Notes on Sarnak's Conjecture

Aaron Benda

November 2022

1 Seminar Notes

1.1 Background

We are all familiar with the classical Birkhoff/Pointwise Ergodic theorem:

Theorem 1 (Pointwise Ergodic Theorem) *Let $(X, \mathcal{B}, \mu)$ be a probability space and $T : X \to X$ an ergodic endomorphism. Then for any $f \in L^1(X)$, we have that for μ-a.e. $x \in X$,*

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \xrightarrow[N\to\infty]{} \int_X f d\mu.$$

Furthermore, for *uniquely* ergodic systems, the above convergence is actually uniform. A natural question can immediately be posed: are there analogues of this theorem if instead we sample along subsequences of natural numbers, as opposed to them all? In particular, it seems like such a statement should hold for "random" sequences, lacking structure (one could imagine a system with a bias in orbits if there is too much structure on which orders are taken).

A classic example of a sequence for which one would hope this to be true is the primes (perhaps the most captivating object of study in number theory). Indeed, this is a theorem of Bourgain.

Theorem 2 (Bourgain '88) *Under the assumptions of the Pointwise Ergodic Theorem, for any $f \in L^1(X)$, we have that for μ-a.e. $x \in X$,*

$$\frac{1}{\pi(N)}\sum_{p\le N} f(T^p x) \xrightarrow[N\to\infty]{} \int_X f d\mu,$$

where p denotes a prime number.

However, for number theoretic reasons, it's interesting to ask when the above convergence happens for every $x \in X$. Unfortunately, it is not known whether the added assumption of unique ergodicity guarantees this convergence for all points the way it does for the full orbit. In fact, there are only a few known examples of systems for which convergence for all points is known. Systems (X, T) satisfying (a similar kind of statement to) the above are said to satisfy a *prime number theorem (PNT)*.

Remark 0.1 *Prominent known cases of systems with PNTs:*
Vinogradov: Irrational Circle Rotations
Green-Tao (2012): Nilmanifolds
Bourgain (2013): Some 3 IETs

1.2 Sarnak's Conjecture

Sarnak's conjecture is a similar problem to the one presented above, but it has a slightly different flavor and more direct connections/ramifications with open problems in number theory. There is also a general philosophy that there should be some relation regarding which systems satisfy a PNT and which satisfy Sarnak's conjecture. However, to date this is mostly speculative, as no firm relationships have been established (though there are number theoretic reasons to hold this suspicion). First, we introduce the classic Möbius function: define

$$\mu(n) = \begin{cases} 0 & \text{if there is a prime } p \text{ for which } p^2 \mid n \\ (-1)^k & \text{if } n \text{ is a product of } k \text{ distinct primes.} \end{cases}$$

Sarnak's conjecture is the following statement:

Conjecture 1 (Sarnak) *Let $T : X \to X$ be a homeomorphism of a compact metric space, and assume T is of zero topological entropy. Then, for every $f \in C(X)$ and $x \in X$, we have*

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

Systems (X, T) satisfying the above conclusion are said to be *Möbius orthogonal*. First note that the truth of this statement for constant functions f is actually equivalent to the classical prime number theorem.

Theorem 3

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n \leq N} \mu(n) = 0 \Leftrightarrow \pi(x) \sim \frac{x}{\log(x)}.$$

The philosophy behind the conjecture is that primes behave randomly, and factorization (in particular, the number of prime factors) of a given integer n is a seemingly random process too. Since μ only oscillates between -1 and 1 and there should be a roughly even splitting between integers with an even versus odd number of distinct prime factors, one would expect these kinds of sums to tend to 0 in general, so long as there is no bias in the values that f takes dependent on the factorization of the integer n involved in computing the iterate $T^n x$. In particular, the conjecture is that zero entropy systems do not have orbits with enough complexity to somehow correlate with the values of the Möbius function.

1.3 More on μ

As you might be able to tell, the Möbius function is of significant interest for number-theoretic reasons. We have already seen that a statement about its sums is equivalent to the prime number theorem, and in fact there is a direct and interesting relationship with the celebrated Riemann zeta function.

Theorem 4 *Let $\zeta(s)$ denote the Riemann zeta function. Then,*

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}.$$

Furthermore, there is an equivalent form of the Riemann hypothesis that can be stated for the function μ.

Theorem 5 (Littlewood) *For arbitrary $\varepsilon > 0$,*

$$\sum_{n=0}^{N} \mu(n) = O_{\varepsilon}(N^{1/2+\varepsilon}) \text{ as } N \to \infty,$$

holds if and only if the Riemann hypothesis does.

So hopefully you are convinced now that the μ function is of significant importance for the sake of number theory. The motivation for Sarnak's conjecture actually comes primarily from the following purely-number theoretic conjecture due to Chowla.

Conjecture 2 (Chowla, '65) *Fix $r \in \mathbb{N}$, integers $1 \leq a_1 \leq ... \leq a_r$, and exponents $i_0, ..., i_r \in \{1, 2\}$ not all equal to 2. Then*

$$\frac{1}{N} \sum_{n \leq N} \mu^{i_0}(n) \cdot \mu^{i_1}(n + a_1) \cdot ... \cdot \mu^{i_r}(n + a_r) \xrightarrow[N \to \infty]{} 0.$$

This conjecture has remained open (and was thought to be mostly unapproachable, until recently) for over 50 years now. It turns out that Sarnak's conjecture is implied by this conjecture of Chowla. The way this is proved is by giving the Chowla conjecture a purely ergodic-theoretic interpretation, and using the approach of invariant measures some (Furstenberg) disjointness results allow us to see the implication (but this is really technical and I don't want to discuss it). Thus, Sarnak's conjecture is seen as an easier, perhaps more tractable result to show. Stronger versions of Sarnak's conjecture are in fact equivalent to the Chowla conjecture, though not the statement in the form I have given it.

1.4 Dynamical Background

Although Sarnak's conjecture is formulated in terms of topological dynamics, there is an easy way to return this question to one residing in ergodic theory (namely, the following theorem):

Theorem 6 (Krylov-Bogolubov) *Let $T : X \to X$ be a continuous map of a compact metric space. Then there is a T-invariant Borel probability measure μ on X. In particular, there are such measures with respect to which T is ergodic.*

The assumption that the system has zero topological entropy can be translated to the language of ergodic theory as well.

Theorem 7 (Variational Principle) *Let $T : X \to X$ be a continuous map of a compact metric space, and let $M(X,T)$ denote the space of all T-invariant Borel probability measures on X. Then*

$$h_{top}(T) = \sup_{\mu \in M(X,T)} h_\mu(T),$$

where h_{top} is the topological entropy, and h_μ is the Kolmogorov-Sinai entropy.

Therefore, by the variational principle, assuming such a system has zero topological entropy is equivalent to assuming the measure-theoretic entropy is zero with respect to any T-invaraint Borel probability measure.

1.5 Almost Everywhere Version

In fact, the statement of Sarnak's conjecture is known to hold for any system (X,T) for almost every $x \in X$, even independent of the assumption of entropy. We have the following theorem.

Theorem 8 *Let T be an automorphism of a standard Borel probability space $(X, \mathcal{B}, \mu)$, and let $f \in L^1(X, \mathcal{B}, \mu)$. Then, for μ-a.e. $x \in X$, we have*

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

We provide a sketch of the proof here.

The first step is to note that we can assume T is ergodic with respect to μ (a general μ is a convex combination of measures with respect to which T is ergodic). Recall also that the operator $U_T : L^2(X, \mathcal{B}, \mu) \to L^2(X, \mathcal{B}, \mu)$ by $f \mapsto f \circ T$ is unitary; in particular, this means that the eigenvalues of the transformation live in $S^1 \subseteq \mathbb{C}$. Then, for a fixed $f \in L^2(X, \mathcal{B}, \mu)$, by application of the Spectral Theorem we have that

$$\left\| \frac{1}{N} \sum_{n \leq N} f(T^n x)\mu(n) \right\|_2 = \left\| \frac{1}{N} \sum_{n \leq N} z^n \mu(n) \right\|_{L^2(\sigma_f)}$$

where σ_f is the spectral measure of f. But now we can apply Davenport's estimate, which says that for every $A > 0$, there is some C_A for which

$$\max_{z \in S^1} \left| \sum_{n \leq N} z^n \mu(n) \right| \leq C_A \cdot \frac{N}{\log^A(N)}.$$

From this point, the proof basically finishes by applying some density results along with the pointwise ergodic theorem to conclude.

The major hurdle here to overcome is that the argument goes by considering the L^2 norm, which in practice can only yield information about μ-a.e. point $x \in X$. Coming up with an argument to demonstrate this for all points $x \in X$ likely requires completely different techniques.

1.6 Positive Entropy Counterexample

If we assume (X, T) has positive entropy, it is natural to expect that it may not be Möbius orthogonal, due to the complexity of the orbit structure. Indeed, the full shift $(\{-1, 0, 1\}^{\mathbb{N}}, \sigma)$, which has entropy $\log(3)$, is not Möbius orthogonal (e.g. take f to be projection onto the zeroth coordinate, and let $x = (..., \mu(-1), \mu(0), \mu(1), \mu(2), ...)$; we can see that the condition is not satisfied here). More generally, any subshift of finite type will have positive entropy and not be Möbius orthogonal (D. Karagulyan, 2017).

Also, it is worth noting that there are known examples of positive entropy systems which are Möbius orthogonal. Indeed, there are some subshifts of positive entropy which are Möbius orthogonal (Downarowicz and Serafin, 2019).

It goes without saying that at this point there are no known examples of zero entropy systems which are not Möbius orthogonal.

1.7 Known Examples of Möbius Disjoint Systems

Systems of algebraic origin: horocycle flows, nilflows

Systems of measure-theoretic origin: finite rank systems, distal systems

Interval exchange transformations

Systems of number-theoretic origin: automatic sequences (and related).

1.8 Best Results

Here we state some of the best and most promising results towards proving the general case.

Theorem 9 (Matomäki, Raziwiłł, Tao (2015)) *As $H, M \to \infty$, $H \ll M$, we have*

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

Corollary 0.1 *Each system (X, T) for which all invariant measures yield measure-theoretic systems with discrete spectrum is Möbius disjoint.*

Theorem 10 (Tao (2016)) *For each $h \neq 0$, as $N \to \infty$,*

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

Corollary 0.2 *Each system (X, T) for which all invariant measures yield measure-theoretic systems with singular spectrum are logarithmically Möbius disjoint.*

Theorem 11 (Frantzinakis and Host (2019)) *Every zero entropy system (X, T) with only countably many ergodic measures is logarithmically Möbius disjoint.*