Correlations of multiplicative functions, without logarithmic averaging

(joint work with Terence Tao)

Joni Teräväinen

University of Turku

Contents

Introduction

- Multiplicative functions
- Chowla's conjecture
- Applications of correlations
- Results on logarithmic correlations
- Applications of logarithmic correlations
- Transferring from logarithmic averages to ordinary ones?

Multiplicative functions

A function $g : \mathbb{N} \to \mathbb{C}$ is said to be multiplicative if $g(mn) = g(m)g(n)$ whenever m and n are coprime.
We will mostly restrict attention to multiplicative functions taking values in $\mathbb{D} := \{z \in \mathbb{C} : \ |z| \leq 1\}$.
Some key examples:

- The Liouville function $\lambda(n) := (-1)^{\Omega(n)}$, where $\Omega(n)$ is the number of prime factors of n (with multiplicities).
- Generalized Liouville functions $g(n) := e^{2\pi i \alpha \Omega(n)}$, where $\alpha \in \mathbb{R}$.
- Indicator function of smooth numbers: $g(n) := 1_{P^+(n) \leq y}$, where $P^+(n)$ is the largest prime factor of n.

Chowla's conjecture

The Liouville sequence $\lambda(n)$ looks like a random sequence of signs ± 1. In particular, $\lambda(n)$ and $\lambda(n+1)$ should not interact with each other. This was made precise by Chowla in the 1960s.

Chowla's conjecture

Let $k \geq 1$, and let $h_1, \ldots, h_k$ be distinct shifts. Then

$$\frac{1}{x}\sum_{n \leq x} \lambda(n+h_1)\cdots\lambda(n+h_k) = o(1).$$

Equivalently, the vector $(\lambda(n), \lambda(n+1), \ldots, \lambda(n+k-1))$ takes every sign pattern in $\{-1,+1\}^k$ with asymptotic probability 2^{-k}.

Chowla's conjecture

Chowla's conjecture is closely connected to the famous Hardy–Littlewood conjecture:

Hardy–Littlewood conjecture

Let $h_1, \ldots, h_k$ be integer shifts. Then we have

$$\frac{1}{x}\sum_{n\le x}\Lambda(n+h_1)\cdots\Lambda(n+h_k) = \mathfrak{S}(h_1,\ldots,h_k) + o(1),$$

where $\mathfrak{S}(h_1,\ldots,h_k)$ is an explicitly computable quantity that is > 0 whenever $(n+h_1)\cdots(n+h_k)$ has no fixed prime factor.

In fact, if we had a very good error term in Chowla's conjecture (say $\ll (\log x)^{-2k-\varepsilon}$), along with some uniformity in the parameters, then the Hardy–Littlewood conjecture would follow.

Applications of correlations

Correlations of multiplicative functions, that is expressions

$$\frac{1}{x}\sum_{n\le x} g_1(n+h_1)\cdots g_k(n+h_k),$$

have been shown to have numerous applications in recent years.

- Tao's resolution of the Erdős discrepancy problem in combinatorics.
- Bounds on the number of sign patterns of $\lambda(n)$ and other multiplicative functions.
- Classification of all multiplicative functions with bounded partial sums (Klurman).
- Rigidity theorems for multiplicative functions (Klurman–Mangerel).
- Partial progress on Sarnak's conjecture in ergodic theory (Frantzikinakis–Host).
- Work on the largest prime factors of consecutive integers (related to a conjecture of Erdős and Pomerance).

Results on logarithmic correlations

In recent years, there has been a burst of research on logarithmically averaged correlations.

- Matomäki–Radziwiłł–Tao (2015): Chowla's conjecture (and a generalization of it) holds on average over the shifts h_i.
- Tao (2015): We have

$$\frac{1}{\log x}\sum_{n\leq x}\frac{g_1(n+h_1)g_2(n+h_2)}{n}=o(1)$$

  whenever one of the multiplicative functions $g_1, g_2 : \mathbb{N} \to \mathbb{D}$ does not "pretend to be" any twisted character $\chi(n)n^{it}$ in a suitable way.
- Tao–T. (2017): We have

$$\frac{1}{\log x}\sum_{n\leq x}\frac{g_1(n+h_1)\cdots g_k(n+h_k)}{n}=o(1)$$

  whenever the product $g_1\cdots g_k$ does not "weakly pretend" to be any character $\chi(n)$ in a suitable sense.

Applications of logarithmic correlations

- Tao–T. (2017): $\lambda(n)$ takes all sign patterns of length 3 with logarithmic density $1/8$.
- T. (2017): The set of n such that $P^+(n) < P^+(n+1)$ has logarithmic density $1/2$.
- Frantzikinakis–Host (2017): If $a(n)$ has linear block complexity, then $a(n)$ satisfies the logarithmic Sarnak conjecture

$$\frac{1}{\log x}\sum_{n\le x}\frac{\lambda(n)a(n)}{n} = o(1).$$

Question.

Can we remove logarithmic weights from these results?

Transferring from logarithmic averages to ordinary ones?

- By partial summation, if $a : \mathbb{N} \to \mathbb{C}$ is bounded, then

$$\frac{1}{x}\sum_{n\leq x} a(n) = o(1) \Longrightarrow \frac{1}{\log x}\sum_{n\leq x} \frac{a(n)}{n} = o(1).$$

Getting from logarithmic averages to ordinary ones can be difficult:

- The logarithmic statement $\frac{1}{\log x}\sum_{n\leq x}\frac{\lambda(n)}{n} = o(1)$ has a short elementary proof (it essentially corresponds to the statement that $\zeta(s)$ has a pole at $s = 1$).
- The ordinary average $\frac{1}{x}\sum_{n\leq x}\lambda(n) = o(1)$ is equivalent to the prime number theorem and contains the deeper information that $\zeta(s)$ is nonzero for $\mathrm{Re}(s) = 1$.

Transferring from logarithmic averages to ordinary ones?

The implication

$$\frac{1}{\log x}\sum_{n\leq x}\frac{a(n)}{n}=o(1)\Longrightarrow\frac{1}{x}\sum_{n\leq x}a(n)=o(1)$$

is even false for some sequences of interest:

- We have $\frac{1}{\log x}\sum_{n\leq x} n^{it}/n = o(1)$ but $\frac{1}{x}\sum_{n\leq x} n^{it}$ diverges (it rotates essentially around a circle of radius $1/|1+it|$).
- If A is the set of integers that start with 1 (in base 10), then $\frac{1}{\log x}\sum_{n\leq x} 1_A(n)/n = \frac{\log 2}{\log 10} + o(1)$, but $\frac{1}{x}\sum_{n\leq x} 1_A(n)$ does not converge.
- If S is the set of $n\in\mathbb{N}$ for which $\pi(n)<\mathrm{Li}(n)$, then Rubinstein and Sarnak showed that the logarithmic density of S is $0.99999973\ldots$, whereas the ordinary density of S fails to exist.

Transferring from logarithmic averages to ordinary ones?

There are however some results about the convergence of logarithmic averages implying convergence of ordinary averages along a subsequence of scales.

- Elementary fact: If $a : \mathbb{N} \to [-1, 1]$ and $\frac{1}{\log x} \sum_{n \leq x} \frac{a(n)}{n} = o(1)$, then $\frac{1}{x_i} \sum_{n \leq x_i} a(n) = o(1)$ for some sequence $x_i \to \infty$.
- Tao (2017): If the logarithmic $2k$-point Chowla conjecture holds, then the k-point Chowla conjecture holds without logarithmic averaging at almost all scales.
- Unfortunately, we do not know the $2k$-point Chowla conjecture for any $k \geq 2$.

Results and applications

- Non-logarithmic correlations
- A structural theorem
- Isotopy formulae
- Applications
- Can we completely remove logarithmic averaging?

Non-logarithmic correlations

Ordinary correlations at almost all scales (Tao–T., 2018)

Let $g_1, \ldots, g_k : \mathbb{N} \to \mathbb{D}$ be multiplicative functions and $h_1, \ldots, h_k$ integer shifts. Suppose one of the following holds:
(i) $k \geq 1$ and the product $g_1 \cdots g_k$ does not weakly pretend to be any twisted character $\chi(n)n^{it}$;
(ii) $k = 2$ and one of g_1 and g_2 does not pretend to be any $\chi(n)n^{it}$ uniformly for $|t| \leq x$.
Then we have

$$\frac{1}{x}\sum_{n\leq x} g_1(n+h_1)\cdots g_k(n+h_k) = o(1) \quad \text{for } x \in \mathcal{X}$$

with $\mathcal{X} \subset \mathbb{N}$ some set of logarithmic density 1.

Moreover, the set $\mathcal{X}$ even has logarithmic Banach density 1 (if we replace $o(1)$ by $O(\varepsilon)$ above).

Non-logarithmic correlations

As a consequence, we can remove logarithmic averaging from the known results on Chowla's conjecture at almost all scales.

Ordinary Chowla conjecture at almost all scales (Tao–T., 2018)

There exists a set $\mathcal{X} \subset \mathbb{N}$ of logarithmic density 1 such that

$$\frac{1}{x}\sum_{n\leq x}\lambda(n+h_1)\cdots\lambda(n+h_k) = o(1)$$

for all $k \in \{2\} \cup (2\mathbb{N}-1)$ and any distinct $h_1, \ldots, h_k$.

A structural theorem

While proving the above results, we actually show that for any multiplicative functions $g_1, \ldots, g_k : \mathbb{N} \to \mathbb{D}$ and any integers $h_1, \ldots, h_k, a$ we have

$$\lim_{i \to \infty} \frac{1}{x_i/d} \sum_{n \leq x_i/d} g_1(n + ah_1) \cdots g_k(n + ah_k) = f(a)d^{-it}$$

for almost all d, where f is an almost periodic function (depending on x_i), t is a suitable real number, and the sequence x_i is chosen so that the limits in question exist (such sequences can always be produced with a diagonal argument).

This generalizes our earlier result on the structure of logarithmic correlations (there the d parameter played no role).

A structural theorem

For any multiplicative functions $g_1, \ldots, g_k : \mathbb{N} \to \mathbb{D}$ and any integers $h_1, \ldots, h_k, a$ we have

$$\lim_{i\to\infty} \frac{1}{x_i/d} \sum_{n\le x_i/d} g_1(n+ah_1)\cdots g_k(n+ah_k) = f(a)d^{-it} \qquad (1)$$

for almost all d, where f is an almost periodic function (depending on x_i), t is a suitable real number, and the sequence x_i is chosen so that the limits in question exist.

The intuition behind this theorem goes as follows:

- If we had $g_j(n) = \chi(n)$ for some Dirichlet character χ, then one could easily show that (1) holds with $t = 0$ for some periodic f.
- If we had $g_j(n) = n^{it_j}$ for some real numbers t_j, then one could show that (1) holds with $t = t_1 + \cdots + t_k$ for f constant.

Isotopy formulae

Using our structural theorem, we analyzed dependencies between the correlations

$$f_a(x) := \frac{1}{x}\sum_{n\le x} g_1(n+ah_1)\cdots g_k(n+ah_k).$$

Isotopy formulae for correlations (Tao–T., 2018)

Let $g_1,\ldots,g_k : \mathbb{N}\to\mathbb{D}$ be multiplicative such that the product $g_1\cdots g_k$ weakly pretends to be $\chi(n)n^{it}$. There exists a set $\mathcal{X}\subset\mathbb{N}$ of logarithmic density 1 such that

$$f_{-a}(x) = \chi(-1)f_a(x) + o(1) \quad \text{for } x\in\mathcal{X}$$

for all $a\in\mathbb{Z}$ (non-Archimedean isotopy), and

$$f_a(x/q) = q^{-it}f_a(x) + o(1) \quad \text{for } x\in\mathcal{X}$$

for all $q\in\mathbb{Q}$ and $a\in\mathbb{Z}$ (Archimedean isotopy).

Isotopy formulae

The isotopy formulae we proved have some applications as well: The non-Archimedean isotopy formula tells us something about even order correlations of $\lambda(n)$ twisted by $\chi(n)$.

Twists of even order Chowla (Tao–T., 2018)

Let $k \geq 2$ be even. Let $\chi \pmod{k-1}$ be an odd Dirichlet character. Then

$$\frac{1}{\log x} \sum_{n \leq x} \frac{\chi(n)\lambda(n)\lambda(n+1)\cdots\lambda(n+k-1)}{n} = o(1).$$

It is crucial in the proof that we twist by a suitable character $\chi(n)$.

Isotopy formulae

The isotopy formulae we proved have some applications as well: The Archimedean isotopy formula tells that if $g_1 \cdots g_k$ pretends to be $\chi(n)n^{it}$ with $t \neq 0$, the argument of the correlation

$$\arg\left(\frac{1}{x}\sum_{n\leq x} g_1(n+h_1)\cdots g_k(n+h_k)\right)$$

equidistributes on $[0, 2\pi)$ when conditioned on those x where the correlation is $\geq \varepsilon$ in modulus (and after performing some technical smoothing).

Applications

Using our non-logarithmic correlation results, we can almost remove logarithmic averaging from our earlier applications.

Sign patterns of the Liouville function (Tao–T., 2018)

There exists a set $\mathcal{X} \subset \mathbb{N}$ of logarithmic density 1 such that

$$\frac{1}{x}|\{n \leq x : \ \lambda(n+i) = \varepsilon_i \text{ for all } 1 \leq i \leq 3\}| = \frac{1}{8} + o(1)$$

for $x \in \mathcal{X}$ and any $\varepsilon_i \in \{-1, +1\}$.

Largest prime factors of consecutive integers (Tao–T., 2018)

There exists a set $\mathcal{X} \subset \mathbb{N}$ of logarithmic density 1 such that

$$\frac{1}{x}|\{n \leq x : \ P^+(n) < P^+(n+1)\}| = \frac{1}{2} + o(1) \quad \text{for } x \in \mathcal{X}.$$

Can we completely remove logarithmic averaging?

Let $s(K)$ be the number of sign patterns of length K that the Liouville function takes.

- Under Chowla's conjecture, it follows that $S(K) = 2^K$.
- Unconditionally, $S(K)/K \to \infty$ (Frnatzikinakis–Host).

We can show that either $s(K)$ grows rapidly, or we can completely remove logarithmic averaging from two-point and odd order Chowlas.

Sign patterns and correlations (Tao–T., 2018)

Either (i) Number of length K sign patterns of $\lambda(n)$ is $\gg \exp(\frac{cK}{\log K})$ for some $c > 0$ and infinitely many K;
or (ii) Binary and odd order Chowla conjectures hold without logarithmic weights.

Ideas from the proof

- The entropy decrement argument
- Reducing to approximate functional equations
- Approximate homomorphisms

The entropy decrement argument

As in previous work on correlations, we use the entropy decrement argument. For logarithmic correlations, it gives the approximate identity (where $G := g_1 \cdots g_k$)

$$\frac{G(p)}{\log x} \sum_{n \le x} \frac{g_1(n+h_1) \cdots g_k(n+h_k)}{n} = \frac{1}{\log x} \sum_{n \le x} \frac{g_1(n+ph_1) \cdots g_k(n+ph_k)}{n} + o(1)$$

for "most" $p \le \log x$.
However, there is also a variant of this for ordinary averages:

$$\frac{G(p)}{x} \sum_{n \le x/p} g_1(n+h_1) \cdots g_k(n+h_k) = \frac{1}{x} \sum_{n \le x} g_1(n+ph_1) \cdots g_k(n+ph_k) + o(1)$$

for "most" $p \le \log x$.
The entropy decrement is based on some inequalities from information theory, combined with an elaborate pigeonholing argument.

Reducing to approximate functional equations

Let

$$f_d(a) := \lim_{i\to\infty} \frac{1}{x_i/d} \sum_{n\le x_i/d} g_1(n+h_1)\cdots g_k(n+h_k).$$

Then the entropy argument gives us a relation $(G := g_1 \cdots g_k)$

$$f_d(ap) = G(p) f_{dp}(a) + o(1)$$

for "most" primes p. By iterating this, we have

$$f_d(ap_1p_2) = G(p_1)G(p_2)f_{dp_1p_2}(a) + o(1)$$

for "most" primes p_1, p_2.

Reducing to approximate functional equations

We have

$$f_d(ap_1p_2) = G(p_1)G(p_2)f_{dp_1p_2}(a) + o(1)$$

for "most" primes p_1, p_2.
Using some deep results from ergodic theory, as in our previous paper, we can approximate $f_1(a)$ by a nilsequence, $F_{\text{nil}}(a)$, which splits into rational and irrational parts: $F_{\text{nil}}(a) = F_{\text{rat}}(a) + F_{\text{irrat}}(a)$. It turns out that $F_{\text{irrat}}(a)$ cancels out in the bilinear average over p_1, p_2, so we get

$$f_{p_1p_2}(a) = \overline{G}(p_1)\overline{G}(p_2)F_{\text{rat}}(ap_1p_2) + o(1)$$

for "most" primes p_1, p_2.

Approximate homomorphisms

Analyzing the approximate functional equation

$$f_{p_1 p_2}(a) = \overline{G}(p_1)\overline{G}(p_2)F_{\text{rat}}(ap_1 p_2) + o(1)$$

gives us a lot of information on $f_d(a)$. In the end, we can show that $f_d(a) = d^{-it} f(a) + o(1)$ for "most" d and for some almost periodic f. This is our structural theorem, and it eventually leads to our results about correlations at almost all scales.

For showing $f_d(a) = d^{-it} f(a) + o(1)$, a key step is:

Lemma (approximate homomorphisms)

Suppose that $\alpha : (0, \infty) \to \mathbb{C}$ is a function such that $\alpha(xy) = \alpha(x)\alpha(y) + O(\varepsilon)$ for all $x, y > 0$ with $\varepsilon > 0$ small but fixed. Suppose also that $\alpha(x) = O(1)$ and $\alpha(x) = 1 + O(\varepsilon)$ for $|x - 1| \leq \varepsilon$. Then there exists a real t such that $\alpha(x) = x^{it} + O(\varepsilon)$.

This in turn is proved by using some standard tools from functional analysis.