Sarnak's Conjecture on Möbius Disjointness

Eliot Hodges

Cambridge University Part III
Department of Mathematics
Joni Teräväinen
May 2026

1. Möbius Pseudorandomness

1.1. **The Möbius function.** First introduced in 1832 by August Ferdinand Möbius, the *Möbius function* $\mu\colon \mathbb{N} \to \{-1, 0, 1\}$, given by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1; \\ 0 & \text{if } n \text{ is divisible by a square;} \\ (-1)^k & \text{if } n \text{ is a product of } k \text{ distinct primes } n = p_1 \cdots p_k, \end{cases}$$

is of central importance in number theory. In particular, understanding the average behavior of μ is equivalent to understanding certain aspects of the prime numbers. Several important arithmetic theorems and conjectures can be translated into statements about the statistics of the Möbius function. August among these are the prime number theorem and the Riemann hypothesis, which we restate below in terms of μ.

Theorem 1.1 (Prime number theorem)**.** *We have that*

$$\sum_{n=1}^{N} \mu(n) = o(N).$$

Conjecture 1.2 (Riemann hypothesis)**.** *For every $\varepsilon > 0$, we have that*

$$\sum_{n=1}^{N} \mu(n) = O_\varepsilon(N^{1/2+\varepsilon}).$$

In the language of the Möbius function, we see immediately the sense in which the Riemann hypothesis can be regarded as a more precise, quantitative version of the prime number theorem. Proof that the statements in Theorem 1.1 and Conjecture 1.2 are indeed equivalent to their classical counterparts is elementary and can be found, for example, in [23, 28].

1.2. **Randomness heuristics.** Closer inspection of the values of $\mu(n)$ reveals seemingly random behavior. As is often the case in number theory, it is helpful to make predictions about the behavior of μ based on the heuristic assumption that the values of μ on squarefree integers are independent and identically distributed random variables assuming 1 and -1 with equal probability.[1] This analogy—called the *Möbius pseudorandomness heuristic*—is not purely philosophical; one of the goals of this paper is to disentangle the several ways in which this heuristic can be made precise.

For example, consider the sum of N independent and identically distributed random variables $X_1, \dots, X_N$ taking values in $\{\pm 1\}$ with uniform probability. The law of the iterated logarithm (see [19]) tells us that

$$\sum_{n=1}^{N} X_n = O(\sqrt{N \log\log N}) = O_\varepsilon(N^{1/2+\varepsilon}) \quad \text{almost surely.} \tag{1}$$

From (1), we see immediately the way in which the Riemann hypothesis supports the Möbius pseudorandomness principle. Despite widespread confidence in the verity of Conjecture 1.2, a proof seems beyond current techniques. Sarnak's conjecture on Möbius disjointness—the subject of this paper—offers a more tractable interpretation of Möbius pseudorandomness.

[1] Of course, as μ is a deterministic sequence, it is not actually random. See the discussion at the bottom of page 1 of [27] for more details.

The heuristic assumption underlying Sarnak's conjecture is that μ does not correlate with functions of low complexity, a notion which is explored in greater detail in Sarnak's lecture notes on the conjecture [27]. Ironing out precisely what is meant by "complexity" is one of the striking accomplishments of Sarnak's work. For any such bounded function ξ, this version of Möbius pseudorandomness can be formulated as

$$\sum_{n=1}^{N} \mu(n)\xi(n) = o(N). \tag{2}$$

That is, we expect that summing μ against any such ξ must lead to significant cancellations. In this language, the prime number theorem and Riemann hypothesis can be reinterpreted as statements about the "orthogonality" of μ and the constant function 1.

A naïve first guess for a viable notion of complexity is *computational complexity*, which comes from theoretical computer science. We say a function ξ has *low computational complexity* if it can be computed in polynomial time. In other words, if there exists some B depending on ξ such that the value $\xi(n)$ can be computed in $O((\log n)^B)$ steps. Immediately, we run into difficulties with this definition. Firstly, it is unknown whether μ itself is of low computational complexity (see [1] for a more detailed discussion). Secondly, consider the function ξ that is -1 on the primes and 0 otherwise. Then ξ is of low computational complexity [2], and we see that

$$\sum_{n=1}^{N} \mu(n)\xi(n) = \sum_{n=1}^{N} |\xi(n)| = \pi(N) \sim \frac{N}{\log N}$$

since $\mu(n)\xi(n) = \mathbb{1}_{n\in\mathbb{P}}$, where $\mathbb{P}$ denotes the set of primes. While the right-hand side of the equation above is $o(N)$, we see that there are no cancellations over the primes: the only reason that μ is disjoint from ξ in this case is that the primes have density zero in the integers.

Thus, computational complexity is not the correct notion of complexity to consider. Instead, Sarnak reaches for a notion of complexity coming from topological dynamics and ergodic theory.

2. Sarnak's Conjecture

2.1. **Flows and entropy.** A *topological dynamical system*, or a *flow*, is a pair (X, T) consisting of a compact metric space (X, d) and a continuous map $T\colon X \to X$. To each $x \in X$ and $f \in C(X)$ we may associate a *sampling sequence* $\xi\colon \mathbb{N} \to \mathbb{C}$ given by applying f to the orbit of x under T. That is, we set $\xi(n) = f(T^n x)$. We say that the system (X, T) is *Möbius disjoint* if (2) holds for every sampling sequence ξ. Precisely, (X, T) is Möbius disjoint if

$$\sum_{n=1}^{N} \mu(n) f(T^n x) = o(N) \tag{3}$$

for every $x \in X$ and $f \in C(X)$.

As a sanity check, we affirm that every sequence taking values in a finite set arises as the sampling sequence of such a dynamical system. Indeed, if $\xi(n)$ is a sequence valued in a finite set $\Omega \subset \mathbb{R}$ of real numbers, then we may realize ξ as a sampling sequence in the following way. Let $X = \Omega^{\mathbb{N}}$, where Ω and X have been equipped with the discrete and product topologies, respectively, and take $T\colon X \to X$ to be the shift map given by $T(x)_n = x_{n+1}$. Let $\pi_1\colon \Omega^{\mathbb{N}} \to \Omega$ denote projection onto the first factor so that $\pi_1(x) = x_1$. Taking x to be the sequence $\xi(1), \xi(2), \ldots$ and $f = \pi_1$, we see that the sequence $\xi(n)$ is the sampling sequence realized by f and x in the system (X, T). An even simpler dynamical system realizing ξ can be constructed as follows. Let X_ξ denote

the closure of the orbit of ξ under T, i.e., the closure of $\{\xi, T(\xi), T^2(\xi), \ldots\}$. Then $F_\xi = (X_\xi, T)$ is a topological dynamical system called the *shift flow* of ξ. From this perspective, the prime number theorem can be rephrased as a statement about the Möbius disjointness of the shift flow on the topological space of a singleton.

The advantage of this topological setup is that we inherit a very natural notion of "complexity" from ergodic theory—namely that of *entropy*. Let (X, T) be a topological dynamical system. For each $n \in \mathbb{N}$, we may define a metric d_n on X by

$$d_n(x, y) = \max_{0 \le i < n} d(T^i(x), T^i(y)). \tag{4}$$

We see that two points $x, y \in X$ are ε-close with respect to d_n if and only if their first n iterates under T are ε-close with respect to d. Given a point x and a neighborhood B of x, we see that d_n distinguishes the points in B that stay close to x under many iterates of T from those whose orbits under T diverge from that of x. A subset E of X is said to be (n, ε)-*separated* if each pair of distinct points in E is at least ε apart with respect to d_n. Let $N(n, \varepsilon)$ denote the maximum cardinality of an (n, ε)-separated set. Note that $N(n, \varepsilon)$ measures the maximum number of length-n orbit segments distinguishable at scale ε.

Definition 2.1. The *topological entropy* of a dynamical system $F = (X, T)$ is defined to be

$$h(F) = \lim_{\varepsilon \to 0} \left(\limsup_{n \to \infty} \frac{\log N(n, \varepsilon)}{n} \right). \tag{5}$$

We see that h measures the exponential growth of the number of distinct orbits of T. A dynamical system is called *deterministic* if it has zero entropy. Roughly speaking, a dynamical system has zero entropy if the number of orbit segments of length n distinguishable at scale ε grows subexponentially in n.

2.2. **The eponymous conjecture.** Sarnak's conjecture makes explicit the notion that *entropy* is a fruitful notion of complexity to consider. Precisely, Sarnak predicts the following:

Conjecture 2.2 (Sarnak's conjecture, [27, Conjecture 4])**.** *Every deterministic topological dynamical system is Möbius disjoint. Precisely, for a dynamical system $F = (X, T)$ with $h(F) = 0$, we have that* (3) *holds for every $x \in X$ and $f \in C(X)$.*

We clarify the notion of entropy and the statement of Sarnak's conjecture with an elementary example. Consider the space $\{\pm 1\}$, which we equip with the discrete metric. Let $X = \{\pm 1\}^{\mathbb{N}}$, and let T be the shift map. Then, as a product space, X is metrizable with metric $d \colon X \times X \to \mathbb{R}_{\ge 0}$ given by

$$d(x, y) = \begin{cases} 0 & \text{if } x = y; \\ 2^{-m(x,y)} & \text{otherwise,} \end{cases}$$

where $m(x, y)$ is the first index where x and y differ. Thus, for ε (depending on n) chosen to be sufficiently small, we see that $d_n(x, y) < \varepsilon$ if and only if the first n entries of x and y agree. For such an ε, an (n, ε)-separated subset of X consists of sequences whose first n terms are pairwise distinct.

Given a sequence $x = (x_k)_k$ in X, a contiguous finite subsequence $x_{k+1}, x_{k+2}, \ldots, x_{k+n}$ of length n is called a *sign pattern* of length n. We restrict our attention to the shift flow X_x determined by the sequence x. For a sequence y, we have that $y \in X_x$ if and only if all of its sign patterns occur in x. Hence, the largest (n, ε)-separated subset of X_x is in bijection with the distinct length-n

sign patterns occurring in x, since two sequences are d_n-close if their first n entries agree. Therefore, we see from (5) that x is deterministic (we abuse notation here and refer to deterministic sequences as those whose shift flow (X_x, T) is deterministic) if and only if the number of distinct length-n sign patterns in x grows subexponentially in n. Sarnak conjectures that μ is disjoint from all such deterministic sequences.

As a final example (which will be important later), we reformulate the prime number theorem in arithmetic progressions as an instance of Sarnak's conjecture. Recall the statement of the prime number theorem in arithmetic progressions.

Theorem 2.3 (Prime number theorem in arithmetic progressions)**.** *We have that*

$$\sum_{\substack{n\leq N\\ n\equiv a \bmod q}} \mu(n) = o(N)$$

for any $q \in \mathbb{N}$ and $a \in \mathbb{Z}/q\mathbb{Z}$.

To rephrase this in terms of dynamics, let (X, T) be the topological dynamical system given by $X = \mathbb{Z}/q\mathbb{Z}$ and $T\colon X \to X$ the map taking $x \mapsto x + 1$. Since X is discrete, we see immediately that $h(X, T) = 0$, as $N(n, \varepsilon) = q$ for any n and any sufficiently small ε. Sarnak's conjecture predicts that

$$\sum_{n=1}^{N} \mu(n) f(T^n x) = o(N)$$

for any q-periodic sequence f and any $x \in X$. We see that the instance of Sarnak's conjecture with $x = 0$ and $f = \mathbb{1}_{\{a\}}$ is precisely the prime number theorem in arithmetic progressions.

2.3. **Roadmap.** The remainder of this essay can be divided thematically into two parts. In the first, corresponding to Section 3, we continue our discussion of Möbius pseudorandomness by relating Sarnak's conjecture to Chowla's conjecture—another famous conjecture from number theory. We first describe the randomness heuristic underlying this conjecture. Then we state Chowla's conjecture (Conjecture 3.1) and make its relation to Sarnak's conjecture (Conjecture 2.2) explicit by showing that Chowla's conjecture implies Sarnak's conjecture (Theorem 3.2). This section and its contents are loosely adapted from Tao's blog [31].

The second section is dedicated to proving a few important instances of Sarnak's conjecture using the orthogonality criterion developed by Bourgain, Sarnak, and Ziegler. Section 4, which is adapted from Section 2 of the paper by the aforementioned authors [4], is dedicated to stating and proving the orthogonality criterion (Theorem 4.1). In Section 5, we use the orthogonality criterion to recover Davenport's classical result (Theorem 5.1) on the Möbius disjointness of Kronecker systems (although we do not offer a rate of convergence). The crux of this section is a proof of Möbius disjointness for polynomial phases (Theorem 5.2) using Fourier analytic tools and equidistribution. A good reference for this section, upon which it is loosely based, is Montgomery's lectures on harmonic analysis and analytic number theory [21].

In Section 6, we demonstrate another application of the orthogonality criterion to give a (slightly abbreviated) proof of Sarnak's conjecture for horocycle flows in the compact case, adapting Sections 3 and 4 of the paper by Bourgain, Sarnak, and Ziegler [4]. In doing so, we introduce several notions from ergodic theory and give a brief exposition of Ratner's theorems for the unfamiliar reader. Finally, in Section 7, we conclude the paper by summarizing the main achievements of the paper and highlighting several future directions of inquiry.

3. Chowla's Conjecture

3.1. **Motivation and statement.** Another Möbius pseudorandomness heuristic concerns the correlations of μ with shifts of itself. Let $X_1, X_2, \ldots$ be a sequence of uniform independent random variables valued in $\{\pm 1\}$. Then for $a_1, \ldots, a_m \in \{1, 2\}$, not all even, the properties of expectation imply that

$$\sum_{n=1}^{N} X_{n+1}^{a_1} \cdots X_{n+m}^{a_m} = o(N).$$

Motivated by the analogy that $\mu(n)$ behaves randomly on squarefree n, we formulate the following conjecture (due to Chowla) about the correlations of μ with itself.

Conjecture 3.1 (Chowla's conjecture; see [6])**.** *Fix an integer m. For $a_1, \ldots, a_m \in \{1, 2\}$, not all even, we have that*

$$\sum_{n=1}^{N} \mu(n+1)^{a_1} \cdots \mu(n+m)^{a_m} = o(N). \tag{6}$$

Thus far, Chowla's conjecture has proved to be unyielding: except in the case where only one of the a_i's is odd—in which the statement of the conjecture is (nearly) equivalent to that of the prime number theorem in arithmetic progressions—the full conjecture is seemingly out of reach. Our reasons for introducing Chowla's conjecture are twofold, aside from its independent arithmetic importance.[2] Firstly, it is a formulation of Möbius pseudorandomness that is more elementary and straightforward than Sarnak's conjecture. Secondly, Sarnak's conjecture is actually a consequence of Chowla.

Theorem 3.2. *Chowla's conjecture implies Sarnak's conjecture.*

3.2. **Proof that Chowla implies Sarnak.** The proof of Theorem 3.2 involves three main steps, which we condense into three lemmas. For the sake of clarity, we present proofs of the lemmas in reverse order. Our strategy is adapted from [31] and can be summarized as follows. Let $f : \mathbb{N} \to \mathbb{C}$ be a sampling sequence associated to a deterministic topological dynamical system. To show (2), we consider

$$\frac{1}{m} \sum_{i=1}^{m} \mu(n+i) f(n+i) \tag{7}$$

as a random variable, where n is drawn uniformly at random from $[1, N] \cap \mathbb{Z}$. Firstly, we show that Sarnak's conjecture follows from showing that the expectation of (7) can be made arbitrarily small as m (and N) get arbitrarily large.

Lemma 3.3. *With notation as in the above,* (2) *holds if*

$$\mathbb{E}\left(\frac{1}{m} \sum_{i=1}^{m} \mu(n+i) f(n+i) \right) \to 0 \tag{8}$$

as $m \to \infty$, where the quantity in the expectation is the random variable given by sampling n uniformly at random from $[1, N] \cap \mathbb{Z}$.

Next, the following lemma gives us conditions for when (8) holds.

[2]There are connections between Chowla's conjecture and the twin prime conjecture; see [31] for more details

Lemma 3.4. *We have that* (8) *holds if the random variable*

$$\frac{1}{m}\left(c_1\mu(n+1)+\cdots+c_m\mu(n+m)\right) \tag{9}$$

concentrates with exponentially high probability, where $c_1,\dots,c_m$ are bounded, deterministic coefficients.

Finally, using the moment method, we prove that (9) indeed concentrates with exponentially high probability.

Lemma 3.5. *Assuming Chowla's conjecture, for any $m \geq 1$, any $\varepsilon > 0$, and any $c_1,\dots,c_m$ with $|c_i| \leq 1$, there exists an absolute constant C such that*

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} c_i\mu(n+i)\right| \geq \varepsilon\right) \leq C\exp(-\varepsilon^2 m/C) + o_{N\to\infty;\varepsilon,m}(1).$$

Here by $o_{N\to\infty;\varepsilon,m}(1)$ we mean that $o_{N\to\infty;\varepsilon,m}(1)$ tends to 0 as $N\to\infty$ for fixed m and ε. As usual n is drawn uniformly at random from 1 to N.

Proof. The main input into the proof is the moment method. Chebyshev's inequality tells us that

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} c_i\mu(n+i)\right| \geq \varepsilon\right) \leq \frac{1}{(\varepsilon m)^k}\mathbb{E}\left(\left|\sum_{i=1}^{m} c_i\mu(n+i)\right|^k\right). \tag{10}$$

The choice of k, which we take to be some large even integer, will be specified later. Expanding out the kth power and using the fact that the c_i's have norm at most 1, the triangle inequality allows us to bound the right-hand side of (10) by

$$\frac{1}{(\varepsilon m)^k}\sum_{1\leq i_1,\dots,i_k\leq m}\left|\mathbb{E}\left(\mu(n+i_1)\cdots\mu(n+i_k)\right)\right|. \tag{11}$$

The lemma will follow from bounding the summands in (11) and some basic combinatorics. Each of the summands in question is bounded trivially from above by 1. Furthermore, our assumption of Chowla's conjecture guarantees that the summands in which at least one of the indices i_j occurs an odd number of times can be bounded above by $o_{N\to\infty;m}(1)$. Thus, the problem reduces to one of counting the tuples of indices $(i_1,\dots,i_k)$ such that no index occurs an odd number of times. We do so in the following paragraph.

Running through the indices left to right, we see that there are two types of i_j: ones that have not previously appeared and ones that are repeats of an index that has occurred before. We call the two types of indices *fresh* and *repeat*, respectively. If the index is fresh, then there are at most m choices for it; if the index is repeat, then there are at most k choices for it. Since each index appears an even number of times, there are at most $k/2$ fresh indices. Thus, we may bound the number of $(i_1,\dots,i_k)$ of the desired type from above by $2^k m^{k/2}k^{k/2}$, where the first multiplicand corresponds to choosing the fresh indices.

Plugging this in, we may bound (11) from above by

$$\left(\frac{4k}{\varepsilon^2 m}\right)^{k/2} + \frac{1}{(\varepsilon m)^k}\sum_{\substack{1\leq i_1,\dots,i_k\leq m\\ i_j \text{ occurs an odd}\\ \text{number of times}}} o_{N\to\infty;m}(1) \leq \left(\frac{4k}{\varepsilon^2 m}\right)^{k/2} + o_{N\to\infty;m,\varepsilon,k}(1).$$

Finally, letting k be the largest even integer less than $\varepsilon^2 m/20$, we may bound the right-hand side of the above by $(1/5)^{\varepsilon^2 m/40} + o_{N\to\infty;\varepsilon,m}(1) \leq C\exp(-\varepsilon^2 m/C) + o_{N\to\infty;\varepsilon,m}(1)$ for some absolute constant C. □

We may prove Lemmas 3.4 and 3.3 using Lemma 3.5, but first we will need the following setup. Let f be a deterministic sequence. By taking real and imaginary parts, we may assume that f is real-valued. Moreover, we may assume that f is bounded in magnitude by 1 since our underlying space is compact. Let $\varepsilon > 0$ and $m \geq 1$. Since f is deterministic, we may realize it as a sampling sequence of some zero-entropy topological dynamical system (X, T) with $f(n) = F(T^n x)$ for $x \in X$ and $F \in C(X)$ with $|F| \leq 1$. Because X is compact, f is uniformly continuous, and thus there exists $\delta > 0$ such that, for any $u, v \in X$ with $d(u, v) < \delta$, we have that $|F(u) - F(v)| < \varepsilon$. Take $E_m \subset X$ to be a maximal (m, δ)-separated set, and define

$$S_m = \{(F(y), F(Ty), \ldots, F(T^{m-1}y)) \mid y \in E_m\} \subset [-1, 1]^m.$$

By definition, we have that

$$\#S_m \leq \#E_m = N(m, \delta) = \exp(o_{m\to\infty;\varepsilon,f}(m)), \tag{12}$$

where the last equality follows from the fact that (X, T) has zero entropy. Moreover, for every $z \in X$ there exists $y \in E_m$ such that $d_m(z, y) < \delta$, as otherwise we could enlarge E_m by z, contradicting its maximality as an (m, δ)-separated set.

The idea is to use the set S_m to approximate the values $f(n + i)$, which we do as follows. Setting $z = T^{n+1}x$ in the discussion above yields $y \in E_m$ such that $d_m(T^{n+1}x, y) < \delta$. Thus, for each $1 \leq i \leq m$, we have that

$$d(T^{n+i}x, T^{i-1}y) < \delta.$$

Therefore, by uniform continuity, we have that $F(T^{i-1}y)$ closely approximates $f(n + i)$: for $1 \leq i \leq m$, we have

$$|f(n + i) - F(T^{i-1}y)| = |F(T^{n+i}x) - F(T^{i-1}y)| < \varepsilon.$$

Proof of Lemma 3.4. With notation as in the above, we see that

$$\left|\frac{1}{m}\sum_{i=1}^{m} f(n + i)\mu(n + i)\right| \leq \sup_{(c_1,\ldots,c_m)\in S_m}\left|\frac{1}{m}\sum_{i=1}^{m} c_i\mu(n + i)\right| + \varepsilon$$

for each n, so

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} f(n + i)\mu(n + i)\right| \geq 2\varepsilon\right) \leq \mathbb{P}\left(\sup_{(c_1,\ldots,c_m)\in S_m}\left|\frac{1}{m}\sum_{i=1}^{m} c_i\mu(n + i)\right| \geq \varepsilon\right). \tag{13}$$

Applying Lemma 3.5, (12), and the union bound, we see that

$$\mathbb{P}\left(\sup_{(c_1,\ldots,c_m)\in S_m}\left|\frac{1}{m}\sum_{i=1}^{m} c_i\mu(n + i)\right| \geq \varepsilon\right) \leq \left(C\exp(-\varepsilon^2 m/C) + o_{N\to\infty;\varepsilon,m}(1)\right)\exp(o_{m\to\infty;\varepsilon,f}(m)),$$

which, combined with (13), tells us that

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} f(n + i)\mu(n + i)\right| \geq 2\varepsilon\right) \leq \left(C\exp(-\varepsilon^2 m/C) + o_{N\to\infty;\varepsilon,m}(1)\right)\exp(o_{m\to\infty;\varepsilon,f}(m)). \tag{14}$$

By choosing m to be so large that $C\exp(-\varepsilon^2 m/C)\exp(o_{m\to\infty;\varepsilon,f}(m)) \le \varepsilon$, we replace (14) by

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} f(n+i)\mu(n+i)\right| \ge 2\varepsilon\right) \le \varepsilon + o_{N\to\infty;\varepsilon,m}(1). \tag{15}$$

Our assumption that f is bounded in magnitude by 1 tells us that the random variable

$$X = \frac{1}{m}\sum_{i=1}^{m} f(n+i)\mu(n+i)$$

is bounded in magnitude by 1. Thus,

$$\mathbb{E}\left(\left|\frac{1}{m}\sum_{i=1}^{m} f(n+i)\mu(n+i)\right|\right) = \mathbb{E}\left(|X|\mathbb{1}_{|X|\le 2\varepsilon}\right) + \mathbb{E}\left(|X|\mathbb{1}_{|X|\ge 2\varepsilon}\right).$$

The first summand we may bound from above by 2ε. The second we may bound from above by $|X|\,\mathbb{P}(|X| \ge 2\varepsilon) \le \mathbb{P}(|X| \ge 2\varepsilon)$, since $|X| \le 1$. Putting these together, (15) implies that

$$\left|\mathbb{E}\left(\frac{1}{m}\sum_{i=1}^{m} f(n+i)\mu(n+i)\right)\right| \le \mathbb{E}\left(\left|\frac{1}{m}\sum_{i=1}^{m} f(n+i)\mu(n+i)\right|\right) \le 3\varepsilon + o_{N\to\infty;\varepsilon,m}(1), \tag{16}$$

as desired. □

Proof of Lemma 3.3. With notation as in the above, we have that

$$\left|\sum_{n=1}^{N} f(n+i)\mu(n+i) - \sum_{n=1}^{N} f(n)\mu(n)\right| \le 2i \sup_{n\le N} |f(n)\mu(n)| \le 2m.$$

Linearity tells us that

$$\left|\mathbb{E}\left(\frac{1}{m}\sum_{i=1}^{m} f(n+i)\mu(n+i)\right) - \frac{1}{N}\sum_{n=1}^{N} f(n)\mu(n)\right| = o_{N\to\infty;m}(1).$$

Combined with (16), this yields

$$\left|\sum_{n=1}^{N} f(n)\mu(n)\right| \le \left(3\varepsilon + o_{N\to\infty;\varepsilon,m}(1)\right) N,$$

which is precisely Sarnak's conjecture. □

Remark 3.6. We note that a careful read of the proof of Theorem 3.2 reveals that we have not used any properties specific to the Möbius function μ other than that it takes values in $\{-1, 0, 1\}$. Replacing μ in Theorem 3.2 by any such sequence satisfying an analogue of Chowla's conjecture will yield an analogous result.

4. The Orthogonality Criterion

4.1. **The BSZ criterion.** In the remainder of the essay, we prove several instances of Möbius disjointness. Namely, we prove Möbius disjointness for polynomial phases and Sarnak's conjecture for horocycle flows. The main tool for doing so is the following "orthogonality criterion" due to Bourgain, Sarnak, and Ziegler [4], which says, roughly, that a sequence is orthogonal to μ if its values on multiples of distinct primes are orthogonal. For more on the orthogonality criterion, we refer the reader to [4, 30].

Theorem 4.1 (Orthogonality criterion [10, Theorem 2.13]). *Let $F\colon \mathbb{N} \to \mathbb{C}$ be a bounded sequence of complex numbers. Assume that, for all pairs of distinct prime numbers p and q, we have*

$$\sum_{n=1}^{N} F(pn)\overline{F(qn)} = o(N). \tag{17}$$

Then, for each multiplicative function ν with $|\nu| \leq 1$, we have that

$$\sum_{n=1}^{N} F(n)\nu(n) = o(N).$$

We see immediately that any sequence satisfying the hypotheses of Theorem 4.1 is Möbius disjoint. The advantage of this criterion is that it reduces analyzing sums of $f(T^n x)\mu(n)$ to analyzing sums $f(T^{pn}x)f(T^{qn}x)$. As we will see in the subsequent sections, we often have a greater number of tools for carrying out the latter analysis. The remainder of this section is devoted entirely to the proof of Theorem 4.1. It will be helpful to prove the following, more precise version of the theorem.

Proposition 4.2 (cf. [4, Theorem 2]). *With notation as in Theorem 4.1, let τ be a small positive real number. Assume that, for all distinct primes $p, q \leq e^{1/\tau}$, we have that*

$$\left|\sum_{n=1}^{N} F(pn)\overline{F(qn)}\right| \leq \tau N \tag{18}$$

for sufficiently large N. Then, for N sufficiently large, we have that

$$\left|\sum_{n=1}^{N} F(n)\nu(n)\right| \leq 6\sqrt{\tau}\log(1/\tau)N.^{3}$$

We begin with a high-level overview of the proof strategy. In the remainder of this section, we abuse notation and use $[1, N]$ to denote $[1, N] \cap \mathbb{Z}$. Also, we take $A \preccurlyeq B$ to denote that $A \leq B$ asymptotically as $N \to \infty$. We begin by cutting up $[1, N]$ into pieces, the number of which will depend on τ. The pieces approximate the interval up to a small fraction and can be written as product sets whose factors are coprime. Using this decomposition, we rewrite the sum

$$\sum_{n=1}^{N} F(n)\nu(n)$$

as a sum over the pieces and leverage the multiplicativity of ν, the fact that $|\nu|, |F| \leq 1$, and some applications of Cauchy–Schwarz to conclude the theorem.

Let $\alpha > 0$ be a small parameter depending on τ that we will specify later. Consider the quantities

$$j_0 = \frac{1}{\alpha}\left(\log\frac{1}{\alpha}\right)^3, \quad j_1 = j_0^2, \quad D_0 = (1+\alpha)^{j_0}, \quad D_1 = (1+\alpha)^{j_1}.$$

[3] Comparing this with Theorem 2 in [4], one sees immediately that the bounds are not the same. This discrepancy is explained in (22) and the footnote that follows it.

Take $S = \{n \in [1, N] \mid n \text{ has a prime factor in } (D_0, D_1)\}$. Applying a sieve, we see that

$$\#[1, N] \setminus S \preccurlyeq \prod_{\substack{D_0 < \ell < D_1 \\ \ell \text{ prime}}} \left(1 - \frac{1}{\ell}\right) N.^4$$

When α is small, we have that D_0 is large, so we may apply Mertens' third theorem to see that

$$\#[1, N] \setminus S \preccurlyeq \prod_{\substack{D_0 < \ell < D_1 \\ \ell \text{ prime}}} \left(1 - \frac{1}{\ell}\right) N \preccurlyeq \frac{\log D_0}{\log D_1} N = \frac{N}{j_0} \preccurlyeq \alpha N. \tag{19}$$

We can interpret (19) as saying that all but a small fraction of $n \in [1, N]$—a fraction of αN-many to be precise—have a prime divisor in (D_0, D_1).

Next, we divide the n in S into smaller subsets based on the smallest prime divisor of n in (D_0, D_1). For $j_0 \le j \le j_1$, let P_j be the set of primes in $[(1+\alpha)^j, (1+\alpha)^{j+1}]$. We may estimate the cardinality of P_j using the prime number theorem (with error estimates) by

$$\#P_j = \frac{(1+\alpha)^{j+1}}{(j+1)\log(1+\alpha)} - \frac{(1+\alpha)^j}{j \log(1+\alpha)} + O((1+\alpha)^j e^{-c\sqrt{\alpha j}})$$

for some constant c. Some algebraic manipulation and the approximation $\log(1+\alpha) \asymp \alpha$ (since α is small) shows that we may bound $\#P_j$ from above by

$$\#P_j = (1+\alpha)^j \left(\frac{\alpha j - 1}{j(j+1)\log(1+\alpha)} + O(e^{-c\sqrt{\alpha j}})\right) \le (1+\alpha)^j \left(\frac{1}{j} + \frac{1}{\alpha j^2} + O(e^{-c\sqrt{\alpha j}})\right). \tag{20}$$

Now, set

$$S_j = \left\{ n \in [1, N] \mid n \text{ has a single divisor in } P_j \text{ and no divisors in } \bigcup_{i<j} P_j \right\}.$$

That is, the smallest prime divisor p of $n \in S_j$ with $p \in (D_0, D_1)$ must lie in P_j, and no other prime in P_j divides n. By definition, the S_j's are disjoint and

$$S \setminus \bigcup_{j_0 \le j \le j_1} S_j \subset \bigcup_{j_0 \le j \le j_1} \{n \in [1, N] \mid n \text{ has at least two prime factors in } P_j\},$$

allowing us to bound the cardinality of this set from above by

$$\#\left(S \setminus \bigcup_{j_0 \le j \le j_1} S_j\right) \preccurlyeq \sum_{j_0 \le j \le j_1} \sum_{\ell_1, \ell_2 \in P_j} \frac{N}{\ell_1 \ell_2} \le N \sum_{j_0 \le j \le j_1} \left(\frac{\#P_j}{(1+\alpha)^j}\right)^2.$$

Plugging in the bound on $\#P_j$ from (20), we see that the above is less than or equal to

$$N \sum_{j_0 \le j \le j_1} \left(\frac{1}{j} + \frac{1}{\alpha j^2} + O\left(e^{-c\sqrt{\alpha j}}\right)\right)^2 \le N \left(\frac{1}{j_0} + \frac{1}{j_0^3 \alpha^2} + O\left(\frac{1}{\alpha}\left(1 + c\sqrt{\alpha j_0}\right) e^{-c\sqrt{\alpha j_0}}\right)\right) \le \alpha N. \tag{21}$$

Thus, up to a fraction of α, the S_j's "partition" $[1, N]$, since they are disjoint sets that nearly cover the entire interval $[1, N]$.

[4] Either the Selberg sieve or the sieve of Eratosthenes–Legendre applies. We apply the sieve with $A = [1, N]$ and $P(z) = \prod_{D_0 \le \ell \le D_1} \ell$. We note that $A_\ell = N/\ell + O(1)$, so the sieve hypothesis holds.

We conclude the setup by approximating S_j as the product of a P_j with another set whose elements' prime factorizations are controlled. Set

$$Q_j = \left\{ m \in [1, N/(1+\alpha)^{j+1}] \mid m \text{ has no prime factors in } \bigcup_{i \le j} P_j \right\}.$$

Immediately we see from the definitions that $P_j Q_j \subset S_j$ for each $j_0 \le j \le j_1$; we will show that the collection of the product sets $P_j Q_j$ covers all but a small fraction of $[1, N]$. By definition,

$$S_j \setminus P_j Q_j \subset P_j \cdot \left[\frac{N}{(1+\alpha)^{j+1}}, \frac{N}{(1+\alpha)^j} \right],$$

implying that

$$\sum_{j_0 \le j \le j_1} \#(S_j \setminus P_j Q_j) \le \sum_{j_0 \le j \le j_1} \#P_j \cdot \frac{\alpha N}{(1+\alpha)^j}.$$

Substituting the upper bound from (20), we can bound the above by

$$N\alpha \sum_{j_0 \le j \le j_1} \left(\frac{1}{j} + \frac{1}{\alpha j^2} + O(e^{-c\sqrt{\alpha j}}) \right) \le N \left(\alpha \log(j_1/j_0) + \frac{1}{j_0} + O\left((1 + c\sqrt{\alpha j_0}) e^{-c\sqrt{\alpha j_0}} \right) \right), \tag{22}$$

where we estimate each term using integration. Working through the definitions of j_0 and j_1, we see that $\alpha \log(j_1/j_0)$, $1/j_0$, and $(1 + c\sqrt{\alpha j_0})e^{-c\sqrt{\alpha j_0}}$ are each bounded above by $\alpha \log(1/\alpha)$. Thus, making α sufficiently small then guarantees that

$$\sum_{j_0 \le j \le j_1} \#(S_j \setminus P_j Q_j) \preccurlyeq 2\alpha \log(1/\alpha) N.^5$$

Combining this with (19) and (21), we have that

$$\# \left([1, N] \setminus \bigcup_{j_0 \le j \le j_1} P_j Q_j \right) \preccurlyeq 4\alpha \log(1/\alpha) N.$$

We now use this "partition" of $[1, N]$ into the $P_j Q_j$'s to estimate the sum

$$\sum_{n=1}^{N} \nu(n) F(n).$$

Since the natural map $P_j \times Q_j \to P_j Q_j$ is bijective, the fact that $|\nu|, |F| \le 1$ implies that

$$\left| \sum_{n=1}^{N} \nu(n) F(n) \right| \preccurlyeq \sum_{j_0 \le j \le j_1} \left| \sum_{x \in P_j, y \in Q_j} \nu(xy) F(xy) \right| + 4\alpha \log(1/\alpha) N.$$

Using the multiplicativity of ν and the fact that $|\nu| \le 1$, we have that

$$\left| \sum_{n=1}^{N} \nu(n) F(n) \right| \preccurlyeq \sum_{j_0 \le j \le j_1} \sum_{y \in Q_j} \left| \sum_{x \in P_j} \nu(x) F(xy) \right| + 4\alpha \log(1/\alpha) N. \tag{23}$$

[5] *Conferatur* [4, Equation 2.13]. Bourgain, Sarnak, and Ziegler conclude that $\sum_{j_0 \le j \le j_1} \#(S_j \setminus P_j Q_j) \preccurlyeq 2\alpha N$. However, it is not clear how they are able to replace the factor of $N\alpha \log(1/\alpha)$ in (22) by $N\alpha$. Thus, I have rewritten the statement of the theorem and its proof to reflect this. We note that these changes do not affect Theorem 4.1 or its subsequent application in later sections.

Cauchy–Schwarz then allows us to bound the sum in the middle:

$$\sum_{y\in Q_j}\left|\sum_{x\in P_j}\nu(x)F(xy)\right| \le \left(\sum_{y\in Q_j}1\right)^{1/2}\left(\sum_{y\in Q_j}\left|\sum_{x\in P_j}\nu(x)F(xy)\right|^2\right)^{1/2}. \tag{24}$$

Then replacing the sum over $y \in Q_j$ and expanding out the square, we see that

$$\begin{aligned}\sum_{y\in Q_j}\left|\sum_{x\in P_j}\nu(x)F(xy)\right| &\le (\#Q_j)^{1/2}\left(\sum_{y\le N/(1+\alpha)^j}\left|\sum_{x\in P_j}\nu(x)F(xy)\right|^2\right)^{1/2}\\ &= (\#Q_j)^{1/2}\left(\sum_{y\le N/(1+\alpha)^j}\sum_{x_1,x_2\in P_j}\nu(x_1)\overline{\nu(x_2)}F(x_1y)\overline{F(x_2y)}\right)^{1/2} \qquad (25)\\ &\le (\#Q_j)^{1/2}\left(\sum_{x_1,x_2\in P_j}\left|\sum_{y\le N/(1+\alpha)^j}F(x_1y)\overline{F(x_2y)}\right|\right)^{1/2},\end{aligned}$$

where the final inequality was deduced from the fact that $|\nu| \le 1$. To bound the sum appearing in the above, we consider the diagonal contribution (where $x_1 = x_2$) separately. Since $|F| \le 1$, the sum over the diagonal in (25) contributes at most

$$(\#Q_j)^{1/2}(\#P_j)^{1/2}\frac{\sqrt{N}}{(1+\alpha)^{j/2}}.$$

Summing over $j_0 \le j \le j_1$ (as in (23)), another application of Cauchy–Schwarz tells us that the diagonal contributes at most

$$\sqrt{N}\left(\sum_{j_0\le j\le j_1}(\#P_j)(\#Q_j)\right)^{1/2}\left(\sum_{j_0\le j\le j_1}\frac{1}{(1+\alpha)^j}\right)^{1/2} \le N\left(\sum_{j_0\le j\le j_1}\frac{1}{(1+\alpha)^j}\right)^{1/2} \le \alpha N \tag{26}$$

in total, where the first inequality follows from the fact that $P_jQ_j \subset S_j$ and that $\bigcup_j S_j \subset [1, N]$; the second from the formula for partial geometric series. In particular, we have the bound

$$\sum_{j_0\le j\le j_1}\frac{1}{(1+\alpha)^j} \le \sum_{j\ge j_0}\frac{1}{(1+\alpha)^j} = \frac{(1+\alpha)^{-j_0}}{1-(1+\alpha)^{-1}} = \frac{1+\alpha}{\alpha}(1+\alpha)^{-j_0}.$$

Writing $(1+\alpha)^{-j_0}$ as $\exp(-j_0\log(1+\alpha))$, the approximation $\log(1+\alpha) \asymp \alpha$ for α small gives us that

$$(1+\alpha)^{-j_0} \le \exp(-j_0\alpha) = \exp(-(\log 1/\alpha)^3).$$

Thus, for α sufficiently small, we have that $(1+\alpha)^{-j_0} \le \alpha^2$, which gives us the bound in (26).

For the off-diagonal contribution, the idea is to apply (18). Begin by noting that, for all $x_1, x_2 \in P_j$, we have $x_1, x_2 < (1+\alpha)^{j_1} < e^{1/\alpha^2}$. Thus, for $\alpha \ge \sqrt{\tau}$, we may apply the hypothesis of Proposition 4.2 (since $x_1 \ne x_2$) to get that

$$\left|\sum_{y\le N/(1+\alpha)^j}F(x_1y)\overline{F(x_2y)}\right| \le \frac{\tau N}{(1+\alpha)^j}.$$

Therefore, the off-diagonal contribution in (23) is at most

$$\begin{aligned}
&\sqrt{\tau N}\sum_{j_0\le j\le j_1}\#P_j(\#Q_j)^{1/2}(1+\alpha)^{-j/2}\\
&\qquad\le \sqrt{\tau N}\left(\sum_{j_0\le j\le j_1}\#P_j\#Q_j\right)^{1/2}\left(\sum_{j_0\le j\le j_1}\#P_j(1+\alpha)^{-j}\right)^{1/2}\\
&\qquad\le N\sqrt{\tau}\left(\log(j_1/j_0)+\frac{1}{j_0\alpha}+\frac{1}{\alpha}O((1+c\sqrt{\alpha j_0})e^{-c\sqrt{\alpha j_0}})\right)^{1/2}\\
&\qquad\preccurlyeq N\sqrt{\tau}\sqrt{\log(1/\alpha)},
\end{aligned}\tag{27}$$

where the first inequality follows from Cauchy–Schwarz; the second from the fact that $\sum_j \#P_j\#Q_j \le N$ and (20); the third from the fact that $\log(j_1/j_0) = \log(1/\alpha) + 3\log\log(1/\alpha) \preccurlyeq \log(1/\alpha)$, that $1/(\alpha j_0) = 1/(\log(1/\alpha))^3 \preccurlyeq \log(1/\alpha)$, and that $O((1+c\sqrt{\alpha j_0})e^{-c\sqrt{\alpha j_0}}) \preccurlyeq \log(1/\alpha)$. Amalgamating (23), (26), and (27), we have that

$$\left|\sum_{n=1}^{N}\nu(n)F(n)\right| \preccurlyeq N(\alpha+\sqrt{\tau\log(1/\alpha)}+4\alpha\log(1/\alpha)).\tag{28}$$

Taking $\alpha=\sqrt{\tau}$ implies Proposition 4.2 and thus Theorem 4.1.

Remark 4.3. We note here that Theorem 4.1 still holds if we relax the hypotheses of the theorem so that (17) holds for all but finitely many pairs of distinct primes. Indeed, tracing through the proof of Proposition 4.2, we notice that (18) is only applied for primes in (D_0, D_1). Thus, we only need (18) to hold for primes that are at least $D_0 = (1+\alpha)^{j_0}$. We can make D_0 arbitrarily large by taking τ to be sufficiently small; hence, in application, we may allow (17) to fail for finitely many prime pairs.

5. Polynomial Phases

5.1. **A quick proof of Davenport's theorem.** Having worked hard to prove the orthogonality criterion, we now leverage it to reap arithmetic profit. Namely, we spend the remainder of the paper discussing applications of the criterion to proving various instances of Sarnak's conjecture. In most cases, proving that the orthogonality criterion applies still requires some serious analytic input. A naïve application of the orthogonality criterion, however, instantly yields the following classical asymptotic, a quantitative version of which is due to Davenport [9].

Theorem 5.1 (cf. [9, Equation 1])**.** *For $\alpha\in\mathbb{R}$, we have that*

$$\sum_{n=1}^{N}\mu(n)e^{2\pi i\alpha n} = o(N).\tag{29}$$

While Davenport's work—which relies on the theory of L-functions and Vinogradov bilinear forms—gives an explicit rate at which the average tends to 0, the advantage of the orthogonality criterion gives us a "quick and dirty" proof of Möbius disjointness that fits in a few lines.[6] When $\alpha = a/b$ is rational, we see that the orthogonality criterion *fails* for pairs of primes p and q

[6]Precisely, Davenport shows that $\sum_{n=1}^{N}\mu(n)e^{2\pi i\alpha n} = O(N(\log N)^{-h})$ uniformly in α for any fixed h, where the implied constant depends only on h.

congruent modulo b. In this case, the proof of (29) is morally equivalent to that of the prime number theorem in arithmetic progressions.

Proof of Theorem 5.1. When α is rational, then a simple reduction shows that (29) is the prime number theorem in arithmetic progressions (see Theorem 2.3 and the subsequent discussion).

When α is irrational, we apply the orthogonality criterion to the sequence $e^{2\pi i\alpha n}$. Let p and q be distinct prime numbers. Then

$$\sum_{n=1}^{N} e^{2\pi i\alpha pn}\overline{e^{2\pi i\alpha qn}} = \sum_{n=1}^{N} e^{2\pi i(p-q)\alpha n}.$$

Since α is irrational, we have that $e^{2\pi i(p-q)\alpha n} \neq 1$. The expression above is then a partial geometric series and is equal to

$$\sum_{n=1}^{N} e^{2\pi i(p-q)\alpha n} = \frac{1 - e^{2\pi i(p-q)\alpha(N+1)}}{1 - e^{2\pi i(p-q)\alpha}} - 1 = o(N),$$

and Theorem 4.1 implies the result. □

We may rephrase this in the language of Sarnak's conjecture as follows. Let $X = \mathbb{R}/\mathbb{Z}$, and let $T\colon X \to X$ be given by $x \mapsto x + \alpha$ for some $\alpha \in \mathbb{R}$. The flow (X, T) has zero entropy because the metrics d_n are all equal, since $d(T^n(x), T^n(y)) = d(x + n\alpha, y + n\alpha) = d(x, y)$. Sarnak's conjecture states that, for any $f \in C(X)$ and $x \in X$, we have

$$\sum_{n=1}^{N} \mu(n) f(T^n x) = \sum_{n=1}^{N} \mu(n) f(x + n\alpha) = o(N). \tag{30}$$

By the Stone–Weierstrass theorem, trigonometric polynomials (finite combinations of the basis functions $e^{2\pi ikx}$) are dense in $C(X)$, so (30) follows from approximating f by a finite linear combination of $e^{2\pi ikx}$'s and applying Theorem 5.1.

5.2. **Möbius disjointness of polynomial phases.** Another advantage of presenting Davenport's theorem through the lens of the orthogonality criterion is that Theorem 5.1 is a first instance of Möbius disjointness for polynomial phases—the subject of the remainder of this section. A *polynomial phase* is a sequence $(u_n)_n = (e^{2\pi iP(n)})_n$ for some polynomial $P(x) \in \mathbb{R}[x]$ with real coefficients. In this language, we see that Theorem 5.1 is simply the statement that μ is orthogonal to polynomial phases of degree 1 (note that the degree-0 case is subsumed by the prime number theorem). Polynomial phases are indeed deterministic sequences, but it is not easy to see this from our setup. Instead, for a justification of this claim, we refer the reader to Example 3.3 of [14], which identifies polynomial phases as nilsequences, and Theorem 3.2 of [16], which states that nilsequences are deterministic. The goal of this section is to prove the following instance of Sarnak's conjecture.

Theorem 5.2. *Let $P(x) \in \mathbb{R}[x]$. Then*

$$\sum_{n=1}^{N} \mu(n) e^{2\pi iP(n)} = o(N). \tag{31}$$

The high-level strategy for the proof of the theorem is outlined below. We first divide into cases based on whether P has irrational coefficients. Assume that all of the coefficients of P are rational. Let q be the least common denominator of the coefficients. Then

$$\sum_{n=1}^{N} \mu(n) e^{2\pi i P(n)} = \sum_{a=0}^{q-1} e^{2\pi i P(a)} \sum_{\substack{n \leq N \\ n \equiv a \bmod q}} \mu(n) = \sum_{a=0}^{q-1} e^{2\pi i P(a)} o(N) = o(N), \tag{32}$$

where the penultimate equality follows from the prime number theorem in arithmetic progressions (see Theorem 2.3). Thus, we have that (31) holds for polynomials with rational coefficients.

Now, we assume that P has a nonconstant irrational coefficient. (If the only irrational coefficient is the constant one, then we may factor this out of $e^{2\pi i P(n)}$ and reduce to the case in the above.) The idea is then to apply the orthogonality criterion (Theorem 4.1) of the previous section. By the orthogonality criterion, Theorem 5.2 follows if, for all pairs of distinct primes p and q,

$$\sum_{n=1}^{N} e^{2\pi i (P(pn) - P(qn))} = o(N). \tag{33}$$

The main tools we will leverage to prove (33) come from the theory of *equidistribution*. The role of this theory is fleshed out further in the following subsection. Let $P(x) = a_d x^d + \cdots + a_1 x + a_0$ for $a_0, a_1, \ldots, a_d \in \mathbb{R}$, and consider the polynomial $\tilde{P}(x) = P(px) - P(qx) \in \mathbb{R}[x]$. Since p and q are distinct, we have that $a_d(p^d - q^d) \neq 0$, so $\tilde{P}$ is also of degree d. Moreover, $\tilde{P}$ has at least one nonconstant irrational coefficient, since $a_j(p^j - q^j)$ is irrational if a_j is. Thus, it suffices to show

$$\sum_{n=1}^{N} e^{2\pi i Q(n)} = o(N) \tag{34}$$

for any polynomial $Q(x) \in \mathbb{R}[x]$ of degree d at least 1 with at least one nonconstant irrational coefficient.

5.3. **Some results in equidistribution theory.** Next, we introduce the necessary concepts from the theory of equidistribution. For a more in-depth treatment of these methods, we refer the reader to Chapters 1 and 2 of Montgomery's lectures [21], from which this subsection is loosely adapted. Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ denote the circle group. A sequence $(u_n)_n$ of points in $\mathbb{T}$ is said to be *uniformly distributed* if, for each $0 \leq \alpha \leq 1$, we have that

$$\lim_{N \to \infty} \frac{\#\{n \in [1, N] \cap \mathbb{Z} \mid 0 \leq u_n \leq \alpha \bmod 1\}}{N} = \alpha. \tag{35}$$

The key ingredient we will need is Weyl's criterion:

Lemma 5.3 (Weyl's equidistribution criterion)**.** *Let $(u_n)_n$ be a sequence of points in $\mathbb{T}$. Then the following are equivalent:*

(*1*) *$(u_n)_n$ is uniformly distributed;*
(*2*) *for each nonzero integer k, we have that*

$$\sum_{n=1}^{N} e^{-2\pi i k u_n} = o(N);$$

(3) *for any properly Riemann-integrable F on $\mathbb{T}$, we have*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} F(u_n) = \int_{\mathbb{T}} F(\alpha)\,\mathrm{d}\alpha.$$

Proof. We begin by proving that (3) implies (2). Setting $F(x) = e^{-2\pi ikx}$, we see that F is Riemann-integrable, so

$$\frac{1}{N}\sum_{n=1}^{N} e^{-2\pi iku_n} = \frac{1}{N}\sum_{n=1}^{N} F(u_n) \to \int_{\mathbb{T}} F(\alpha)\,\mathrm{d}\alpha = 0.$$

To see why (2) implies (3), we begin by proving the implication for F continuous. Now, recall that the trigonometric polynomials—the space of functions spanned by $e_k(x) = e^{-2\pi ikx}$—are uniformly dense in the space of continuous functions on $\mathbb{T}$. That is, for each continuous F on $\mathbb{T}$ and $\varepsilon > 0$ there exists a trigonometric polynomial T such that $\|F - T\|_\infty < \varepsilon$. It follows that

$$\left|\int_{\mathbb{T}} F(\alpha) - T(\alpha)\,\mathrm{d}\alpha\right| \le \varepsilon \quad \text{and} \quad \left|\frac{1}{N}\sum_{n=1}^{N} F(u_n) - \frac{1}{N}\sum_{n=1}^{N} T(u_n)\right| \le \varepsilon.$$

We see that it suffices to show the implication for trigonometric polynomials, as then $\frac{1}{N}\sum_{n=1}^{N} T(u_n) \to \int_{\mathbb{T}} T(\alpha)\,\mathrm{d}\alpha$ as $N \to \infty$, from which we may conclude that

$$\limsup_{N\to\infty} \left|\frac{1}{N}\sum_{n=1}^{N} F(u_n) - \int_{\mathbb{T}} F(\alpha)\,\mathrm{d}\alpha\right| \le 2\varepsilon.$$

Given a trigonometric polynomial $T(x) = \sum_{|k|\le K} c_k e_k(x)$, we have that

$$\frac{1}{N}\sum_{n=1}^{N} T(u_n) = \frac{1}{N}\sum_{|k|\le K} c_k \sum_{n=1}^{N} e_k(u_n) = \frac{1}{N}\left(Nc_0 + \sum_{0\ne|k|\le K} c_k \sum_{n=1}^{N} e^{-2\pi iku_n}\right).$$

Now, (2) implies that, as $N \to \infty$, the above approaches c_0, which is precisely the value of $\int_{\mathbb{T}} T(\alpha)\,\mathrm{d}\alpha$.

Finally, assume that F is Riemann integrable. Recall that the integrability of F guarantees the following: for any $\varepsilon > 0$, there exist continuous functions F^+ and F^- with $F^- \le F \le F^+$ and $\int_{\mathbb{T}} F^+ - F^- < \varepsilon$. Then

$$\frac{1}{N}\sum_{n=1}^{N} F^-(u_n) \le \frac{1}{N}\sum_{n=1}^{N} F(u_n) \le \frac{1}{N}\sum_{n=1}^{N} F^+(u_n).$$

Since F^+ and F^- are continuous, our work in the above shows that the left-hand side tends to $\int_{\mathbb{T}} F^-(\alpha)\,\mathrm{d}\alpha \ge \int_{\mathbb{T}} F(\alpha)\,\mathrm{d}\alpha - \varepsilon$, and the right-hand side tends to $\int_{\mathbb{T}} F^+(\alpha)\,\mathrm{d}\alpha \le \int_{\mathbb{T}} F(\alpha)\,\mathrm{d}\alpha + \varepsilon$. It follows that

$$\left|\frac{1}{N}\sum_{n=1}^{N} F(u_n) - \int_{\mathbb{T}} F(\alpha)\,\mathrm{d}\alpha\right| \le \varepsilon.$$

We easily see that (3) implies (1) by taking F to be $\mathbb{1}_{[0,\alpha]}$ for $0 \le \alpha \le 1$, the indicator function of the interval $[0, \alpha]$.

Lastly, we prove that (1) implies (3). Note that (1) gives us (3) for $F = \mathbb{1}_{[0,\alpha]}$. The result follows from recalling the fact that any Riemann integrable function F can be approximated from above and below by step functions (linear combinations of indicators of intervals) and by emulating the arguments from the previous paragraph. □

We see immediately that if we can show that $Q(n)$ is uniformly distributed modulo 1, then Weyl's criterion with $k = -1$ implies (34) and thus Theorem 5.2. To show equidistribution of $Q(n)$, the idea is to induct on the degree of the polynomial. The key step in the induction invokes Van der Corput's theorem, which says that a sequence $(u_n)_n$ is uniformly distributed if the difference sequence $(u_{n+h} - u_n)_n$ is uniformly distributed for each $h \in \mathbb{N}$.

Lemma 5.4 (Van der Corput's theorem)**.** *Let $(u_n)_n$ be a sequence of points in $\mathbb{T}$. If, for each positive integer h, the sequence $(u_{n+h} - u_n)_n$ is uniformly distributed, then the sequence $(u_n)_n$ is uniformly distributed.*

We remark briefly that the converse of Van der Corput's theorem is false: the proof of Theorem 5.1 shows that the sequence $u_n = \alpha n$ for $\alpha \in \mathbb{R}$ irrational is uniformly distributed; however, $u_{n+h} - u_n = \alpha h$ is constant (and thereby not uniformly distributed). The main input to the proof of Van der Corput's theorem is Van der Corput's lemma, whose statement and proof follow.

Lemma 5.5 (Van der Corput's lemma)**.** *Let H be a positive integer. Then for any complex numbers $y_1, \dots, y_N$, we have that*

$$\left|\sum_{n=1}^{N} y_n\right|^2 \le \frac{N+H}{H+1} \sum_{n=1}^{N} |y_n|^2 + \frac{2(N+H)}{H+1} \sum_{h=1}^{H} \left(1 - \frac{h}{H+1}\right) \left|\sum_{n=1}^{N-h} y_{n+h}\overline{y_n}\right|. \tag{36}$$

Proof. The result is essentially a clever application of Cauchy–Schwarz. For simplicity of notation, we set $y_n = 0$ when $n < 1$ or $n > N$. Set $\alpha_m = 1/(H+1)$ for $0 \le m \le H$ so that

$$\sum_n y_n = \left(\sum_m \alpha_m\right)\left(\sum_n y_n\right) = \sum_m \alpha_m \sum_n y_{m+n} = \sum_{n=-H+1}^{N} \sum_m \alpha_m y_{m+n}.$$

Cauchy–Schwarz then tells us that

$$\left|\sum_n y_n\right|^2 \le (N+H) \sum_n \left|\sum_m \alpha_m y_{m+n}\right|^2 = (N+H) \sum_{m_1} \sum_{m_2} \alpha_{m_1}\overline{\alpha_{m_2}} \sum_n y_{m_1+n}\overline{y_{m_2+n}}, \tag{37}$$

where the equality follows from expanding the square and rearranging order of summation. Reindexing, we set $m = m_2$ and $h = m_1 - m_2$ so that the right-hand side of (37) can be expressed as

$$(N+H) \sum_h \sum_m \alpha_{m+h}\overline{\alpha_m} \sum_n y_{m+h+n}\overline{y_{m+n}} = (N+H) \sum_h \left(\sum_m \alpha_{m+h}\overline{\alpha_m}\right)\left(\sum_n y_{n+h}\overline{y_n}\right), \tag{38}$$

where the equality follows by reindexing, which is valid because the value of $\sum_n y_{m+n+h}\overline{y_{m+n}}$ depends only on h. Because $\alpha_m = 1/(H+1)$ for $0 \le m \le H$ and is 0 otherwise,

$$(H+1)^2 \sum_m \alpha_{m+h}\overline{\alpha_m}$$

simply counts the $m \in [0, H] \cap \mathbb{Z}$ such that $m + h \in [0, H] \cap \mathbb{Z}$ as well. We must have $|h| \le H$ (otherwise this quantity is zero); thus,

$$\sum_m \alpha_{m+h}\overline{\alpha_m} = \frac{H + 1 - |h|}{(H+1)^2}.$$

Now, substituting this into (38) and noting that the contributions of the h and $-h$ terms (for $h \ne 0$) are conjugate to each other implies (36). □

Proof of Lemma 5.4. An application of Weyl's criterion tells us that it suffices to show that

$$\sum_{n=1}^{N} e^{-2\pi i k u_n} = o(N)$$

for each nonzero integer k. Fix $k \neq 0$. The idea is to apply Van der Corput's lemma with $y_n = e_k(u_n) = e^{-2\pi i k u_n}$. Thus, $|y_n|^2 = 1$, so by (36) we have that, for each positive integer H,

$$\left|\sum_{n=1}^{N} e^{-2\pi i k u_n}\right|^2 = \left|\sum_{n=1}^{N} y_n\right|^2 \leq \frac{N(N+H)}{H+1} + \frac{2(N+H)}{H+1}\sum_{h=1}^{H}\left(1 - \frac{h}{H+1}\right)\left|\sum_{n=1}^{N-h} y_{n+h}\overline{y_n}\right|. \tag{39}$$

We compute that

$$\sum_{n=1}^{N-h} y_{n+h}\overline{y_n} = \sum_{n=1}^{N-h} e_k(u_{n+h} - u_n).$$

By hypothesis, $(u_{n+h} - u_n)_n$ is equidistributed for each h, so Weyl's criterion tells us that the sum on the right-hand side of the above is $o(N)$ for each h. Substituting this into (39), we get that

$$\left|\sum_{n=1}^{N} e^{-2\pi i k u_n}\right|^2 \leq \frac{N(N+H)}{H+1} + \frac{2(N+H)}{H+1}\sum_{h=1}^{H}\left(1 - \frac{h}{H+1}\right)o(N) \leq \frac{N(N+H)}{H+1} + \frac{2(N+H)H}{H+1}o(N),$$

where the last inequality follows from the fact that $1 - h/(H+1) \leq 1$. Dividing each side by N^2 yields

$$\left|\frac{1}{N}\sum_{n=1}^{N} e^{-2\pi i k u_n}\right|^2 \leq \frac{N+H}{N(H+1)} + \frac{2(N+H)H}{N^2(H+1)}o(N),$$

and letting $N \to \infty$, we deduce that

$$\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{N} e^{-2\pi i k u_n}\right| \leq \frac{1}{\sqrt{H+1}}.$$

Because this holds for every positive integer H (see Lemma 5.4), it follows that

$$\frac{1}{N}\sum_{n=1}^{N} e^{-2\pi i k u_n} = o(1),$$

as desired. □

5.4. **Back to polynomial phases.** Finally, we are ready to show that polynomial sequences $Q(n)$ are equidistributed, which we will use to show that polynomial phases are Möbius disjoint.

Proposition 5.6. *Let $Q(x) = a_d x^d + \cdots + a_1 x + a_0 \in \mathbb{R}[x]$ be a polynomial of degree $d \geq 1$ with at least one nonconstant irrational coefficient. Then $Q(n)$ is uniformly distributed modulo 1.*

Proof. We induct on the degree d of Q. In the base case $d = 1$, we have $Q(n) = a_1 n + a_0$ for irrational a_1. Arguing as in the proof of Theorem 5.1 shows that

$$e^{-2\pi i a_0 k}\sum_{n=1}^{N} e^{-2\pi i k a_1 n} = o(N)$$

for each nonzero integer k, since the sum in question is a geometric series. Thus, the Weyl criterion implies uniform distribution modulo 1.

For the inductive step, assume that $R(n)$ is uniformly distributed for all polynomials $R(x)$ of degree $d-1$ with at least one nonconstant irrational coefficient. Let Q be as in the statement of the result. Let j be the largest index $1 \leq j \leq d$ such that a_j is irrational. By Van der Corput's theorem, we have that $Q(n)$ is uniformly distributed if the shift sequence $Q(n+h) - Q(n)$ is uniformly distributed for each positive integer h. We see that

$$Q(x+h) - Q(x) = a_d(x+h)^d + \cdots + a_1(x+h) + a_0 - a_d x^d - \cdots - a_1 x - a_0,$$

so $Q(x+h) - Q(x)$ is a polynomial of degree $d-1$ in x. The x^{j-1}-coefficient of $Q(x+h) - Q(x)$ is given by

$$jha_j + \sum_{j<k\leq d} \binom{k}{j-1} a_k h^{k-j+1}.$$

The sum in the above is a rational combination of the a_k's, all of which must be rational by the definition of j. Therefore, the x^{j-1}-coefficient of $Q(x+h) - Q(x)$ is irrational.

From here we divide into cases based on the value of j. If $j = 1$, then let q be the least common denominator of $a_2, \ldots, a_d$, which are rational by the definition of j. Write

$$\sum_{n=1}^{N} e^{-2\pi i k Q(n)} = \sum_{r=0}^{q-1} \sum_{\substack{m \\ qm+r\leq N}} e^{-2\pi i k Q(qm+r)}.$$

Expanding out $Q(qm+r)$ as a polynomial in m, we see that

$$Q(qm+r) = a_d(qm+r)^d + \cdots + a_1(qm+r) + a_0 = \beta_r + (a_1 q + \gamma_r)m + \delta_r,$$

where β_r is an integer-valued polynomial in m, γ_r is a rational number, and δ_r is a real number, all depending on r. We note that β_r is integer-valued because the terms of degree at least 2 in m have coefficients that are integer multiples of $a_j q^\ell$ for some $\ell \geq 2$. Thus, we have that

$$\sum_{r=0}^{q-1} \sum_{\substack{m \\ qm+r\leq N}} e^{-2\pi i k Q(qm+r)} = \sum_{r=0}^{q-1} \sum_{\substack{m \\ qm+r\leq N}} e^{-2\pi i k((a_1 q+\gamma_r)m+\delta_r)}.$$

Applying the base case, the above is $o(N)$ for each integer k; Weyl's criterion implies that $Q(n)$ is equidistributed, as desired.

If $1 < j \leq d$, then we have that $Q(x+h)-Q(x)$ is a real polynomial of degree $d-1$ with at least one nonconstant irrational coefficient. By the inductive hypothesis, we have that $Q(n+h)-Q(n)$ is uniformly distributed for each positive integer h. Van der Corput's theorem then tells us that $Q(n)$ must be uniformly distributed. □

Armed with this result, we are poised to prove that polynomial phases are Möbius disjoint.

Proof of Theorem 5.2. We divide into cases. First, consider the case where at least one of the nonconstant coefficients is irrational. Recall that by the orthogonality criterion (Theorem 4.1), it suffices to show that such polynomial phases are uniformly distributed mod 1. This reduction is the content of (33), (34), and an application of Weyl's equidistribution criterion (Lemma 5.3). In the case in which all of the nonconstant coefficients are rational is handled by (32) □

6. Horocycle Flows

6.1. **Möbius disjointness of horocycle flows.** The orthogonality criterion of Section 4 was developed by Bourgain, Sarnak, and Ziegler to prove Sarnak's conjecture for horocycle flows. An exposition of this result and their methods is the subject of this section; for more details, we refer the reader to Sections 3 and 4 of [4].

Let $G = \mathrm{SL}_2(\mathbb{R})$, and let $\mu = \mathrm{d}g$ be the unique (up to scaling) left-invariant Haar measure on G. Let Γ be a lattice in G, i.e., a discrete subgroup of G such that $X = \Gamma\backslash G$ has finite volume.[7] It is a standard fact that μ descends to a G-invariant probability measure on $\Gamma\backslash G$; we will often abuse notation by denoting this by $\mu = \mathrm{d}g$. For $\lambda \in \mathbb{R}$, let

$$u^\lambda = \begin{bmatrix} 1 & \lambda \\ & 1 \end{bmatrix}$$

denote the corresponding element of the one-parameter subgroup of upper-triangular unipotent matrices in G. For convenience, we will take $u = u^1$. To each such u^λ, we may associate a *discrete time horocycle flow* (X, T^λ), where X is equipped with the quotient topology and T^λ is the map $X \to X$ given by

$$T^\lambda(\Gamma x) = \Gamma x u^\lambda. \tag{40}$$

Likewise, we set $T = T^1$. A result of Gurevich [15] tells us that (X, T^λ) is a deterministic flow; the main theorem of Bourgain, Sarnak, and Ziegler is that the flow (X, T) is Möbius disjoint.

Theorem 6.1 ([4, Theorem 1])**.** *Let (X, T) be the horocycle flow. For all $x \in X$ and $f \in C(X)$, we have that*

$$\sum_{n=1}^{N} \mu(n) f(T^n x) = o(N).^8 \tag{41}$$

For the sake of simplicity, we assume henceforth that $X = \Gamma\backslash G$ is compact. This case is appreciably simpler and best for illustrating the main ideas and techniques involved in the proof of Theorem 6.1, which are largely the same in both cases. The curious reader can find an explanation for how to adapt the proof of the compact case to the noncompact setting in Remark 6.13 and in [4, Sections 3 and 4].

Fix $f \in C(X)$ and $x \in X$. By the orthogonality criterion (Theorem 4.1 and Remark 4.3), Theorem 6.1 will follow if we can show, for all but finitely many pairs of distinct primes p and q, that

$$\sum_{n=1}^{N} f(T^{pn} x)\overline{f(T^{qn} x)} = o(N). \tag{42}$$

As in Section 5, we will need serious analytic input in order to evaluate this sum. The tools we use to analyze (42) are Ratner's theorems—a series of fundamental results from ergodic theory.

---

[7] By finite volume, we mean that any fundamental domain $\mathcal{F}$ for the action of Γ on G has $\mu(\mathcal{F}) < \infty$.

[8] The careful reader will note that we have abused notation and that if $X = \Gamma\backslash G$ is not compact, then (X, T) is not a topological dynamical system as defined in Section 2. For example, if $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, then the classical work of Siegel shows that $X = \Gamma\backslash G$ is noncompact and has a cusp "going off to infinity." This discrepancy is handled as follows: Since X is locally compact and Hausdorff, it has a one-point compactification $\widehat{X}$; in the case where X is noncompact, we assume implicitly that $f \in C(\widehat{X})$.

6.2. **Ratner's theorem and ergodic theory.** Ratner's main theorem tells us that each orbit of the horocycle flow on X is either closed or dense.[9] We may further refine this statement using the language of equidistribution: in addition to classifying the orbit closures of unipotent flows, Ratner showed that each orbit is in fact uniformly distributed.

Theorem 6.2 (Ratner orbit closure theorem with equidistribution [26, Theorem B])**.** *Let L be a Lie group, and let Λ be a lattice in L. Set $Y = \Lambda\backslash L$, and let $\ell \in L$ be a unipotent element. Then, for each $y \in Y$, there exists a connected, closed subgroup $H \leq L$ such that $\ell \in H$, the subset $yH \subset Y$ is closed and has finite H-invariant volume, and*

$$\overline{\{y\ell^n \mid n \in \mathbb{N}\}} = yH.$$

Let μ_H denote the unique H-invariant probability measure on yH. Then, for every $\Phi \in C_c(Y)$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Phi(y\ell^n) = \int_{yH} \Phi \, d\mu_H.$$

This theorem begets the introduction of the following notions from ergodic theory.

Definition 6.3. Let (Y, T) be a flow. A probability measure ν on Y is said to be *T-invariant* if $\nu(T^{-1}(A)) = \nu(A)$ for every Borel subset $A \subset X$. A T-invariant probability measure ν on Y is said to be *ergodic* if every T-invariant Borel set has measure 0 or 1. The system (Y, T) is *uniquely ergodic* if it has a unique T-invariant Borel probability measure.

Given a flow (Y, T) with a T-invariant probability measure ν, a sequence $x_1, x_2, \ldots$ in Y is said to be *uniformly distributed* with respect to ν if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \varphi(x_n) = \int_Y \varphi \, \mathrm{d}\nu$$

for all $\varphi \in C(Y)$.

Definition 6.4. Let (Y, T) be a flow. Given a T-invariant probability measure ν, a point $y \in Y$ is said to be *generic* for the triple (Y, T, ν) if its forward orbit $y, Ty, T^2y, \ldots$ is uniformly distributed. That is, if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \varphi(T^n y) = \int_Y \varphi \, \mathrm{d}\nu$$

for every $\varphi \in C(Y)$. A standard result from ergodic theory tells us that every point is generic in a uniquely ergodic system (see [29, Exercise 10]).

The following result of Furstenberg tells us that, for compact X, the horocycle flow is uniquely ergodic.

Lemma 6.5 (Furstenberg [11, Corollary 6.13])**.** *Suppose X is compact. Then the horocycle flow (X, T^λ) for any nonzero λ is uniquely ergodic. Thus, every point of X is generic for $(X, T^\lambda, \mathrm{d}g)$.*

[9] We qualify this statement by noting that this is a specification of Ratner's results to the setting $G = \mathrm{SL}_2(\mathbb{R})$. For a fully general (and complete) treatment of Ratner's results, see [22].

6.3. **Joinings.** Next, we introduce the notion of joinings—a powerful tool in ergodic theory that is essential to the proof of Theorem 6.1. Informally, joinings allow us to couple together two measure preserving flows into a dynamical system on the product space so that the projections onto each factor return the component flows.

Definition 6.6. Let (Y_1, T_1, ν_1) and (Y_2, T_2, ν_2) be measure-preserving flows. A *joining* of these systems is a Borel probability measure ρ on $Y_1 \times Y_2$ with the following significance:

(1) ρ is $(T_1 \times T_2)$-invariant;
(2) $(\pi_i)_*\rho = \nu_i$ for $1 \le i \le 2$, where $\pi_i \colon Y_1 \times Y_2 \to Y_i$ is the ith coordinate projection.

When $\rho = \nu_1 \times \nu_2$ we say it is *trivial*; otherwise, it is *nontrivial*.

One of the most important consequences of Ratner's theorem is a classification of the possible joinings of (X, T^λ) with itself. For simplicity, we will state a specialized version of Ratner's result for self-joinings of horocycle flows.

Theorem 6.7 (Ratner self-joinings theorem [25, Theorem 6])**.** *Consider G, Γ, and X as in the above, where X is equipped with its G-invariant probability measure, denoted μ. Then, for $\lambda_1, \lambda_2 \in \mathbb{R}$, every ergodic joining ν of (X, T^{λ_1}, μ) and (X, T^{λ_2}, μ) is* algebraic*. That is, either $\nu = \mu \times \mu$ is the trivial joining, or there exists a morphism $\psi \colon G \to G \times G$ of the form $\psi(g) = (\psi_1(g), \psi_2(g))$ with*

$$\psi_1(u) = u^{\lambda_1} \quad \text{and} \quad \psi_2(u) = u^{\lambda_2},$$

and there exists $(\Gamma x_1, \Gamma x_2) \in X \times X$ such that the orbit

$$\{(\Gamma x_1 \psi_1(g), \Gamma x_2 \psi_2(g)) \mid g \in G\} \subset X \times X$$

is closed, and ν is the unique H-invariant probability measure supported on this orbit, where we set $H = \psi(G)$.

We will apply Theorem 6.7 to the following scenario. Fix $\xi \in \Gamma\backslash G$, $\lambda_1, \lambda_2 \in \mathbb{R}_+$, and $\varphi, \psi \in C(X)$. Also fix a representative of ξ in G; we will occasionally abuse notation by denoting this by ξ as well. Taking $L = G \times G$, $\Lambda = \Gamma \times \Gamma$, $\ell = (u^{\lambda_1}, u^{\lambda_2})$, $y = (\xi, \xi)$, and $\Phi(x_1, x_2) = \varphi(x_1)\psi(x_2)$, Ratner's orbit-closure theorem, in particular the statement about equidistribution (see Theorem 6.2), tells us that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \varphi(\xi u^{\lambda_1 n})\psi(\xi u^{\lambda_2 n}) = \int_{\Gamma\backslash G \times \Gamma\backslash G} \Phi \, \mathrm{d}\nu, \tag{43}$$

where ν is an algebraic H-invariant Haar measure supported on $(\xi \times \xi)H$, and $(\xi \times \xi)H$ is closed in $\Lambda\backslash L$. In fact, ν is a joining.

Lemma 6.8. *The measure ν is a joining of (X, T^{λ_1}, μ) and (X, T^{λ_2}, μ).*

Proof. For all $\Omega \in C(X \times X)$, we have by Theorem 6.2 that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Omega(T^{\lambda_1 n}(\xi), T^{\lambda_2 n}(\xi)) = \int_{\Gamma\backslash G \times \Gamma\backslash G} \Omega \, \mathrm{d}\nu. \tag{44}$$

For $\omega \in C(X)$, set $\Omega_i(x_1, x_2) = \omega(x_i) = (\omega \circ \pi_i)(x_1, x_2)$. Plugging this into (44), we see that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \omega(T^{\lambda_i n}(\xi)) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Omega_i(T^{\lambda_1 n}(\xi), T^{\lambda_2 n}(\xi)) = \int_{\Gamma\backslash G \times \Gamma\backslash G} \Omega_i \, \mathrm{d}\nu.$$

The right-hand side of the above can then be expressed as

$$\int_{\Gamma\backslash G\times\Gamma\backslash G} \Omega_i \, \mathrm{d}\nu = \int_{X\times X} \omega \circ \pi_i \, \mathrm{d}\nu = \int_X \omega \, \mathrm{d}((\pi_i)_*\nu).$$

Now, by Lemma 6.5, ξ is generic for $(X, T^{\lambda_i}, \mathrm{d}g)$, from which it follows that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \omega(T^{\lambda_i n}(\xi)) = \int_X \omega \, \mathrm{d}g.$$

Putting these all together, we get

$$\int_X \omega \, \mathrm{d}g = \int_X \omega \, \mathrm{d}((\pi_i)_*\nu)$$

for all $\omega \in C(X)$. Thus, the projections of ν onto the first and second factors are exactly $\mathrm{d}g$, and we conclude that ν is a joining of (X, T^{λ_1}, μ) and (X, T^{λ_2}, μ). □

Using Theorem 6.7, we may classify precisely when ν is a nontrivial joining of (X, T^{λ_1}, μ) and (X, T^{λ_2}, μ). In order to do so, we must first develop the group-theoretic notions of commensurator and correlator subgroups, which play crucial roles in this characterization. Recall that two subgroups H_1 and H_2 of G are said to be *commensurable* if their intersection $H_1 \cap H_2$ has finite index in both H_1 and H_2.

Definition 6.9. Let Γ be a lattice in G. The *commensurator subgroup* $\mathrm{Comm}(\Gamma)$ is the set of elements $g \in G$ such that Γ and its conjugate by g are commensurable. That is,

$$\mathrm{Comm}(\Gamma) = \{g \in G \mid g^{-1}\Gamma g \cap \Gamma \text{ has finite index in both } \Gamma \text{ and } g^{-1}\Gamma g\}.$$

Now, recall that there is a natural action of $G = \mathrm{SL}_2(\mathbb{R})$ on the projective line $\mathbb{P}^1(\mathbb{R})$ by fractional linear transformations. In coordinates, an element $g \in \mathrm{SL}_2(\mathbb{R})$ takes $z \in \mathbb{R}$ to

$$g \cdot z = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot z = \frac{az+b}{cz+d}.$$

For $z \in \mathbb{P}^1(\mathbb{R})$, let $P_z \subset G$ denote the stabilizer of z with respect to this action. We compute explicitly using the change of coordinates $t \mapsto 1/t$ that

$$P_\infty = \left\{ \begin{bmatrix} \alpha & \beta \\ & \alpha^{-1} \end{bmatrix} \mid \alpha \neq 0 \right\}.$$

Since the stabilizer subgroups are conjugate, it follows that $P_z = g^{-1}P_\infty g$, where $g \in G$ is such that $g \cdot z = \infty$. We define a character $\chi \colon P_\infty \to \mathbb{R}_+$ by

$$\chi\left(\begin{bmatrix} \alpha & \beta \\ & \alpha^{-1} \end{bmatrix} \right) = \alpha^2;$$

since P_z is of the form $P_z = g^{-1}P_\infty g$, we may define χ in exactly the same way on P_z for any z.

Definition 6.10. Given a subgroup Δ_z of P_z, we define its *correlator subgroup* $C(\Delta_z)$ to be $\chi(\Delta_z)$, i.e., the subgroup of $\mathbb{R}_+$ given by the image of Δ_z under χ.

Set $C(\Gamma, z) = C(\mathrm{Comm}(\Gamma) \cap P_z)$. As we will see momentarily, the correlator subgroup is fundamental in characterizing when ν is a nontrivial joining.

Lemma 6.11 ([4, Lemma 4])**.** *With notation as in the above, ν is a nontrivial joining of (X, T^{λ_1}, μ) and (X, T^{λ_2}, μ) if and only if $\lambda_2/\lambda_1 \in C(\Gamma, \xi(\infty))$.*

Proof. By the Ratner joinings theorem (Theorem 6.7), we have that ν is nontrivial if and only if there exists a morphism $\psi\colon G \to G \times G$ such that $\psi = (\psi_1, \psi_2)$ with $\psi_1(u) = u^{\lambda_1}$ and $\psi_2(u) = u^{\lambda_2}$ and the orbit $\{(\xi\psi_1(g), \xi\psi_2(g)) \mid g \in G\}$ is closed in $X \times X$.[10] Now, each $\psi_i\colon G \to G$ is a nontrivial homomorphism of Lie groups, from which it follows that each $\mathrm{d}\psi_i\colon \mathfrak{sl}_2(\mathbb{R}) \to \mathfrak{sl}_2(\mathbb{R})$ is a nontrivial homomorphism of Lie algebras. Since $\mathfrak{sl}_2(\mathbb{R})$ is a simple Lie algebra, it follows that $\mathrm{d}\psi_i$ is an isomorphism of Lie algebras (simplicity implies that $\ker(\mathrm{d}\psi_i)$ is trivial; the domain and codomain are the same). Therefore, ψ_i is a Lie group automorphism of $G = \mathrm{SL}_2(\mathbb{R})$. Now, recall that every automorphism of $\mathrm{SL}_2(\mathbb{R})$ is given by conjugation by some element of $\mathrm{GL}_2(\mathbb{R})$, so $\psi_i(g) = \alpha_i g \alpha_i^{-1}$ for $\alpha_i \in \mathrm{GL}_2(\mathbb{R})$. Rescaling, we may assume that $\det(\alpha_i)^2 = 1$. In particular, we have that

$$\alpha_i u \alpha_i^{-1} = u^{\lambda_i},$$

from which we may conclude that $\alpha_i \in \mathrm{SL}_2(\mathbb{R})$[11] and that $\alpha_1, \alpha_2 \in P_\infty$. Writing

$$\alpha_i = \begin{bmatrix} a & b \\ & a^{-1} \end{bmatrix}$$

(which is possible since $\alpha_i \in P_\infty$), some algebra shows that

$$\alpha_i u \alpha_i^{-1} = u^{a^2},$$

so $a^2 = \lambda_i$. Thus, we conclude that

$$\chi(\alpha_1) = \lambda_1 \quad \text{and} \quad \chi(\alpha_2) = \lambda_2.$$

Now, reindexing shows that the orbit

$$\{(\xi\alpha_1 g \alpha_1^{-1}, \xi\alpha_2 g \alpha_2^{-1}) \mid g \in G\}$$

is closed in $X \times X$ if and only if the orbit

$$\{(\Gamma h \alpha_1^{-1}, \Gamma\xi\alpha_2\alpha_1^{-1}\xi^{-1} h \alpha_2^{-1}) \mid h \in G\}$$

is closed in $X \times X$. Setting $\delta = \xi\alpha_2\alpha_1^{-1}\xi^{-1}$, we want to show that $\{(\Gamma h\alpha_1^{-1}, \Gamma\delta h\alpha_2^{-1}) \mid h \in G\}$ is closed. Because right-multiplication by an (α_1, α_2) is a homeomorphism, we see that it suffices to show that $D = \{(\Gamma h, \Gamma\delta h) \mid h \in G\}$ is closed. Notably, D is the orbit of $(\Gamma, \Gamma\delta)$ under the action of the diagonal image Δ of G in $G \times G$. Now, a result of Raghunathan (see [24, Theorem 1.13] or [21, Exercises 22 and 24]) tells us that the orbit $(\Gamma, \Gamma\delta)$ under Δ is closed if and only if $\mathrm{Stab}_\Delta((\Gamma, \Gamma\delta))$ is a lattice in $\Delta \simeq G$. Now, we see that $(\Gamma, \Gamma\delta) \cdot (g, g) = (\Gamma g, \Gamma\delta g) = (\Gamma, \Gamma\delta)$ if and only if $g \in \Gamma$ and $g \in \delta^{-1}\Gamma\delta$. That is, $\mathrm{Stab}_\Delta(\Gamma, \Gamma\delta) \simeq \Gamma \cap \delta^{-1}\Gamma\delta$, so D is closed if and only if $\Gamma \cap \delta^{-1}\Gamma\delta$ is a lattice in G. Since $\Gamma \cap \delta^{-1}\Gamma\delta$ is a subgroup of Γ and $\delta^{-1}\Gamma\delta$, we see that $\Gamma \cap \delta^{-1}\Gamma\delta$ is a lattice in G if and only if $\Gamma \cap \delta^{-1}\Gamma\delta$ has finite index in Γ and $\delta^{-1}\Gamma\delta$. We see that this occurs precisely when $\delta \in \mathrm{Comm}(\Gamma)$.

Since $\delta = \xi\alpha_2\alpha_1^{-1}\xi^{-1} \in P_{\xi(\infty)}$, we compute that

$$\chi(\delta) = \chi(\alpha_2)\chi(\alpha_1)^{-1} = \lambda_2/\lambda_1.$$

From this we may conclude that ν is nontrivial if and only if $\lambda_2/\lambda_1 \in C(\Gamma, \xi(\infty))$. For the converse, given $\lambda_2/\lambda_1 \in C(\Gamma, \xi(\infty))$, tracing backwards through the proof allows us to construct δ, α_1, α_2, and thus ψ as in the above, from which it follows that ν is nontrivial. □

[10] Note that we are not applying Theorem 6.7 exactly as stated, as we have no reason to assume that ν should be ergodic. Theorem 6.7 was stated for ergodic joinings for the sake of brevity and expositional clarity. However, this can be rectified by decomposing ν into its ergodic components and applying the theorem to each component.

[11] Otherwise, one can explicitly compute that $\alpha_i u \alpha_i^{-1} \neq u^{\lambda_i}$ for $\lambda_i > 0$.

The final ingredient needed before we can prove Theorem 6.1 for $\Gamma\backslash G$ compact is a characterization of $C(\Gamma, z)$, which is given in the lemma below. While its proof is relatively elementary, it is beyond the scope of this paper—a more detailed account can be found in [4].

Lemma 6.12 ([4, Lemmas 1 and 2])**.** *If $\Gamma\backslash G$ is compact, then*

$$\left\{\frac{q}{p} \mid p, q \in \mathbb{P}, p \neq q\right\} \cap C(\Gamma, z)$$

is finite for all $z \in \mathbb{P}^1(\mathbb{R})$.

6.4. **Proof of the main theorem.** Having classified when ν is a nontrivial joining, we are now ready to prove Theorem 6.1 for $X = \Gamma\backslash G$ compact.

Proof of Theorem 6.1 in the compact case. We begin by writing

$$c = \int_{\Gamma\backslash G} f(g)\,\mathrm{d}g$$

so that we may express f as $f(x) = f_1(x)+c$ for some $f_1 \in C(X)$ with mean 0. That is, $\int_X f_1(g)\,\mathrm{d}g = 0$. Thus,

$$\sum_{n=1}^{N} \mu(n) f(T^n x) = \sum_{n=1}^{N} \mu(n) f_1(T^n x) + c\sum_{n=1}^{N} \mu(n) = \sum_{n=1}^{N} \mu(n) f_1(T^n x) + o(N),$$

where the final equality follows from the prime number theorem. We have reduced the theorem to showing that f_1 is Möbius disjoint. By the orthogonality criterion (Theorem 4.1) and Remark 4.3, it suffices to show that

$$\sum_{n=1}^{N} f_1(T^{pn}x)\overline{f_1(T^{qn}x)} = o(N) \tag{45}$$

for all but finitely many primes p and q.

We do so by showing that, given $x \in X$ and $\varphi \in C(X)$, the following holds for all but finitely many pairs of distinct primes p and q:

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \varphi(xu^{pn})\overline{\varphi(xu^{qn})} = \left|\int_{\Gamma\backslash G} \varphi(g)\,\mathrm{d}g\right|^2. \tag{46}$$

We see immediately that applying (46) with $\varphi = f_1$ will imply (45). In order to achieve (46), we will show something marginally stronger: that for $x \in X$ and $\varphi, \psi \in C(X)$,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \varphi(xu^{pn})\psi(xu^{qn}) = \left(\int_{\Gamma\backslash G} \varphi(g)\,\mathrm{d}g\right)\left(\int_{\Gamma\backslash G} \psi(g)\,\mathrm{d}g\right) \tag{47}$$

holds for all but finitely many pairs of distinct primes p and q. Substituting $\varphi = f_1$ and $\psi = \overline{f_1}$ into (47) will imply (46) and hence the theorem.

To that end, fix $\xi \in \Gamma\backslash G$, and let p and q be distinct primes. Taking $L = G \times G$, $\Lambda = \Gamma \times \Gamma$, $\ell = (u^p, u^q)$, $y = (\xi, \xi)$, and $\Phi(x_1, x_2) = \varphi(x_1)\psi(x_2)$, Ratner's orbit-closure theorem tells us that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \varphi(\xi u^{pn})\psi(\xi u^{qn}) = \int_{\Gamma\backslash G\times\Gamma\backslash G} \Phi\,\mathrm{d}\nu, \tag{48}$$

where ν is an H-invariant algebraic Haar measure supported on $(\xi\times\xi)H$, which is closed in L/Λ. By Lemma 6.8, we see that ν is a joining of (X, T^p, μ) with (X, T^q, μ).

In the above, we characterized exactly when ν is a nontrivial joining; we now profit from this hard work by showing that there are only finitely many pairs of primes p and q for which ν is a nontrivial joining of (X, T^p, μ) and (X, T^q, μ). Indeed, Lemma 6.11 tells us that ν is nontrivial if and only if $q/p \in C(\Gamma, \xi(\infty))$, and by Lemma 6.12, the set of such q/p is finite. Hence, by Remark 4.3, we may assume that ν is the trivial joining.

If ν is the trivial joining, then $\nu = \mu \times \mu = \mathrm{d}g \times \mathrm{d}g$, and (48) simplifies to

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \varphi(\xi u^{pn}) \psi(\xi u^{qn}) = \int_{\Gamma\backslash G \times \Gamma\backslash G} \Phi \, \mathrm{d}\nu = \left(\int_{\Gamma\backslash G} \varphi(g) \, \mathrm{d}g \right) \left(\int_{\Gamma\backslash G} \psi(g) \, \mathrm{d}g \right). \tag{49}$$

It follows that (47) holds for all but finitely many pairs of primes p, q. Therefore, (47) and thus (46) and (45) hold for all but finitely many pairs of primes, allowing us to conclude the theorem. □

Remark 6.13. The noncompact case of Theorem 6.1 is treated similarly but requires additional casework based on whether ξ is generic and Γ is arithmetic. When ξ is generic, the proofs of Lemmas 6.8 and 6.11 apply essentially without modification. From there, one must prove analogues of Lemma 6.12 classifying $C(\Gamma, \xi(\infty)) \cap \mathbb{Q}$, which is done via casework on whether Γ is arithmetic. When ξ is not generic, the theorem follows from a combination of the work of Dani [7] and Davenport [9].

The case in which X is noncompact is treated in greater detail by Bourgain, Sarnak, and Ziegler; see [4, Sections 3 and 4] for more details.

7. Conclusion

Möbius pseudorandomness remains one of the driving open problems in number theory, due to both its intrinsic interest and connections to keynote arithmetic conjectures such as the Riemann hypothesis and Chowla's conjecture. Sarnak's conjecture on Möbius disjointness (Conjecture 2.2) offers a compelling prediction for the way in which μ behaves pseudorandomly via the language of topological dynamical systems and entropy. The conjecture is particularly striking because it is a consequence of Chowla's conjecture (Theorem 3.2) and because, in several cases, it is actually tractable (unlike other formulations of Möbius pseudorandomness, such as Conjectures 1.2 and 3.1, that are still largely unresolved, despite significant effort).

In addition to motivating Sarnak's conjecture and explaining its connections to Chowla's conjecture, this essay was largely dedicated to giving an exposition of the orthogonality criterion of Bourgain, Sarnak, and Ziegler—one of the primary tools used in proving instances of Sarnak's conjecture. The orthogonality criterion can be used to prove Möbius disjointness in a variety of settings, ranging from Kronecker systems (Theorem 5.1) to polynomial phases (Theorem 5.2) to horocycle flows (Theorem 6.1). Each application of the orthogonality criterion requires serious additional analytic input, ranging from Fourier analysis to equidistribution theory to Ratner's theorems on unipotent flows. Another variant of the orthogonality criterion, due to Kátai, gives Sarnak's conjecture for the Thue–Morse sequence.

Let $(t_n)_n$ denote the Thue–Morse sequence, which is a sequence in $\{0, 1\}$ defined as follows. For $n \in \mathbb{N}$, write $n-1$ in binary and consider the number of ones in this expression. If this number is odd, then $t_n = 1$; otherwise, $t_n = 0$. The first 30 entries of t_n are given below:

$$0110100110010110100101100110 10\ldots$$

The Thue–Morse sequence is deterministic because it is 2-automatic: the number of distinct length-n sign patterns in an automatic sequence is linear in n; in particular, asymptotics for this quantity were computed in [5]. Another exciting instance of Sarnak's conjecture is that

the Möbius sequence is orthogonal to the Thue–Morse sequence. Shown by Indlekofer–Kátai in 2001 [17] and refined by Dartyge–Tenenbaum in 2005 [8], this result analyzes correlations of $(t_n)_n$ with itself and uses Kátai's variant of the orthogonality criterion to conclude that $(t_n)_n$ is Möbius disjoint. Later, without appealing to the orthogonality criterion, Mauduit and Rivat [20] proved a prime number theorem for the Thue–Morse sequence, from which they derived an explicit rate of convergence for the correlations of $(t_n)_n$ with μ.

While the orthogonality criterion is extraordinarily useful, it is not the only weapon in our arsenal. Indeed, several instances of Sarnak's conjecture have surrendered to other techniques. Of particular importance is the Möbius disjointness of polynomial nilsequences, which was shown by Green and Tao in 2012 [12]. Green and Tao showed the following instance of Sarnak's conjecture, which can be thought of as a vast generalization of Theorem 5.2.

Theorem 7.1 ([12, Theorem 1.1])**.** *Let G be a simply connected nilpotent Lie group, and let Γ be a discrete, cocompact subgroup so that $\Gamma\backslash G$ is a nilmanifold. Let $F\colon \Gamma\backslash G \to \mathbb{C}$ be a Lipschitz function, and let $g\colon \mathbb{Z} \to G$ be a polynomial sequence. Then*

$$\sum_{n=1}^{N} \mu(n) F(\Gamma g(n)) = O_{F,G,\Gamma,A}\left(\frac{N}{\log^A N}\right)$$

for all $A > 0$.

Polynomial nilsequences are deterministic by Theorem 3.2 of [16]. Letting $G = \mathbb{R}$ and $\Gamma = \mathbb{Z}$, we see that Theorem 7.1 recovers a more quantitative version—in that it offers an explicit rate of convergence—of Theorem 5.2 immediately. The proof of Theorem 7.1 is dense and requires advanced tools to prove precise quantitative asymptotics for the distribution of polynomial orbits on nilmanifolds from [13]. We refer the reader to [12] and [13] for an in-depth treatment.

We conclude the paper by offering several further avenues of exploration, all of which are, to our knowledge, open.[12] The first of these concerns Sarnak's conjecture for horocycle flows. One of the drawbacks of the orthogonality criterion is that it offers no explicit rate of convergence. While Sarnak's conjecture only asks that the correlations of a sequence with μ are $o(N)$, as we saw in Conjecture 1.2, one might draw sharper arithmetic conclusions when given an explicit rate of convergence. Thus, a natural follow-up to Theorem 6.1 is asking whether obtaining an explicit rate in (41) is possible. Because Ratner's theorems offer no such quantitative asymptotic, the methods used in Section 6 would either need to be abandoned or refined in order to achieve such an asymptotic. One way of making this explicit is to ask whether a prime number theorem for horocycle flows exists. See [10, Question 6] for a more detailed exposition.

Another interesting open problem concerns generalizations of the Thue–Morse sequence aptly named *generalized Morse sequences*. For a precise definition of generalized Morse sequences, which we omit for the sake of brevity, we refer the reader to the introduction of [18]. Inspired by the orthogonality of the Möbius function and the Thue–Morse sequence, Mauduit asked whether generalized Morse sequences are Möbius disjoint. Bourgain [3] proved an instance of this conjecture for a subclass of generalized Morse sequences called *Kakutani sequences*, which contains the simplest generalized Morse sequences after Thue–Morse itself. For a more in-depth survey on the Möbius disjointness of Kakutani sequences, we refer the reader to [10, Section 5.5]. Despite this progress, the Möbius disjointness of generalized Morse sequences remains largely open.

[12] A more complete—and more detailed—list of the recent progress on Sarnak's conjecture and several related open problems can be found in Section 5 of [10].

References

[1] Adleman, L. M., and McCurley, K. S. Open problems in number theoretic complexity, ii. In *Algorithmic Number Theory* (Berlin, Heidelberg, 1994), L. M. Adleman and M.-D. Huang, Eds., Springer Berlin Heidelberg, pp. 291–322.

[2] Agrawal, M., Kayal, N., and Saxena, N. PRIMES is in P. *Annals of Mathematics 160*, 2 (2004), 781–793.

[3] Bourgain, J. Möbius-Walsh correlation bounds and an estimate of Mauduit and Rivat. *Journal d'Analyse Mathématique 119*, 1 (2013), 147–163.

[4] Bourgain, J., Sarnak, P., and Ziegler, T. Disjointness of Möbius from horocycle flows. In *From Fourier Analysis and Number Theory to Radon Transforms and Geometry*, vol. 28 of *Developments in Mathematics*. Springer, New York, 2013, pp. 67–83.

[5] Brlek, S. Enumeration of factors in the thue-morse word. *Discrete Applied Mathematics 24*, 1-3 (1989), 83–96.

[6] Chowla, S. *The Riemann Hypothesis and Hilbert's Tenth Problem*, vol. 4 of *Mathematics and Its Applications*. Gordon and Breach Science Publishers, New York, 1965.

[7] Dani, S. G. On uniformly distributed orbits of certain horocycle flows. *Ergodic Theory and Dynamical Systems 2*, 2 (1982), 139–158.

[8] Dartyge, C., and Tenenbaum, G. Sommes des chiffres de multiples d'entiers. In *Annales de l'institut Fourier* (2005), vol. 55, pp. 2423–2474.

[9] Davenport, H. On some infinite series involving arithmetical functions. II. *The Quarterly Journal of Mathematics 8* (1937), 313–320.

[10] Ferenczi, S., Kułaga-Przymus, J., and Lemańczyk, M. *Sarnak's Conjecture: What's New*. Springer International Publishing, Cham, 2018, pp. 163–235.

[11] Furstenberg, H. The unique ergodicity of the horocycle flow. In *Recent Advances in Topological Dynamics* (Berlin, Heidelberg, 1973), A. Beck, Ed., Springer Berlin Heidelberg, pp. 95–115.

[12] Green, B., and Tao, T. The Möbius function is strongly orthogonal to nilsequences. *Annals of Mathematics 175*, 2 (2012), 541–566.

[13] Green, B., and Tao, T. The quantitative behaviour of polynomial orbits on nilmanifolds. *Annals of Mathematics* (2012), 465–540.

[14] Green, B., Tao, T., and Ziegler, T. An inverse theorem for the gowers Us+1 [n]-norm. *Annals of Mathematics 176*, 2 (2012), 1231–1372.

[15] Gurevich, B. M. The entropy of horocycle flows. *Doklady Akademii Nauk SSSR 136*, 4 (1961), 768–770. In Russian. Available at `https://www.mathnet.ru/eng/dan24575`.

[16] Host, B., Kra, B., and Maass, A. Complexity of nilsystems and systems lacking nilfactors. *Journal d'analyse mathématique 124*, 1 (2014), 261–295.

[17] Indlekofer, K.-H., and Kátai, I. Investigations in the theory of q-additive and q-multiplicative functions. i. *Acta Mathematica Hungarica 91*, 1 (2001), 53–78.

[18] Keane, M. Generalized Morse sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 10*, 4 (1968), 335–353.

[19] Ledoux, M., and Talagrand, M. *The Law of the Iterated Logarithm*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1991, pp. 196–234.

[20] Mauduit, C., and Rivat, J. Sur un problème de gelfond: la somme des chiffres des nombres premiers. *Annals of Mathematics* (2010), 1591–1646.

[21] Montgomery, H. L. *Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis*, vol. 84 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, Providence, RI, 1994. Published for the Conference Board of the Mathematical Sciences.

[22] Morris, D. W. *Ratner's Theorems on Unipotent Flows*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, 2005.

[23] Mussardo, G., and LeClair, A. Randomness of Möbius coefficients and Brownian motion: growth of the Mertens function and the Riemann hypothesis. *Journal of Statistical Mechanics: Theory and Experiment 2021*, 11 (nov 2021), 113106.

[24] Raghunathan, M. S. *Discrete Subgroups of Lie Groups*, vol. 68 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 2. Folge*. Springer-Verlag, Berlin, Heidelberg, 1972.

[25] Ratner, M. Horocycle flows, joinings and rigidity of products. *Annals of Mathematics 118*, 2 (1983), 277–313.

[26] Ratner, M. Raghunathan's topological conjecture and distributions of unipotent flows. *Duke Mathematical Journal 63*, 1 (1991), 235–280.

[27] Sarnak, P. Three lectures on the Möbius function, randomness and dynamics. https://publications.ias.edu/sites/default/files/MobiusFunctionsLectures%282%29.pdf, 2011. Lecture notes, Institute for Advanced Study.

[28] Shapiro, H. Some assertions equivalent to the prime number theorem for arithmetic progressions. *Communications on Pure and Applied Mathematics 2* (11 2006), 293 – 308.

[29] Tao, T. 254a, Lecture 9: Ergodicity. https://terrytao.wordpress.com/2008/02/04/254a-lecture-9-ergodicity/, Feb. 2008. Blog post, What's new.

[30] Tao, T. The Katai–Bourgain–Sarnak–Ziegler orthogonality criterion. https://terrytao.wordpress.com/2011/11/21/the-bourgain-sarnak-ziegler-orthogonality-criterion/, Nov. 2011. Blog post, What's new.

[31] Tao, T. The Chowla conjecture and the Sarnak conjecture. https://terrytao.wordpress.com/2012/10/14/the-chowla-conjecture-and-the-sarnak-conjecture/, Oct. 2012. Blog post, What's new.