(1),\ldots)$)$\in X_h$. Suppose that $\frac{1}{N_m}\sum_{n=0}^{N_m-1}\delta_{(\sigma_A)^n x_0}$ weak* converges to a Borel probability measure ν as $m\to\infty$. Choose a free ultrafilter ω in the weak* closure of $\{N_m: m=1,2,3,\ldots\}$ in $\beta\mathbb{N}$. Then applying Theorems 5.1 to the mean state E depending on ω, we obtain for any $\widetilde{f}(n)\in\mathcal{A}_h$,

$$E(\widetilde{f})=\lim_{m\to\infty}\frac{1}{N_m}\sum_{n=0}^{N_m-1}\widetilde{f}(n)=\int_{X_f}\widetilde{f}(x)d\nu(x),$$

where $\widetilde{f}(x)$ is the image of $\widetilde{f}(n)$ under the Gelfand transform (see equation (10)). Then for any $\widetilde{f}(x)\in C(X_h)$, $\lim_{j\to\infty}\|\widetilde{f}\circ(\sigma_A)^{n_j}(x)-\widetilde{f}(x)\|_{L^2(\nu)}=0$. So (X_h,σ_A,x_0) satisfies the condition in Problem 2. Hence

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(n)h((\sigma_A)^n x_0)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu(n)h(n)=0. \qquad \square$$

Appendix A. Mean and large values theorems

In this section, we list some lemmas that are used in the proof of Lemma C.1. They are hybrid versions of the corresponding results in [28]. We refers readers to [24, Section 3] or [31, Theorems 6.4; 8.3] for detailed proofs about Lemmas A.1, A.2, and A.4.

Lemma A.1. *Let $T,N,k\geq 1$ and $\{a_n\}_{n=1}^{\infty}$ be a sequence of complex numbers. Then*

$$\sum_{\chi(mod\ k)}\int_0^T|\sum_{n\leq N}a_n\chi(n)n^{it}|^2dt\ll(\varphi(k)T+\frac{\varphi(k)}{k}N)\sum_{\substack{n\leq N\\(n,k)=1}}|a_n|^2$$

Lemma A.2. *Let $T, N, k \geq 1$ and $\{a_n\}$ be any complex numbers. Let $\mathcal{E}$ be a subset of $\{\chi(\text{mod } k)\} \times [-T, T]$ satisfying that $|t-u| \geq 1$ whenever $(\chi,t), (\chi,u) \in \mathcal{E}$ with $t \neq u$. Then*

$$\sum_{(\chi,t)\in\mathcal{E}} |\sum_{n\leq N} a_n\chi(n)n^{it}|^2 \ll \Big(\varphi(k)T + \frac{\varphi(k)}{k}N\Big)\log(3k) \sum_{\substack{n\leq N \\ (n,k)=1}} |a_n|^2.$$

Applying the above lemma with an argument similar to the proof of [28, Lemma 8], we have the following.

Lemma A.3. *Let $P, T \geq 2$, $k \geq 1$ and $V > 0$. Write*

$$P_\chi(s) = \sum_{P\leq p\leq 2P} \frac{a_p\chi(p)}{p^s}$$

with $|a_p| \leq 1$ for $p \leq 2P$. Let $\mathcal{R}(\mathcal{T}, V)$ be a subset of $\{(\chi,t) \in \{\chi(\text{mod } k)\} \times [-T, T] : P_\chi(1+it) \geq V^{-1}\}$ satisfying $|t-u| \geq 1$ whenever $(\chi,t), (\chi,u) \in \mathcal{R}(\mathcal{T}, V)$ with $t \neq u$. Then

$$\#\mathcal{R}(\mathcal{T}, V) \ll (kT)^{2\frac{\log V}{\log P}} V^2 \exp\Big(2\frac{\log(kT)}{\log P}\log\log(kT)\Big).$$

The following is a hybrid version of "Halász inequality for integers" stated in [28, Lemma 9].

Lemma A.4. *With the same assumptions as Lemma A.2. We have*

$$\sum_{(\chi,t)\in\mathcal{E}} |\sum_{n\leq N} a_n\chi(n)n^{it}|^2 \ll \Big(\frac{\varphi(k)}{k}N + |\mathcal{E}|(kT)^{\frac{1}{2}}\log(2kT)\Big) \sum_{\substack{n\leq N \\ (n,k)=1}} |a_n|^2.$$

When a_n is supported on the set of primes, we have the following hybrid version of "Halász inequality for primes" stated in [28, Lemma 11].

Lemma A.5. *Let $P, T \geq 2$ and $k < (\log P)^{\frac{4}{3}-\epsilon}$. Let $\mathcal{E}$ be a subset of $\{\chi(\text{mod } k)\} \times [-T, T]$ satisfying that $|t-u| \geq 1$ whenever $(\chi,t), (\chi,u) \in \mathcal{E}$ with $t \neq u$. Then*

$$\sum_{(\chi,t)\in\mathcal{E}} |\sum_{P\leq p\leq 2P} a_p\chi(p)p^{it}|^2$$
$$\ll \Big(\varphi(k)P + |\mathcal{E}|P\exp(-\frac{\log P}{(\log(P+T))^{\frac{2}{3}+\epsilon}})(\log(P+T))^5\Big) \sum_{P\leq p\leq 2P} \frac{|a_p|^2}{\log P},$$

where ϵ is a sufficiently small positive number.

Proof. By the duality principle applied to $(\chi(p)p^{it})_{P\le p\le 2P,(\chi,t)\in\mathcal{E}}$, it is enough to prove that for any complex numbers $\eta_{\chi,t}$,

$$\sum_{P\le p\le 2P}\log p|\sum_{(\chi,t)\in\mathcal{E}}\eta_{\chi,t}\chi(p)p^{it}|^2$$
$$\ll\Big(|\mathcal{E}|P\exp(-\frac{\log P}{(\log(P+T))^{\frac{2}{3}+\epsilon}})(\log(P+T))^5\Big)+\varphi(k)P\Big)\sum_{(\chi,t)\in\mathcal{E}}|\eta_{\chi,t}|^2.$$

Let $f(x)$ be a smooth compactly supported function on $[1/2,5/2]$ such that $f(x)=1$ for $1\le x\le 2$ and f decays to zero outside of the interval $[1,2]$. Let $\widetilde{f}$ denote the Mellin transform of f. Then $\widetilde{f}(x+iy)\ll_A(1+|y|^{-2})$ uniformly in $|x|\le A$. Then

$$\sum_{P\le p\le 2P}\log p|\sum_{(\chi,t)\in\mathcal{E}}\eta_{\chi,t}\chi(p)p^{it}|^2\le\sum_{p^l}\log p|\sum_{(\chi,t)\in\mathcal{E}}\eta_{\chi,t}\chi(p^l)p^{ilt}|^2f(\frac{p^l}{P})$$
$$\le\sum_{(\chi,t),(\chi,t_1)\in\mathcal{E}}|\eta_{\chi,t}\eta_{\chi,t_1}|\sum_{p^l}(\log p)p^{il(t-t_1)}\chi_0(p^l)f(\frac{p^l}{P})|$$
$$+\sum_{\substack{(\chi,t),(\chi_1,t_1)\in\mathcal{E}\\ \chi_1\ne\chi}}|\eta_{\chi,t}\eta_{\chi_1,t_1}|\sum_{p^l}(\log p)p^{il(t-t_1)}\chi(p^l)\overline{\chi_1}(p^l)f(\frac{p^l}{P})|.$$

When χ is not a principal character modulo k, Perron's formula with the zero-free region for $L(s,\chi)$ gives for $|\alpha|\le T$,

$$\sum_{P<p<2P}p^{i\alpha}\chi(p)\ll P\exp(-\frac{\log P}{(\log(P+T))^{\frac{2}{3}+\epsilon}})(\log(P+T))^4.$$

Combining with $ab\le\frac{a^2+b^2}{2}$, we have

$$\sum_{\substack{(\chi,t),(\chi_1,t_1)\in\mathcal{E}\\ \chi_1\ne\chi}}|\eta_{\chi,t}\eta_{\chi_1,t_1}|\sum_{p^l}(\log p)p^{il(t-t_1)}\chi(p^l)\overline{\chi_1}(p^l)f(\frac{p^l}{P})|$$
$$\ll\sum_{\substack{(\chi,t),(\chi_1,t_1)\in\mathcal{E}\\ \chi_1\ne\chi}}(|\eta_{\chi,t}|^2+|\eta_{\chi_1,t_1}|^2)P\exp(-\frac{\log P}{(\log(P+T))^{\frac{2}{3}+\epsilon}})(\log(P+T))^5$$
$$\ll|\mathcal{E}|P\exp(-\frac{\log P}{(\log(P+T))^{\frac{2}{3}+\epsilon}})(\log(P+T))^5\sum_{(\chi,t)\in\mathcal{E}}|\eta_{\chi,t}|^2.$$

It follows from a similar argument to the proof of Lemma 11 in [28] that

$$\sum_{(\chi,t),(\chi,t_1)\in\mathcal{E}} |\eta_{\chi,t}\eta_{\chi,t_1}|\Big|\sum_{p^l}(\log p)p^{il(t-t_1)}\chi_0(p^l)f(\frac{p^l}{P})\Big|$$
$$\ll \sum_{(\chi,t),(\chi,t_1)\in\mathcal{E}} (|\eta_{\chi,t}|^2+|\eta_{\chi,t_1}|^2)\Big(\Big|\sum_{p^l}(\log p)p^{il(t-t_1)}f(\frac{p^l}{P})\Big|+\log P\log k\Big)$$
$$\ll (\varphi(k)P+|\mathcal{E}|P\exp(-\frac{\log P}{(\log T)^{\frac{2}{3}+\epsilon}})(\log T)^2+|\mathcal{E}|\log k\log P)\sum_{(\chi,t)\in\mathcal{E}}|\eta_{\chi,t}|^2.$$

□

The proofs of the next two lemmas are almost the same as the proofs of Lemmas 12, 13 in [28] with the following small differences: instead of the standard mean value theorem for Dirichlet polynomials, we apply Lemma A.1; one obtains the extra factor $\varphi(k)/k$ due to the coefficients are supported on the integers $(n,k)=1$.

Lemma A.6. *Let $X,H\geq 1$ and $Q>P\geq 2$. Suppose that $a_{mp}=b_mc_p$, $p\nmid m$, $P\leq p\leq Q$, where the sequences $\{a_m\}_m,\{b_m\}_m,\{c_p\}_p$ are bounded. Let $k\geq 1$ and $\mathcal{M}$ be a collection of Dirichlet characters modulo k. Let*

$$Q_{v,H}(\chi,s)=\sum_{\substack{P\leq p\leq Q\\ e^{\frac{v}{H}}\leq p\leq e^{\frac{v+1}{H}}}}\frac{c_p\chi(p)}{p^s}$$

and

$$R_{v,H}(\chi,s)=\sum_{Xe^{-\frac{v}{H}}\leq 2Xe^{-\frac{v}{H}}}\frac{b_m\chi(m)}{m^s}\frac{1}{\#\{P\leq q\leq Q: q|m, q\ is\ a\ prime\}+1}.$$

Let $\mathcal{T}_\chi\subseteq[-T,T]$, and $\mathcal{I}=\{j\in\mathbb{N}:\lfloor H\log P\rfloor\leq j\leq H\log Q\}$. Then

$$\sum_{\chi\in\mathcal{M}}\int_{\mathcal{T}_\chi}|\sum_{X\leq m\leq 2X}\frac{a_m\chi(m)}{m^{1+it}}|^2dt$$
$$\ll H\log(\frac{Q}{P})\times\sum_{\chi\in\mathcal{M}}\sum_{j\in\mathcal{I}}\int_{\mathcal{T}_\chi}|Q_{j,H}(\chi,1+it)R_{j,H}(\chi,1+it)|^2dt$$
$$+\frac{\varphi(k)}{k}\frac{\varphi(k)T+(\varphi(k)/k)X}{X}\Big(\frac{1}{H}+\frac{1}{P}\Big)$$

$$+\frac{\varphi(k)T+(\varphi(k)/k)X}{X}\sum_{\substack{X\le m\le 2X\\(m,k\prod_{P\le p\le Q}p)=1}}\frac{|a_m|^2}{m}.$$

Lemma A.7. *Let $k, T \geq 1$, $Y_2 \geq Y_1 \geq 2$ and $l = \lceil \frac{\log Y_2}{\log Y_1} \rceil$. Let $\{a_m\}_m$ and $\{c_p\}_p$ be bounded sequences. Suppose that X is sufficiently large. Let*

$$Q(\chi,s)=\sum_{Y_1\le p\le 2Y_1}\frac{c_p\chi(p)}{p^s}$$

and

$$R(\chi,s)=\sum_{X/Y_2\le m\le 2X/Y_2}\frac{a_m\chi(m)}{m^s}.$$

Then

$$\sum_{\chi(mod\ \ k)}\int_{-T}^{T}|Q(\chi,1+it)^l R(\chi,1+it)|^2dt \ll \frac{\varphi(k)}{k}(\varphi(k)\frac{T}{X}+\frac{\varphi(k)}{k}2^lY_1)(l+1)!^2.$$

The following Parseval bound follows exactly in the same way as [28, Lemma 14] with no need to consider the difference of two averages as the integral function.

Lemma A.8. *Suppose that $\{a_m\}_{m=1}^{\infty}$ be a bounded sequence. Assume that $X \geq 2$ and $1 \leq h \leq X$. Write*

$$A(s):=\sum_{X\le m\le 4X}\frac{a_m}{m^s}.$$

Then

$$\begin{aligned}(57)\qquad &\frac{1}{X}\int_X^{2X}|\frac{1}{h}\sum_{x\le n\le x+h}a_n|^2dx\\ &\ll\int_1^{1+iX/h}|A(s)|^2|ds|+\max_{T\ge X/h}\frac{X/h}{T}\int_{1+iT}^{1+2iT}|A(s)|^2|ds|.\end{aligned}$$

Appendix B. Lemmas on multiplicative functions

In this section, we give some lemmas on the pointwise bounds of Dirichlet polynomials with coefficients supported on integers coprime to a fixed number. We start from the following lemma which has almost identical proof to that of [1, Corollary 2.2] with the small modification: one applies the refinement of the Halász-Montgometry-Tenenbaum result ([16, Corollary 1]), rather than the Halász inequality. This leads to that the bound $O(\frac{1}{\sqrt{T}})$ is improved by $O(\frac{1}{T})$.

Lemma B.1. *Let $x \geq 3$, $1 \leq k \leq x$ and $1 \leq T \leq (\log x)^{\frac{1}{4}}$. Let $f(n)$ be a multiplicative function with $|f(n)| \leq 1$ for all $n \in \mathbb{N}$. Then*

$$\frac{1}{x}\sum_{\substack{n\leq x\\(n,k)=1}} f(n) \ll \frac{\varphi(k)}{k}\Big((M_k(f;x;T)+1)\exp(-M_k(f;x;T)) + \frac{1}{T}\Big).$$

While for large T in the above lemma, it follows directly from [24, Lemma 2.2] that

Lemma B.2. *Let $x \geq 3$, $1 \leq k \leq x$ and $(\log x)^{\frac{1}{4}} < T \leq x$. Let $f(n)$ be a multiplicative function with $|f(n)| \leq 1$ for all $n \in \mathbb{N}$. Then*

$$\frac{1}{x}\sum_{\substack{n\leq x\\(n,k)=1}} f(n) \ll \frac{\varphi(k)}{k}\Big((M_k(f;x;T)+1)\exp(-M_k(f;x;T)) + (\log x)^{-\frac{5}{64}}\Big).$$

Combining with Lemmas B.1 and B.2, we have the following Halász-type inequality for the mean values of multiplicative functions.

Lemma B.3. *Let $x \geq 3$ and $1 \leq k, T \leq x$. Let $f(n)$ be a multiplicative function with $|f(n)| \leq 1$ for all $n \in \mathbb{N}$. Then*

$$\frac{1}{x}\sum_{\substack{n\leq x\\(n,k)=1}} f(n) \ll \frac{\varphi(k)}{k}\Big((M_k(f;x;T)+1)\exp(-M_k(f;x;T)) + \frac{1}{T} + (\log x)^{-\frac{5}{64}}\Big).$$

The following lemma follows immediately from Lemma B.3 and partial summation.

Lemma B.4. *Let $x \geq 3$ and $1 \leq k, T_0 \leq x$. Suppose that χ is a Dirichlet character modulo k. Let $f(n)$ be a multiplicative function with $|f(n)| \leq 1$, and let*

$$F(\chi, s) = \sum_{x\leq n\leq 2x} \frac{f(n)\chi(n)}{n^s}.$$

Let

$$L(f\chi;x;T_0) = \inf_{|t_0|\leq T_0} \mathbb{D}_k(f\chi, n \mapsto n^{it+it_0};x)^2. \tag{58}$$

Then

$$|F(\chi, \sigma + it)|$$
$$\ll x^{1-\sigma} \frac{\varphi(k)}{k} \Big((L(f\chi; x; T_0) + 1) \exp(-L(f\chi; x; T_0)) + \frac{1}{T_0} + (\log x)^{-\frac{5}{64}} \Big).$$

Actually, in the proof of Theorem C.1, we also need to apply the Halász-type inequality to a Dirichlet polynomial of the form $F_{v,H}(\chi, s)$ in Lemma A.6 with the coefficients not quite multiplicative. Using Lemma B.4, a similar argument to the proof of Lemma 3 in [28] gives the following result.

Proposition B.5. *Let $X \geq Q > P \geq 2$. Let $1 \leq k, T_0 \leq X$ and χ be a Dirichlet character modulo k. Let $f(n)$ be a multiplicative function with $|f(n)| \leq 1$ and*

$$R(\chi, s) = \sum_{X \leq n \leq 2X} \frac{f(n)\chi(n)}{n^s} \frac{1}{\#\{P \leq q \leq Q : q|m, q \text{ is a prime}\} + 1}.$$

Suppose that $\delta(n)$ is the characteristic function supported on the set of all integers between 1 and $2X$ which is coprime to $\prod_{P \leq p \leq Q} p$. Then for any t,

$$|R(\chi, 1 + it)|$$
$$\ll \frac{\log Q}{\log P} \frac{\varphi(k)}{k} \Big((L(\delta f\chi; X; T_0) + 1) \exp(-L(\delta f\chi; X; T_0)) + \frac{1}{T_0} + (\log x)^{-\frac{5}{64}} \Big)$$
$$+ (\log X) \exp(-\frac{\log X}{3 \log Q} \log \frac{\log X}{\log Q}),$$

where $L(\delta f\chi; X; T_0)$ is defined as equation (58).

Appendix C. Proof of Proposition 6.1

In this section we shall first prove Proposition 6.1, which states that the average of a 1-bounded multiplicative function is small for almost all short arithmetic progressions when it it not $\chi(p)p^{it}$ pretentious. The proof of this result can be reduced to proving the following lemma.

Lemma C.1. *Let X be large enough such that $1 \leq k \leq (\log X)^{1/32}$. Suppose that $2 \leq h \leq X/k$. Let $f(n)$ be a multiplicative function with $|f(n)| \leq 1$ for all $n \geq 1$, and let*

$$F(\chi, s) = \sum_{X \leq n \leq 2X} \frac{f(n)\chi(n)}{n^s}.$$

Then, for any $T \geq 1$,

$$\begin{aligned}&\sum_{\chi (mod\ k)} \int_0^T |F(\chi, 1+it)|^2 dt \\ \ll &\frac{\varphi(k)}{k}\Big(\frac{\varphi(k)T}{X/h} + \frac{\varphi(k)}{k}\Big)\Big(\frac{k}{\varphi(k)}\frac{\log\log h}{\log h} + \frac{1}{(\log X)^{1/300}}\Big) \\ &+ \frac{\varphi^2(k)}{k^2}\Big(M_k(f;k;X;2X)+1\Big)\exp(-M_k(f;k;X;2X)\Big).\end{aligned}$$

Some results used below are given in Appendices A and B. We first show that the above lemma implies Proposition 6.1.

Proof of Proposition 6.1 (Assume that Lemma C.1 holds). By the Parseval bound stated in formula (57) and Lemma C.1,

$$\begin{aligned}&\frac{1}{k^2h^2X}\sum_{\chi(mod\ k)}\int_X^{2X}\Big|\sum_{n=x}^{x+hk} f(n)\chi(n)\Big|^2 dx \\ \ll &\sum_{\chi(mod\ k)}\int_1^{1+i\frac{X}{kh}}\Big|\sum_{X\leq m\leq 4X}\frac{f(m)\chi(m)}{m^s}\Big|^2|ds| \\ &+ \max_{T\geq \frac{X}{kh}}\frac{X/kh}{T}\sum_{\chi(mod\ k)}\int_{1+iT}^{1+2iT}\Big|\sum_{X\leq m\leq 4X}\frac{f(m)\chi(m)}{m^s}\Big|^2|ds| \\ \ll &\frac{\varphi^2(k)}{k^2}\Big(\frac{k}{\varphi(k)}\frac{\log\log h}{\log h} + \frac{1}{(\log X)^{1/300}}\Big) + \frac{\varphi^2(k)}{k^2}\frac{M_k(f;k;X;2X)+1}{\exp(M_k(f;k;X;2X))}.\end{aligned}$$

Hence

$$\begin{aligned}&\sum_{\substack{a=1\\(a,k)=1}}^{k}\sum_{x=X}^{2X}\Big|\sum_{\substack{n=x\\ n\equiv a(mod\ k)}}^{x+hk} f(n)\Big|^2 \\ =&\frac{1}{\varphi^2(k)}\sum_{\substack{a=1\\(a,k)=1}}^{k}\sum_{x=X}^{2X}\Big|\sum_{\chi(mod\ k)}\overline{\chi}(a)\sum_{n=x}^{x+hk} f(n)\chi(n)\Big|^2 = \frac{1}{\varphi^2(k)}\sum_{x=X}^{2X}\sum_{\chi_1,\chi_2(mod\ k)} \\ &\Big(\sum_{\substack{a=1\\(a,k)=1}}^{k}\overline{\chi_1}(a)\chi_2(a)\Big)\Big(\sum_{n=x}^{x+hk} f(n)\chi_1(n)\Big)\Big(\sum_{n=x}^{x+hk}\overline{f}(n)\overline{\chi_2}(n)\Big) \\ =&\frac{1}{\varphi(k)}\sum_{x=X}^{2X}\sum_{\chi(mod\ k)}\Big|\sum_{n=x}^{x+hk} f(n)\chi(n)\Big|^2\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{\varphi(k)}\sum_{\chi(mod\ k)}\int_X^{2X}\Big|\sum_{n=x}^{x+hk} f(n)\chi(n)\Big|^2 dx\\
\ll& h^2X\varphi(k)\Big(\frac{k}{\varphi(k)}\frac{\log\log h}{\log h}+\frac{1}{(\log X)^{1/300}}+\frac{M_k(f;k;X;2X)+1}{\exp(M_k(f;k;X;2X))}\Big). \qquad \square
\end{aligned}
$$

Now we start to prove Lemma C.1.

Proof of Lemma C.1. Since the hybrid mean value theorem (see Lemma A.1) gives the bound $O\Big(\frac{\varphi(k)}{k}(\frac{\varphi(k)T}{X}+\frac{\varphi(k)}{k})\Big)$, we can assume that $T \leq X$. Let χ_1 be the character modulo k minimizing the distance $\inf_{|t|\leq 2X}\mathbb{D}_k(f\chi, n\mapsto n^{it}; X)$. Let t_1 be the real number minimizing $\mathbb{D}_k(f\chi_1, n\mapsto n^{it}; X)$. Then for any $\chi(mod\ k)$ and $|t|\leq 2X$, $\mathbb{D}_k(f\chi, n\mapsto n^{it}; X)\geq \mathbb{D}_k(f\chi_1, n\mapsto n^{it_1}; X)$. Next we claim that for $\chi\neq\chi_1$ and any t with $|t|\leq 2X$,

$$
2\mathbb{D}_k(f\chi, n\mapsto n^{it}; X)\geq (\frac{1}{\sqrt{3}}-\epsilon)\sqrt{\log\log X}+O(1) \tag{59}
$$

and for $\chi=\chi_1$ and $|t-t_1|\geq 1$,

$$
2\mathbb{D}_k(f\chi_1, n\mapsto n^{it}; X)\geq (\frac{1}{\sqrt{3}}-\epsilon)\sqrt{\log\log X}+O(1), \tag{60}
$$

where $\epsilon>0$ is sufficiently small. In fact, suppose first that f is unimodular, i.e., $|f(n)|=1$ for all $n\geq 1$. By the triangle inequality of $\mathbb{D}_k$ (see, e.g., [17, Lemma 3.1]),

$$
\begin{aligned}
2\mathbb{D}_k(f\chi, n\mapsto n^{it}; X) &\geq \mathbb{D}_k(f\chi, n\mapsto n^{it}; X)+\mathbb{D}_k(f\chi_1, n\mapsto n^{it_1}; X)\\
&= \mathbb{D}_k(\overline{f}; n\mapsto \chi(n)n^{-it}; X)+\mathbb{D}_k(f, n\mapsto \overline{\chi_1}(n)n^{it_1}; X)\\
&\geq \mathbb{D}_k(\overline{f}f, n\mapsto \overline{\chi_1}\chi(n)n^{i(t_1-t)}; X)\\
&= \mathbb{D}_k(1, n\mapsto \overline{\chi_1}\chi(n)n^{i(t_1-t)}; X).
\end{aligned}
$$

If f is not unimodular, by means of the method used in [24, Lemma 2.2], we can model f by a stochastic multiplicative function $\mathbf{f}$ such that $\{\mathbf{f}(n)\}_n$ being a sequence of unimodular random variables defined on certain probability space, and for each prime p the expectation $\mathbb{E}\mathbf{f}(p)=f(p)$. By linearity of the expectation, we thus have

$$
\mathbb{D}_k(f\chi, n\mapsto n^{it}; X)^2=\sum_{\substack{p\leq x\\ p\nmid k}}\frac{1-\mathrm{Re}(p^{-it}\chi(p)\mathbb{E}\mathbf{f}(p))}{p}=\mathbb{E}\Big(\mathbb{D}_k(\mathbf{f}\chi, n\mapsto n^{it}; X)^2\Big).
$$

Since $\mathbf{f}$ is unimodular, $2\mathbb{D}_k(\mathbf{f}\chi, n \mapsto n^{it}; X) \geq \mathbb{D}_k(1, n \mapsto \overline{\chi_1}\chi(n)n^{i(t_1-t)}; X)$. Hence formulas (59) and (60) hold.

Write $[0,T] = \mathcal{L}_1 \cup \mathcal{L}_2$, where

$$\mathcal{L}_1 = \{0 \leq t \leq T : |t-t_1| < (\log X)^{\frac{5}{64}}\},$$

$$\mathcal{L}_2 = \{0 \leq t \leq T : |t-t_1| \geq (\log X)^{\frac{5}{64}}\}.$$

We now first estimate $\sum_{\chi (mod\ k)} \int_{\mathcal{L}_2} |F(\chi, 1+it)|^2 dt$. By means of similar ideas in the proof of [28, Proposition 1], we first split the integral over $\mathcal{L}_2$ into several parts according to the typical factorization when n is restricted to a dense subset $\mathcal{S} \subseteq [X, 2X]$. Recall that $\mathcal{S}$ in [28] is defined to be the set of all integers $X \leq n \leq 2X$ having at least one prime factor in each interval $[P_j, Q_j]$ for $j \leq J$, where J is chosen to be the largest index j such that $Q_j \leq \exp((\log X)^{\frac{1}{2}})$. The choice of P_j, Q_j needs to satisfy some requirements as in [28]. Now we set the same parameters $\alpha_j := \frac{1}{4} - \eta(1+\frac{1}{2j})$, $\eta := 1/150$, $H_j := j^2 P_1^{\frac{1}{6}-\eta}/(\log Q_1)^{\frac{1}{3}}$, $\mathcal{I}_j := [v \in \mathbb{N} : \lfloor H_j \log P_j \rfloor \leq v \leq H_j \log Q_j]$ as in [28]. Define for $v \in \mathcal{I}_j$,

$$R_{v,H_j}(\chi, 1+it) := \sum_{Xe^{-v/H_j} \leq m \leq 2Xe^{-v/H_j}} \frac{f(m)\chi(m)}{m^s} \frac{1}{\sharp\{P_j \leq p \leq Q_j : p|m\}+1}$$

and

$$Q_{v,H_j}(\chi, s) := \sum_{\substack{P_j \leq q \leq Q_j \\ e^{v/H_j} \leq q \leq e^{(v+1)/H_j}}} \frac{f(q)\chi(q)}{q^s}.$$

Let $\mathcal{T}_j$ denote the set of all $(\chi, t) \in \{\chi (mod\ k)\} \times \mathcal{L}_2$ with j the smallest index such that for all $v \in \mathcal{I}_j$, $|Q_{v,H_j}(\chi, 1+it)| \leq e^{-\alpha_j v/H_j}$. Let $\mathcal{U}$ be the complement of union of $\mathcal{T}_j$. We may also write that for some sets $\mathcal{T}_{j,\chi}, \mathcal{U}_\chi \subseteq \mathcal{L}_2$, $\mathcal{T}_j = \bigcup_{\chi (mod\ k)} \{\chi\} \times \mathcal{T}_{j,\chi}$ and $\mathcal{U} = \bigcup_{\chi (mod\ k)} \{\chi\} \times \mathcal{U}_\chi$. Then

$$\sum_{\chi (mod\ k)} \int_{\mathcal{L}_2} |F(\chi, 1+it)|^2 dt = \sum_{j=1}^{J} \sum_{\chi (mod\ k)} \int_{\mathcal{T}_{j,\chi}} |F(\chi, 1+it)|^2 dt + \sum_{\chi (mod\ k)} \int_{\mathcal{U}_\chi} |F(\chi, 1+it)|^2 dt. \tag{61}$$

By the the fundamental lemma of the sieve,

$$\sum_{\substack{X\le m\le 2X\\ (m,k\prod_{P_j\le p\le Q_j}p)=1}} 1 \ll X\frac{\varphi(k)}{k}\frac{\log P_j}{\log Q_j}\prod_{\substack{P_j\le p\le Q_j\\ p|k}}(1-\frac{1}{p})^{-1}\le X\frac{\varphi(k)}{k}\frac{\log P_j}{\log Q_j}\frac{k}{\varphi(k)}.$$

Using Lemma A.6 with $H=H_j, P=P_j, Q=Q_j$ and $a_m=b_m=f(m)\chi(m)$, $c_p=f(p)\chi(p)$ and the above inequality, we obtain

$$\begin{aligned}&\sum_{\chi(mod\ k)}\int_{\mathcal{T}_{j,\chi}}|F(\chi,1+it)|^2dt\\ &\ll H_j\log(\frac{Q_j}{P_j})\sum_{\chi(mod\ k)}\sum_{v\in\mathcal{I}_j}\int_{\mathcal{T}_{j,\chi}}|Q_{v,H}(\chi,1+it)R_{v,H}(\chi,1+it)|^2dt\\ &+\frac{\varphi(k)}{k}\frac{\varphi(k)T+(\varphi(k)/k)X}{X}\Big(\frac{1}{H_j}+\frac{1}{P_j}+\frac{k}{\varphi(k)}\frac{\log P_j}{\log Q_j}\Big).\end{aligned}$$

Here the second term contributes totally to the right-hand side of formula (61),

$$\begin{aligned}&\frac{\varphi(k)}{k}\frac{\varphi(k)T+(\varphi(k)/k)X}{X}\sum_{j=1}^{J}\Big(\frac{1}{H_j}+\frac{1}{P_j}+\frac{k}{\varphi(k)}\frac{\log P_j}{\log Q_j}\Big)\\ &\ll\frac{\varphi(k)}{k}\frac{\varphi(k)T+(\varphi(k)/k)X}{X}\Big(\frac{(\log Q_1)^{\frac13}}{P_1^{\frac16-\eta}}+\sum_{j=1}^{J}\frac{1}{P_1^{j^2}}+\frac{k}{\varphi(k)}\frac{\log P_1}{\log Q_1}\Big)\\ &\ll\frac{\varphi(k)}{k}\frac{\varphi(k)T+(\varphi(k)/k)X}{X}\Big(\frac{(\log Q_1)^{\frac13}}{P_1^{\frac16-\eta}}+\frac{k}{\varphi(k)}\frac{\log P_1}{\log Q_1}\Big).\end{aligned}\tag{62}$$

In the above, we use the relation that $H_j=j^2P_1^{\frac16-\eta}/(\log Q_1)^{\frac13}$ and $\log P_j\ge 8j^2/\eta\log Q_{j-1}+16j^2/\eta\log j$.

Now for $1\le j\le J$, we focus on bounding

$$E_j:=H_j\log Q_j\sum_{\chi(mod\ k)}\sum_{v\in\mathcal{I}_j}\int_{\mathcal{T}_{j,\chi}}|Q_{v,H_j}(\chi,1+it)R_{v,H_j}(\chi,1+it)|^2dt.$$

Estimate of E_1. We repeat the argument in [28, Section 8.1] with the difference that the standard mean-value theorem is replaced by the "hybrid mean-value theorem" (Lemma A.1),

$$E_1 \ll (\frac{\varphi(k)T}{X/Q_1} + \frac{\varphi(k)}{k})\frac{(\log Q_1)^{\frac{1}{3}}}{P_1^{\frac{1}{6}-\eta}}\frac{\varphi(k)}{k}. \tag{63}$$

Estimate of E_j for $2 \leq j \leq J$. Let $\mathcal{T}_{j,\chi}^r = \{t \in \mathcal{T}_{j,\chi}, |Q_{r,H_{j-1}}(\chi, 1+it)| > e^{-\frac{r\alpha_{j-1}}{H_{j-1}}}\}$ for $r \in \mathcal{I}_{j-1}$. Then $\mathcal{T}_{j,\chi} = \bigcup_{r\in\mathcal{I}_{j-1}} \mathcal{T}_{j,\chi}^r$. If $\mathcal{T}_{j,\chi}^r = \emptyset$, we set $\int_{T_{j,\chi}^r} |Q_{r,H_{j-1}}(\chi, 1+it)|^2 dt = 0$. Then

$$E_j \ll H_j \log Q_j \sum_{v\in\mathcal{I}_j} \sum_{r\in\mathcal{I}_{j-1}} \sum_{\chi (mod\ k)} e^{-2\frac{\alpha_j v}{H_j}} \int_{\mathcal{T}_{j,\chi}^r} |R_{v,H_j}(\chi, 1+it)|^2 dt.$$

By an argument similar to [28, Section 8.2] and Lemmas A.1, A.7, we obtain

$$E_j \ll \frac{\varphi(k)}{k}(\frac{\varphi(k)T}{X} + \frac{\varphi(k)}{k})\frac{1}{j^2 P_1}. \tag{64}$$

Estimate of $\sum_{\chi (mod\ k)} \int_{\mathcal{U}_\chi} |F(\chi, 1+it)|^2 dt$. Let $P = \exp((\log X)^{\frac{63}{64}}), Q = \exp(\frac{\log X}{\log\log X}), H = (\log X)^{\frac{1}{64}}$. Set $\mathcal{I} = [\lfloor H \log P\rfloor, H \log Q]$. For $v \in \mathcal{I}$, write

$$Q_{v,H}(\chi, s) = \sum_{\substack{P\leq p\leq Q \\ e^{v/H}\leq p\leq e^{(v+1)/H}}} \frac{f(p)\chi(p)}{p^s},$$

and

$$R_{v,H}(\chi, s) = \sum_{Xe^{-v/H}\leq n\leq 2Xe^{-v/H}} \frac{f(n)\chi(n)}{n^s} \frac{1}{\sharp\{p \in [P, Q] : p|n\} + 1}.$$

Note that $k < \log X$ and then $(k, \prod_{P\leq p\leq Q} p) = 1$. Applying Lemma A.6 with $a_m = b_m = f(m)\chi(m), c_p = f(p)\chi(p)$, we have that for some $v_0 \in \mathcal{I}$,

$$\begin{aligned}
&\sum_{\chi (mod\ k)} \int_{\mathcal{U}_\chi} \Big| \sum_{X\leq m\leq 2X} \frac{f(m)\chi(m)}{m^{1+it}} \Big|^2 dt \\
\ll& H^2 \log^2(\frac{Q}{P}) \sum_{\chi (mod\ k)} \int_{\mathcal{U}_\chi} |Q_{v_0,H}(\chi, 1+it) R_{v_0,H}(\chi, 1+it)|^2 dt \\
&+ \frac{\varphi(k)}{k} \frac{\varphi(k)T + (\varphi(k)/k)X}{X}\Big(\frac{1}{H} + \frac{1}{P}\Big) + \frac{\varphi(k)T + (\varphi(k)/k)X}{X} \frac{\log P}{\log Q} \frac{\varphi(k)}{k}.
\end{aligned}$$

Recall that $\mathcal{W} \subseteq [0, T]$ is called a set of well-spaced points if for any $t_1, t_2 \in \mathcal{W}$, we have $|t_1 - t_2| \geq 1$. There is a well-spaced set $\mathcal{L}_\chi \subseteq \mathcal{U}_\chi$ such that

$$\int_{\mathcal{U}_\chi} |Q_{v_0,H}(\chi, 1+it)R_{v_0,H}(\chi, 1+it)|^2 dt \ll \sum_{t \in \mathcal{L}_\chi} |Q_{v_0,H}(\chi, 1+it)R_{v_0,H}(\chi, 1+it)|^2.$$

Let

$$\mathcal{U}' = \bigcup_{\chi (mod\ k)} \{\chi\} \times \mathcal{L}_\chi.$$

Since $\log P_J - 1 \geq \frac{4j^2}{\eta} \log\log Q_{J+1} \geq \frac{2}{\eta} \log\log X$, $P_J > (\log X)^{\frac{2}{\eta}}$. By definition of $\mathcal{U}'$, for each $(\chi, t) \in \mathcal{U}'$, there is a $v \in \mathcal{I}_J$ such that $|Q_{v,H_J}(\chi, 1 + it)| > e^{-\alpha_J v/H_J}$. By Lemma A.3,

$$|\mathcal{U}'| \ll |\mathcal{I}_J|(kT)^{2\alpha_J + o(1)}(kT)^\eta X^{o(1)} \ll T^{\frac{1}{2}-\eta} X^{o(1)}.$$

We now also consider separately the cases

$$\mathcal{U}_S := \{(\chi, t) \in \mathcal{U}' : |Q_{v_0,H}(\chi, 1 + it)| < (\log X)^{-100}\},$$

$$\mathcal{U}_L := \{(\chi, t) \in \mathcal{U}' : |Q_{v_0,H}(\chi, 1 + it)| \geq (\log X)^{-100}\}.$$

For $\mathcal{U}_S$, applying Lemma A.4,

$$\begin{aligned} &\sum_{(\chi,t) \in \mathcal{U}_S} |Q_{v_0,H}(\chi, 1 + it)R_{v_0,H}(\chi, 1 + it)|^2 dt \\ \ll &\frac{1}{(\log X)^{200}} \sum_{(\chi,t) \in \mathcal{U}_S} |R_{v_0,H}(\chi, 1 + it)|^2 \\ \ll &\frac{1}{(\log X)^{200}} (Xe^{-v/H} + |\mathcal{U}_S|(kT)^{\frac{1}{2}})(\log 2kT) \frac{1}{Xe^{-v/H}} \ll \frac{1}{(\log X)^{199}}. \end{aligned}$$

Now it remains to estimate

$$\sum_{(\chi,t) \in \mathcal{U}_L} |Q_{v_0,H}(\chi, 1 + it)R_{v_0,H}(\chi, 1 + it)|^2.$$

By Lemma A.3, we obtain $|\mathcal{U}_L| \leq \exp((\log X)^{1/64+o(1)})$. We now give a pointwise bound to $R_{v_0,H}(\chi, 1 + it)$ for $(\chi, t) \in \mathcal{U}'$ as follows.

$$\max_{(\chi,t) \in \mathcal{U}_L} |R_{v_0,H}(\chi, 1 + it)| \ll \frac{\varphi(k)}{k} (\log X)^{-\frac{1}{16}+o(1)} \frac{\log Q}{\log P}. \tag{65}$$

We mainly use Proposition B.5 to prove the above inequality. Suppose $\delta(n) = 1_{(n,\prod_{P\le p\le Q} p)=1}(n)$. Then

$$\begin{aligned}\mathbb{D}_k(f\delta\chi, n\mapsto n^{it}; X)^2 =& \sum_{\substack{p\le x\\ p\nmid k}} \frac{1-\mathrm{Re}(p^{-it}\delta(p)\chi(p)f(p))}{p}\\ \ge& \mathbb{D}_k(f\chi, n\mapsto n^{it}; X)^2 - \sum_{p=P}^{Q}\frac{1}{p}\\ >& \mathbb{D}_k(f\chi, n\mapsto n^{it}; X)^2 - \frac{1}{64}\log\log X.\end{aligned} \tag{66}$$

By bounds (59) and (60), for $\chi \neq \chi_1$ and any t with $|t| \le 2X$,

$$\mathbb{D}_k(f\delta\chi, n\mapsto n^{it}; X)^2 > (1/16)\log\log X \tag{67}$$

and for $\chi = \chi_1$ and $|t - t_1| \ge 1$,

$$\mathbb{D}_k(f\delta\chi_1, n\mapsto n^{it}; X)^2 > (1/16)\log\log X. \tag{68}$$

Hence applying the above bounds and Proposition B.5 with $T_0 = \frac{1}{2}(\log X)^{\frac{5}{64}}$, we conclude

$$\max_{\substack{\chi(mod\ k)\\ \chi\neq\chi_1}} \max_{|t|\le X} |R_{v_0,H}(\chi, 1+it)| \ll \frac{\varphi(k)}{k}(\log X)^{-\frac{1}{16}+o(1)}\frac{\log Q}{\log P}. \tag{69}$$

and for $\chi = \chi_1$,

$$\max_{|t|\le X, |t-t_1|\ge(\log x)^{\frac{5}{64}}} |R_{v_0,H}(\chi_1, 1+it)| \ll \frac{\varphi(k)}{k}(\log X)^{-\frac{1}{16}+o(1)}\frac{\log Q}{\log P}. \tag{70}$$

Note that $\mathcal{U}_L \subseteq \mathcal{L}_2 = \{t \in [0,T] : |t - t_1| \ge (\log X)^{\frac{5}{64}}\}$. Hence we obtain formula (65). Based on the Halász bound (65) and the condition that $k \le (\log X)^{1/32}$, it follows, from the similar process in [28, Section 8.3] with the Halász inequality for primes replaced by a hybrid version of it (Lemma A.5), that

$$\begin{aligned}&\sum_{\chi(mod\ k)} \int_{\mathcal{U}_\chi} \Big|\sum_{X\le m\le 2X} \frac{f(m)\chi(m)}{m^{1+it}}\Big|^2 dt\\ &\ll \frac{\varphi(k)}{k}(\varphi(k)T/X + (\varphi(k)/k))(\log X)^{-\frac{1}{64}+o(1)}.\end{aligned} \tag{71}$$

Combining bounds (62), (63), (64), (71) with formula (61), we obtain

$$\begin{aligned}(72)\qquad &\sum_{\chi(mod\ k)} \int_{\mathcal{L}_2} |F(\chi, 1+it)|^2 dt \\ &\ll \frac{\varphi(k)}{k}\Big(\frac{\varphi(k)T}{X/Q_1} + \frac{\varphi(k)}{k}\Big)\Big(\frac{(\log Q_1)^{\frac{1}{3}}}{P_1^{\frac{1}{6}-\eta}} + \frac{k}{\varphi(k)}\frac{\log P_1}{\log Q_1} + \frac{1}{(\log X)^{\frac{1}{65}}}\Big).\end{aligned}$$

Thanks to equation (69), an argument similar to the proof of (72) leads to

$$\begin{aligned}(73)\qquad &\sum_{\substack{\chi(mod\ k)\\ \chi\neq\chi_1}} \int_{\mathcal{L}_1} |F(\chi, 1+it)|^2 dt \\ &\ll \frac{\varphi(k)}{k}\Big(\frac{\varphi(k)T}{X/Q_1} + \frac{\varphi(k)}{k}\Big)\Big(\frac{(\log Q_1)^{\frac{1}{3}}}{P_1^{\frac{1}{6}-\eta}} + \frac{k}{\varphi(k)}\frac{\log P_1}{\log Q_1} + \frac{1}{(\log X)^{\frac{1}{65}}}\Big).\end{aligned}$$

Now we are just left with estimating

$$\int_{\mathcal{L}_1} |F(\chi_1, 1+it)|^2 dt.$$

We first assume that

$$(74)\qquad (M_k(f\chi_1; X; 2X)+1)\exp(-M_k(f\chi_1; X; 2X)) > (\log X)^{-\frac{5}{64}}.$$

Now we write $\mathcal{L}_1 = \mathcal{L}_{0,1} \cup \mathcal{L}_{0,2}$ as a disjoint union, where

$$\mathcal{L}_{0,1} = \{t \in \mathcal{L}_1 : |t - t_1| < (M_k(f\chi_1; X; 2X)+1)^{-1}\exp(M_k(f\chi_1; X; 2X))\},$$

and

$$\begin{aligned}\mathcal{L}_{0,2} \;\; = \;\; &\{t \in \mathcal{L}_1 : (M_k(f\chi_1; X; 2X)+1)^{-1}\exp(M_k(f\chi_1; X; 2X)) \\ &\leq |t - t_1| \leq (\log X)^{\frac{5}{64}}\}.\end{aligned}$$

For $t \in \mathcal{L}_{0,1}$, by Lemma B.4 with $T_0 = (\log X)^{\frac{5}{64}}$, we have for $|t| \leq T \leq X$,

$$F(\chi_1, 1+it) \ll \frac{\varphi(k)}{k}(M_k(f\chi_1; X; 2X)+1)\exp(-M_k(f\chi_1; X; 2X)).$$

For $t \in \mathcal{L}_{0,2}$, by Lemma B.4 with $T_0 = \frac{|t-t_1|}{2}$, we have for $|t| \leq T \leq X$,

$$F(\chi_1, 1+it) \ll \frac{\varphi(k)}{k}\frac{1}{|t - t_1|},$$

this is because that $(M_k(f\chi_1;X;2X)+1)^{-1}\exp(M_k(f\chi_1;X;2X))\geq 1$ and $(L(f\chi_1;X;T_0)+1)\exp(-L(f\chi_1;X;T_0))\ll(\log X)^{-\frac{1}{12}+o(1)}$ by equation (60). Hence

$$\int_{\mathcal{L}_1}|F(\chi_1,1+it)|^2dt\ll\frac{\varphi^2(k)}{k^2}(M_k(f\chi_1;X;2X)+1)\exp(-M_k(f\chi_1;X;2X)). \tag{75}$$

Note that $M_k(f\chi_1;X;2X)=M_k(f;k;X;2X)$. Therefore, collecting equations (72), (73) and (75), we conclude that

$$\begin{aligned}&\sum_{\chi(mod\ k)}\int_0^T|F(\chi,1+it)|^2dt\\ &\ll\frac{\varphi(k)}{k}(\frac{\varphi(k)T}{X/Q_1}+\frac{\varphi(k)}{k})\Big(\frac{(\log Q_1)^{\frac{1}{3}}}{P_1^{\frac{1}{6}-\eta}}+\frac{k}{\varphi(k)}\frac{\log P_1}{\log Q_1}+\frac{1}{(\log X)^{\frac{1}{65}}}\Big)\\ &\quad+\frac{\varphi^2(k)}{k^2}(M_k(f;k;X;2X)+1)\exp(-M_k(f;k;X;2X)).\end{aligned} \tag{76}$$

If condition (74) does not hold, then

$$M_k(f\chi_1;X;2X)\geq(5/64-o(1))\log\log X.$$

So equation (68) holds for any $|t|\leq 2X$ by equation (66). Further using (67) and Proposition B.5 with $T_0=(\log X)^{\frac{1}{16}}$,

$$\max_{\chi(mod\ k)}\max_{|t|\leq X}|R_{v_0,H}(\chi,1+it)|\ll\frac{\varphi(k)}{k}(\log X)^{-\frac{1}{16}+o(1)}\frac{\log Q}{\log P}.$$

By the above pointwise bound, an argument similar to the proof of equation (72) leads to

$$\begin{aligned}\sum_{\chi(mod\ q)}\int_0^T|F(\chi,1+it)|^2dt\ \ll\ &\frac{\varphi(k)}{k}(\frac{\varphi(k)T}{X/Q_1}+\frac{\varphi(k)}{k})\Big(\frac{(\log Q_1)^{\frac{1}{3}}}{P_1^{\frac{1}{6}-\eta}}\\ &+\frac{k}{\varphi(k)}\frac{\log P_1}{\log Q_1}+\frac{1}{(\log X)^{\frac{1}{65}}}\Big),\end{aligned}$$

which implies formula (76).

Note that $\eta=\frac{1}{150}$. In case $h\leq\exp((\log X)^{1/2})$, we choose $Q_1=h$ and $P_1=(\log h)^{\frac{40}{\eta}}$; in case $\exp((\log X)^{\frac{1}{2}})\leq h\leq X$, we choose $Q_1=\exp((\log X)^{\frac{1}{2}})$, $P_1=Q_1^{(1/4)(\log h)^{-1/100}}$. Hence from the formula (76), we obtain the inequality in the statement of this lemma. $\square$

Acknowledgements

I am indebted to my advisor Liming Ge for encouraging me in this research. I am very grateful to Peter Sarnak for insightful comments on the manuscript. I heartily thank Jinxin Xue for providing helpful comments and suggestions on the manuscript. I would like to thank Arthur Jaffe for his support, and Boqing Xue and Wei Yuan for valuable discussions. This research was supported in part by the University of New Hampshire, by the AMSS of the Chinese Academy of Sciences, and by Grant TRT 0159 from the Templeton Religion Trust, and by the fellowship of China Postdoctoral Science Foundation 2020M670273.

References

[1] A. Balog, A. Granville and K. Soundararajan, Multiplicative functions in arithmetic progressions. *Ann. Math. Qué.* (1) **37** (2013), 3–30. MR3117735

[2] J. Bourgain, On the correlation of the Moebius function with rank-one systems. *J. Anal. Math.* (1) **120** (2013), 105–130. MR3095150

[3] J. Bourgain, P. Sarnak and T. Ziegler, Disjointness of Moebius from horocycle flows, in *From Fourier Analysis and Number Theory to Radon Transforms and Geometry*, Developments in Mathematics, vol. 28 (Springer-Verlag, New York, 2013), 67–83. MR2986954

[4] E. Čech, On bicompact spaces. *Ann. of Math.* (4) **38** (1937), 823–844. MR1503374

[5] F. Cellarosi and Y. G. Sinai, Ergodic properties of square-free numbers. *J. Eur. Math. Soc.* (4) **15** (2013), 1343–1374. MR3055764

[6] S. Chowla, *The Riemann Hypothesis and Hilbert's tenth problem.* Mathematics and Its Applications, vol. 4 (Gordon and Breach Science Publishers, New York, NY, 1965). MR0177943

[7] J. H. Conway, *A Course in Functional Analysis.* Gratuate Texts in Mathematics, vol. 96 (Springer-Verlag, New York, 1990). MR1070713

[8] W. F. Eberlein, Abstract Ergodic theorems and weak almost periodic functions. *Trans. Amer. Math. Soc.* (1) **67**(1949), 217–240. MR0036455

[9] E. H. El Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk and T. de la Rue, The Chowla and the Sarnak conjectures from ergodic

theory point of view. *Discrete Contin. Dyn. Syst* (6) **37** (2017), 2899–2944. MR3622068

[10] P. D. T. A. Elliott, On the correlation of multiplicative functions. *Notas Soc. Mat. Chile* (1) **11** (1992), 1–11. MR1292619

[11] A. Fan and Y. Jiang, Oscillating sequences, MMA and MMLS flows and Sarnak's conjecture. *Ergodic Theory Dynam. Systems* (5) **38** (2018), 1709–1744. MR3819999

[12] N. Frantzikinakis and B. Host, The logarithmic Sarnak conjecture for ergodic weights. *Ann. of Math.* (3) **187** (2018), 869–931. MR3779960

[13] 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. MR3546682

[14] E. Glasner, *Ergodic Theory via Joinings.* Mathematical surveys and monographs, vol. 101 (American Mathematical Society, Providence, RI, 2002). MR1958753

[15] E. Glasner and M. Megrelishvili, Hereditarily non-sensitive dynamical systems and linear representations. *Colloq. Math.* (2) **104** (2006), 223–283. MR2197078

[16] A. Granville and K. Soundararajan, Decay of mean-values of multiplicative functions. *Canad. J. Math.* (6) **55** (2003), 1191–1230. MR2016245

[17] A. Granville and K. Soundararajan, Large character sums: pretentious characters and the Pólya-Vinogradov theorem. *J. Amer. Math. Soc.* (2) **20** (2007), 357–384. MR2276774

[18] B. Green and T. Tao, The Möbius function is strong orthogonal to nilsequences. *Ann. of Math.* (2) **175** (2012), 541–566. MR2877066

[19] W. Gu and F. Wei, Möbius disjointness for a class of exponential functions. *Quart. J. Math.* https://doi.org/10.1093/qmath/haac009, (2022).

[20] W. Huang, Z. Wang and X. Ye, Measure complexity and Möbius disjointness. *Adv. Math.* **347** (2019), 827–858. MR3920840

[21] H. Iwaniec and E. Kowalski, *Analytic Number Theory.* American Mathematical Society Colloquium Publications, vol. 53 (American Mathematical Society, Providence, RI, 2004). MR2061214

[22] R. Kadison and J. Ringrose, *Fundamentals of the theory of Operator Algebras, Volume I and II.* Graduate Studies in Mathematics vol. 15 and vol. 16 (Academic Press, Orlando, 1983 and 1986). MR1468229

[23] A. Kanigowski, M. Lemańczyk and M. Radziwiłł, Rigidity in dynamics and Möbius disjointness. *Fund. Math.* (3) **255** (2021), 309–336. MR4324828

[24] O. Klurman, A. Mangerel and J. Teräväinen, Multiplicative functions in short arithmetic progressions. Preprint (2019), arXiv:1909.12280v3.

[25] M. Lemańczyk and C. Mauduit, Ergodicity of a class of cocycles over irrational rotations. *J. London Math. Soc.* (1) **49** (1994), 124–132. MR1253017

[26] J. Li, P. Oprocha and G. Zhang, On dynamics of quasi-graph maps. Preprint (2018), arXiv:1809.05617v1.

[27] J. Liu and P. Sarnak, The Möbius function and distal flows. *Duke Math. J.* (7) **164** (2015), 1353–1399. MR3347317

[28] K. Matomäki and M. Radziwiłł, Multiplicative functions in short intervals. *Ann. of Math.* (3) **183** (2016), 1015–1056. MR3488742

[29] K. Matomäki, M. Radziwiłł and T. Tao, An average form of Chowla's conjecture. *Algebra Number Theory* (9) **9** (2015), 2167–2196. MR3435814

[30] L. Mirsky, On the frequency of pairs of square-free numbers with a given difference. *Bull. Amer. Math. Soc.* (10) **55** (1949), 936–939. MR0031507

[31] H. Montgomery, *Topics in multiplicative number theory.* Lecture notes in Mathematics, vol. 227 (Springer-Verlag, New York, 1971). MR0337847

[32] R. Peckner, Uniqueness of the measure of maximal entropy for the squarefree flow. *Isr. J. Math.* (1) **210** (2015), 335–357. MR3430278

[33] R. Peckner, Möbius disjointness for homogeneous dynamics. *Duke Math. J.* (14) **167** (2018), 2745–2792. MR3859364

[34] K. Petersen, *Ergodic Theory.* Cambridge Studies in Advanced Mathematics, vol. 2 (Cambridge University Press, Cambridge, 1983). MR0833286

[35] P. Sarnak, *Three Lectures on the Möbius Function, Randomness and Dynamics.* Lecture Notes. IAS (2010).

[36] W. Veech, A fixed point theorem-free approach to weak almost periodicity. *Trans. Amer. Math. Soc.* **177** (1973), 353–362. MR0343254

[37] W. Veech, Möbius orthogonality for generalized Morse-Kakutani flows. *Amer. J. Math.* (5) **139** (2017), 1157–1203. MR3702497

[38] P. Walters, *An Introduction to Ergodic Theory*. Graduate Texts in Mathematics, vol. 79 (Springer-Verlag, New York, 1982). MR0648108

[39] Z. Wang, Möbius disjointness for analytic skew products. *Invent. Math.* (1) **209** (2017), 175–196. MR3660308

[40] F. Wei, *Entropy of arithmetic functions and Sarnak's Möbius disjointness conjecture*. Ph.D. Thesis, The University of Chinese Academy of Sciences, 2016.

[41] F. Wei, Anqie entropy and arithmetic compactifications of natural numbers. *Banach J. Math. Anal.* (1) **16** (2022), no. 11. MR4356225

[42] F. Wei, On the correlation between Möbius and polynomial phases in short arithmetic progressions. Preprint (2021), arXiv:2101.10134.

[43] L. Xu, Möbius disjointness for topological models of ergodic measure-preserving systems with quasi-discrete spectrum. *J. Differential Equations* (7) **266** (2019), 3883–3905. MR3912703

Fei Wei
Yau Mathematical Sciences Center
Tsinghua University
Beijing 100084
China
E-mail: weif@mail.tsinghua.edu.cn