PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 2 (339–362)

MULTIPLICATIVE FUNCTIONS IN SHORT INTERVALS, AND CORRELATIONS OF MULTIPLICATIVE FUNCTIONS

KAISA MATOMÄKI AND MAKSYM RADZIWIŁŁ

Abstract

Our goal in this note is two-fold. In part I, we motivate and explain the ideas behind a recent theorem of ours.

Theorem 1 (Matomäki-Radziwiłł). *Let f be a real-valued multiplicative function with $|f| \leq 1$. Then, for all $X < x \leq 2X$, with at most $o(X)$ exceptions,*

$$\frac{1}{H} \sum_{x < x \leq x+H} f(n) - \frac{1}{X} \sum_{X < n \leq 2X} f(n) = o(1)$$

as soon as $H \to \infty$ with $X \to \infty$.

In part II, which can be read independently, our goal is to survey some of the recent developments connected to Theorem 1. These have been by far and large related to Chowla's conjecture.

Conjecture 1 (Chowla). *Let μ denote the Möbius function. Then, for any set of distinct integers $h_1, \dots, h_k$,*

$$\sum_{n \leq X} \mu(n + h_1) \dots \mu(n + h_k) = o(X)$$

as $X \to \infty$.

Part I

We will be interested throughout in *multiplicative* functions, that is $f : \mathbb{N} \to \mathbb{C}$ such that $f(ab) = f(a)f(b)$ for all co-prime a, b. A basic example is the *Möbius function* μ, defined by $\mu(p) = -1$ and $\mu(p^\alpha) = 0$ for all $\alpha > 1$ and primes p. The Möbius function is closely connected to the primes as we now explain. Let Λ denote the von Mangoldt

MSC2010: primary 11N25; secondary 11N37.

function $\Lambda(n)$ defined by setting $\Lambda(n) = \log p$ whenever $n = p^\alpha$ and $\Lambda(n) = 0$ otherwise. Then, by inclusion-exclusion,

$$\Lambda(n) = -\sum_{d \mid n} \mu(d) \log d.$$

Using this relationship one can show that,

$$\sum_{n \le x} \Lambda(n) \sim x \iff \sum_{n \le x} \mu(n) = o(x) \tag{1}$$

and

$$\sum_{n \le x} \Lambda(n) = x + O_\varepsilon(x^{1/2+\varepsilon}) \quad \text{for all } \varepsilon > 0 \iff \sum_{n \le x} \mu(n) = O_\varepsilon(x^{1/2+\varepsilon}) \quad \text{for all } \varepsilon > 0. \tag{2}$$

The statement on the left-hand side of (1) is known as the Prime number theorem, while the statement on the left of (2) is an equivalent reformulation of the Riemann Hypothesis.

An immediate consequence of the Riemann Hypothesis is that the prime number theorem holds in all intervals $[x, x + x^\alpha]$ with $\alpha > \frac{1}{2}$. In principle deeper information than the Riemann Hypothesis is contained in the *explicit formula*,

$$\sum_{n \le x} \Lambda(n) = x - \sum_{\rho} \frac{x^\rho}{\rho} + O(1)$$

where the sum over ρ corresponds to a sum over zeros of the Riemann zeta-function. However from this formula it is possible to see that without understanding cancellations between zeros of the Riemann zeta-function nothing can be said about intervals with $\alpha \le \frac{1}{2}$. For $\alpha \in (0, \frac{1}{2}]$ it is thus natural to relax the question and ask whether the prime number theorem holds in most intervals $[x, x + x^\alpha]$ with $\alpha \in (0, \frac{1}{2}]$. Conditionally on the Riemann Hypothesis, Selberg [1943] succeeded in providing a positive answer to this question. Here we state a weak version of his result.

Theorem 2 (Selberg)**.** *Assume the Riemann Hypothesis. Let $\varepsilon > 0$ be given. The number of integers $x \in [0, X]$ for which*

$$\Big| \sum_{x < n \le x + x^\alpha} \Lambda(n) - x^\alpha \Big| > x^{\alpha/2+\varepsilon}$$

is $o(X)$ as $X \to \infty$.

The proof of this result is not difficult : Covering the interval $[0, X]$ by dy-adic intervals, the result will follow from Chebyschev's inequality if we can show that,

$$\int_X^{2X} \Big| \sum_{x<n\le x+x^\alpha} \Lambda(n) - x^\alpha \Big|^2 dx \ll_\varepsilon X^{1+\alpha+\varepsilon}.$$

By the explicit formula, for $X \le x \le 2X$,

$$\sum_{x<n\le x+x^\alpha} \Lambda(n) - x^\alpha \approx x^\alpha \sum_{|\rho|\le X^{1-\alpha}} x^{\rho-1}.$$

where ρ is a sum over the zeros of the Riemann zeta-function, and here and later $\approx$ means that the statement is "morally true". Therefore,

$$\int_X^{2X} \Big| \sum_{x<n\le x+x^\alpha} \Lambda(n) - x^\alpha \Big|^2 dx \approx X^{2\alpha} \int_X^{2X} \Big| \sum_{|\rho|\le X^{1-\alpha}} x^{\rho-1} \Big|^2 dx \tag{3}$$

Expanding the square, and executing the integration over $x \in [X, 2X]$ we obtain that the integral over x is

$$\approx \sum_{\substack{|\rho|\le X^{1-\alpha} \\ |\rho'|\le X^{1-\alpha}}} \frac{X^{\rho+\overline{\rho'}-1}}{\rho+\overline{\rho'}-1}$$

We then bound this trivially by

$$\sum_{\substack{|\rho|\le X^{1-\alpha} \\ |\rho'|\le X^{1-\alpha}}} \frac{1}{|\rho+\overline{\rho}-1|} \ll X^{1-\alpha+\varepsilon}. \tag{4}$$

Altogether this shows that the left-hand side of (3) is $\ll_\varepsilon X^{1+\alpha+\varepsilon}$, as needed. A similar but more complicated argument establishes a corresponding theorem for the Möbius function.

The Riemann Hypothesis is used crucially in the upper bound (4) which uses that $\Re\rho = \Re\rho' = \frac{1}{2}$. In order to run this argument unconditionally one needs a *zero-density estimate* for the number of zeros of the Riemann zeta function in the strip $\Re s > \sigma$ and with height $|\Im s| \le T$. Using Huxley's zero-density estimate Huxley [1972] allows one to prove the following theorem (see Ramachandra [1976] for details).

Theorem 3 (Huxley). *Let $a(n) = \Lambda(n) - 1$ or $a(n) = \mu(n)$. Then, for $H > X^{1/6+\varepsilon}$, we have, for almost all $X < x < 2X$,*

$$\Big| \sum_{x<n\le x+H} a(n) \Big| \ll_A H(\log X)^{-A}.$$

Two features are worth noticing : Compared to the conditional Theorem 2, the saving that we obtain is weaker, and the range of H is worse. Crucial in the proof of Theorem 3 is the relationship of $\Lambda(n)$ or $\mu(n)$ with the zeros of the Riemann zeta-function.

Until recently this is where things stood. In a recent result we have obtained an improvement of Theorem 3 for arbitrary multiplicative function which is optimal in terms of H. We will begin with a very special case of our result for the Möbius function in the range $H = X^{\varepsilon}$.

Theorem 4 (Matomäki-Radziwiłł). *Let $\varepsilon > 0$ be given. Let $H = X^{\varepsilon}$. Then, for almost all $1 \leq x \leq X$,*

$$\sum_{x<n\leq x+H} \mu(n) = o(H).$$

Unlike Huxley's result our theorem depends directly on the multiplicativity of $\mu(\cdot)$. While the proof that we will give at first will depend on the fact that $\mu(p) = -1$ for all primes p, we will soon see that this is in no way crucial. A complete account of this special case can be found in Matomäki and Radziwiłł [2016a].

0.1 Sketch of the proof of Theorem 4. Consider

$$\int_X^{2X} \Big| \sum_{x<n\leq x+H} \mu(n) \Big|^2 dx \tag{5}$$

Our goal is to show that this is $o(XH^2)$. By an application of Plancherel, (5) is (essentially) equivalent to

$$\int_0^{X/H} \Big| \sum_{X<n\leq 2X} \frac{\mu(n)}{n^{1+it}} \Big|^2 dt = o(1). \tag{6}$$

We have at our disposal two distributional estimates for

$$\sum_{X<n\leq 2X} \frac{\mu(n)}{n^{1+it}}.$$

On the one hand the prime number theorem in the form of Vinogradov-Korobov Korobov [1958] implies that the above Dirichet polynomial is less than $O_A((\log X)^{-A})$ for all $|t| \leq \exp((\log X)^{3/2-\varepsilon})$. On the other hand a result of Montgomery and Vaughan [1974] shows that for arbitrary complex coefficients $a(n)$ and $T, X \geq 1$,

$$\int_0^T \Big| \sum_{X<n\leq 2X} a(n) n^{it} \Big|^2 dt = (T + O(X)) \sum_{X<n\leq 2X} |a(n)|^2. \tag{7}$$

Neither result is directly sufficient for obtaining (6). The first result allows us to only handle the range $|t| \leq (\log X)^B$ for any fixed B. This is not sufficient to obtain the result (unless $H > X/(\log X)^B$). However, in any case it allows to reduce our attention to showing that

$$\int_{(\log X)^B}^{X/H} \Big| \sum_{X \leq n \leq 2X} \frac{\mu(n)}{n^{1+it}} \Big|^2 dt = o(1). \tag{8}$$

for any fixed $B > 0$. Applying (7) to (8) shows that (8) is $O(1)$. This barely fails to be non-trivial. The situation is reminiscent with what one encounters in the proof of the Bombieri-Vinogradov theorem Vaughan [1981] [1]. Similarly, the missing additional input is a bilinear structure.

We create the bilinear structure using *Ramaré's identity*,

$$\mu(n) = \sum_{\substack{n = pm \\ P \leq p \leq Q \\ (m,p)=1}} \mu(p) \cdot \frac{\mu(m)}{1 + \#\{P \leq q \leq Q : q | m\}} + \mathbf{1}_{(n, \prod_{P \leq p \leq Q} p) = 1} \cdot \mu(n) \tag{9}$$

valid for any interval $[P, Q]$. If the parameters P, Q are chosen so that

$$\sum_{P \leq p \leq Q} \frac{1}{p} \to \infty$$

as $X \to \infty$, then all but $o(X)$ of the integers $X < n \leq 2X$ have a prime factor in $[P, Q]$. Consequently the second term in Ramaré's identity (9) is typically zero. The point of Ramaré's identity is that it roughly allows us to write

$$\sum_{X < n \leq 2X} \frac{\mu(n)}{n^{1+it}} \approx$$

$$\approx \frac{1}{\log \frac{\log Q}{\log P}} \sum_{P \leq R = 2^k \leq Q} \Big(\sum_{R < p \leq 2R} \frac{\mu(p)}{p^{1+it}} \cdot \sum_{X/R < m \leq 2X/R} \frac{\mu(m)}{m^{1+it}} \Big) + \text{“small in } L^2\text{”}.$$

[1] If one attempts to prove Bombieri-Vinogradov without a combinatorial decomposition of $\Lambda(n)$ then the available tools turn out to be insufficient: Siegel-Walfisz can handle moduli with $Q \leq (\log X)^A$ while the large sieve recovers the trivial bound. This is similar to the situation we are facing here.

where "small in L^2" is a Dirichlet polynomial that is small on average. As a result, we roughly have

$$\int_{(\log X)^B}^{X/H} \Big| \sum_{X<n\le 2X} \frac{\mu(n)}{n^{1+it}} \Big|^2 dt$$
$$\text{“}\ll\text{”} \frac{\log Q}{\log R} \sum_{P \le R = 2^k \le Q} \int_{(\log X)^B}^{X/H} \Big| \sum_{R<p\le 2R} \frac{1}{p^{1+it}} \cdot \sum_{X/R<m\le 2X/R} \frac{\mu(m)}{m^{1+it}} \Big|^2 dt + o(1).$$

Now we have a bilinear decomposition and we can use both distributional estimates alluded to above. First of all, if the parameter P is chosen so that $P \ge \exp((\log X)^{2/3+\varepsilon})$, then the prime number theorem of Vinogradov-Korobov applies to the Dirichlet polynomial over the primes, and gives us

$$\sum_{R<p\le 2R} \frac{1}{p^{1+it}} \ll (\log X)^{-B/2}$$

for all $R = 2^k$ in the range $[P, Q]$ and all $|t| > (\log X)^B$. Secondly, (7) shows that,

$$\int_{(\log X)^B}^{X/H} \Big| \sum_{X/R<m\le 2X/R} \frac{\mu(m)}{m^{1+it}} \Big|^2 dt \ll \Big(\frac{X}{H} + \frac{X}{R} \Big) \sum_{X/R<m\le 2X/R} \frac{1}{m^2} \ll \frac{R}{H} + 1.$$

Applying both to (8), we conclude that (8) is

$$\ll \sum_{P\le R=2^k\le Q} (\log X)^{-B} \cdot \Big(\frac{R}{H} + 1 \Big) + o(1).$$

This is $o(1)$ as long $R/H \le (\log X)^{B/2}$. We collect all our requirements to see if they can be satisfied at once. We require that

1. $P \ge \exp((\log X)^{2/3+\varepsilon})$ (so that the prime number theorem is applicable).
2. $P/H \ll (\log X)^{B/2}$ (so that the mean-value theorem is efficient).
3. Q is chosen so that $\sum_{P\le p\le Q} p^{-1} \to \infty$ as $X \to \infty$ (so that in Ramaré's identity the second term is negligible).

To meet all these requirements it suffices to pick $Q = H \ge X^{\varepsilon}$ and $P = \exp((\log X)^{2/3+\varepsilon})$.

0.2 General multiplicative functions. The main thrust of the previous argument still came from a harmonious relationship between μ and prime numbers (manifested for example in the property that $\mu(p) = -1$ for all primes p). If one wishes to extend the result to general multiplicative functions, this is a bottleneck. But more generally, before we can proceed we need to understand what is the analogue of the prime number theorem for general multiplicative functions.

Since the prime number theorem can be expressed as

$$\sum_{n \leq x} \mu(n) = o(x)$$

it is natural to expect that "having a prime number theorem" for $f(n)$ would amount to knowing the behavior of

$$\sum_{n \leq x} f(n).$$

as $x \to \infty$. There is at present a rather well-developed theory of such mean-values. The central result is due to Halász (see also Montgomery and Vaughan [2001] and Granville, Harper, and Soundararajan [2017]).

Theorem 5 (Halász)**.** *Let f be multiplicative with $|f| \leq 1$. Then, for all $X, T \geq 1$,*

$$\frac{1}{X} \sum_{X < m \leq 2X} f(m) \ll M \exp(-M) + \frac{\log \log X}{\log X} + \frac{1}{T}$$

where

$$M := \min_{|t| \leq T} \sum_{p \leq X} \frac{1 - \Re f(p) p^{-it}}{p}$$

The quantity M singles out cases in which $f(p)$ is close to p^{it} for most primes p. This is indeed important. If $f = m^{iT}$ then the mean-value is of size approximately X/T showing that the last term is optimal. The other two terms are unfortunately also optimal. For instance the second term takes into account multiplicative functions f such that for example $f(p) = 0$ for $p < X/2$ and $f(p) = 1$ for $X/2 \leq p \leq X$. In recent literature the function M is frequently denoted by $\mathbb{D}^2(f, p^{it})$ and refered to as a distance function.

Going back to our main problem, we would like to show that for almost all $X \leq x \leq 2X$,

$$\frac{1}{H} \sum_{x < n \leq x + H} f(n) - \frac{1}{X} \sum_{X < n \leq 2X} f(n) = o(1). \tag{10}$$

and for H as small as possible, compared to X. Since we aim to prove this for an arbitrary multiplicative function f with $|f| \leq 1$ one should think of (10) as a statement about the factorization of the integers, rather than about multiplicative functions. Indeed, (10) states that in most short intervals $[x, x+H]$ integers factorize in a way that is similar to the long interval $[X, 2X]$. Thus in the way that integers factorize in short intervals, there can be no consistently pathological behavior that would work out to still be consistent with, for example, the prime number theorem. An example of such a pathological behavior is the existence of a positive proportion of intervals $[x, x+H]$ on which for instance $\mu(n) = 1$ and a positive proportion of intervals $[x, x+H]$ on which $\mu(n) = -1$. The intervals could be arranged in such a way so that their existence would be consistent with the prime number theorem $\sum_{n \leq x} \mu(n) = o(x)$. What (10) achieves is that it rules out such a possibility.

Let us try to prove (10). Expressing both expressions in (10) in terms of Mellin transform and applying Plancherel, reveals that showing (10) amounts to proving that

$$\int_{(\log X)^{\varepsilon}}^{X/H} \Big| \sum_{X < n \leq 2X} \frac{f(n)}{n^{1+it}} \Big|^2 dt = o(1).$$

Notice that the integral starts at $(\log X)^{\varepsilon}$, this is important and corresponds to the fact that the Mellin transform of the short sum and the long sum, coincide at the small frequencies $|t| \leq (\log X)^{\varepsilon}$.

We now run the same argument as before. The crucial issue is to show that,

(11)
$$\int_{(\log X)^{\varepsilon}}^{X/H} \Big| \sum_{R < p \leq 2R} \frac{f(p)}{p^{1+it}} \Big|^2 \cdot \Big| \sum_{X/R < m \leq 2X/R} \frac{f(m)}{m^{1+it}} \Big|^2 dt = O\Big((\log X)^{-\varepsilon/2} \cdot \Big(\frac{\log Q}{\log P} \Big)^{-2} \Big)$$

for some $\varepsilon > 0$. One could get away with a smaller saving. Notice that the coefficients $f(p)$ are arbitrary, so there are no point-wise bounds for $\sum_{R < p \leq 2R} f(p) p^{-1-it}$. In particular the prime number theorem is of no use!

However the sum over primes is still small for "most" t. Indeed by (7),

$$\int_{0}^{X/H} \Big| \sum_{R < p \leq 2R} \frac{f(p)}{p^{1+it}} \Big|^2 dt \sim \frac{X}{H} \cdot \frac{1}{R \log R}.$$

This shows that for "most" t the sum over the primes is of size $R^{-1/2+o(1)}$. Thus for most $t \in [(\log X)^{\varepsilon}, X/H]$ our previous argument works. As a result we can focus on the set $\mathcal{U}$ of those $t \in [(\log X)^{\varepsilon}, X/H]$ for which,

$$\Big| \sum_{R < p \leq 2R} \frac{1}{p^{1+it}} \Big|^2 \geq (\log X)^{-\varepsilon} \sum_{R < p \leq 2R} \frac{1}{p}.$$

We can find a 1-spaced set $\mathcal{V} \subset U$ such that

$$\int_{t\in\mathcal{U}} \Big| \sum_{R<p\le 2R} \frac{f(p)}{p^{1+it}} \cdot \sum_{X/R<m\le 2X/R} \frac{f(m)}{m^{1+it}} \Big|^2 dt \tag{12}$$
$$\ll \sum_{t\in\mathcal{V}} \Big| \sum_{R<p\le 2R} \frac{f(p)}{p^{1+it}} \Big|^2 \cdot \Big| \sum_{X/R<m\le 2X/R} \frac{f(m)}{m^{1+it}} \Big|^2$$

A first task is to understand the cardinality of $\mathcal{V}$. If $Q < \exp((\log X)^{1-2\varepsilon})$ then one can compute moments

$$\int_0^{X/H} \Big| \sum_{R<p\le 2R} \frac{f(p)}{p^{1+it}} \Big|^{2k} dt$$

with a judiciously chosen k, to conclude that $|\mathcal{V}|$ is very small.

On the set $t \in \mathcal{V} \subset [(\log X)^{\varepsilon}, X/H]$ we cannot expect cancellations in the sum over the primes. In fact if there exists a single t for which there are no cancellations in both sums over p and m then we cannot succeed. Thankfully, for real-valued multiplicative functions Halász's theorem ensures that the sum over m is always non-trivially small. In fact we can hope for a saving of a small power of the logarithm in the sum over m. This leads to a bound for (12) of the form,

$$(\log X)^{-\varepsilon} \sum_{t\in\mathcal{V}} \Big| \sum_{R<p\le 2R} \frac{f(p)}{p^{1+it}} \Big|^2 \tag{13}$$

A variant of a distributional result of Halász and Turán [1969] (see Matomäki and Radziwiłł [2016b, Lemma 11]) then shows that once the set $\mathcal{V}$ is small enough, the sum over $t \in \mathcal{V}$ behaves as if there was exactly one term $t \in \mathcal{V}$ at which there are no cancellations. Therefore (13) is

$$\ll (\log X)^{-\varepsilon} \cdot \Big(\frac{1}{\log R}\Big)^2 \ll (\log X)^{-\varepsilon} \cdot (\log Q/\log P)^{-2}$$

provided that P, Q are choosen so that $P \ge \exp(\sqrt{\log Q})$, which we can assume. For instance, simply choose $Q = \exp((\log X)^{1-2\varepsilon})$ and $P = \exp((\log X)^{2/3+\varepsilon})$. This gives (11) and proves

Corollary 1 (Matomäki-Radziwiłł). *Let $f : \mathbb{N} \to [-1, 1]$ be multiplicative. Let $\varepsilon > 0$ be given and set $H = X^{\varepsilon}$. Then for almost all $X < x \le 2X$,*

$$\frac{1}{H} \sum_{x<n\le x+H} f(n) - \frac{1}{X} \sum_{X<n\le 2X} f(n) = o(1)$$

as $X \to \infty$.

0.3 The full result. In our joint work Matomäki and Radziwiłł [2016b] we succeeded in pushing the dependence on H to its optimal form. In particular we obtained the following result.

Theorem 6 (Matomäki-Radziwiłł). *Let $|f| \leq 1$ be multiplicative and real-valued. Let $\delta > 0$ be fixed. Then,*

$$\Big|\frac{1}{H}\sum_{x<n\leq x+H} f(n) - \frac{1}{X}\sum_{X<n\leq 2X} f(n)\Big| < \delta.$$

for all $X < x \leq 2X$ with $\ll X/H^{\delta/30} + X/(\log X)^{1/50}$ exceptions.

This implies that if H goes to infinity with X, no matter how slowly, then,

$$\frac{1}{H}\sum_{x<n\leq x+H} f(n) = \frac{1}{X}\sum_{X\leq n\leq 2X} f(n) + o(1) \tag{14}$$

for almost all $X \leq x \leq 2X$. This is optimal. If H were bounded then (14) is not true: for instance take $H = p+1$ with p prime and, and f equal to a quadratic character $\mod p$. Then the short sum is always equal to $1/H$ in absolute value, while the long sum is tiny. In any case when H is bounded the statement is conjectured to be false for all multiplicative functions f. However such counterexamples are open even for $f = \mu$.

It is not a matter of simple technique to go from $H = X^{\varepsilon}$ to $H \to \infty$. This can be most clearly seen in the fact that previously even on the assumption of the Riemann Hypothesis a result like (6) for $f = \mu$ was only known for $H > (\log X)^A$ for some large A.

There is in fact a very conceptual reason for this. For simplicity let's focus on the case of the Möbius function. After converting the problem to Dirichlet polynomials, for the method to work we have to create bilinear forms with Dirichlet polynomials over primes p in the range of H (or smaller, but for simplicity let's focus on p of size H). Then we have to understand the size of

$$\sum_{H\leq p\leq 2H} \frac{1}{p^{1+it}}$$

Ideally one would like to say that this Dirichlet polynomial is always small. The Riemann Hypothesis guarantees this to be the case for $H > (\log X)^{2+\varepsilon}$. When $(\log X)^{1+\varepsilon} < H < (\log X)^{2+\varepsilon}$ there is a gap in our knowledge: we still expect the Dirichlet polynomial to be always small, but even the Riemann Hypothesis is unable to confirm this. Finally the range $H < (\log X)^{1-\varepsilon}$ is particularly difficult : Using diophantine approximation one can show that there are arbitrarily large $X < t < 2X$, such that,

$$\Big|\sum_{H\leq p\leq 2H} \frac{1}{p^{1+it}}\Big| \geq \frac{1}{4}\sum_{H\leq p\leq H} \frac{1}{p}$$

Moreover the number of such t up to X is expected to be a small power of X. Therefore any argument that works in scales $H < (\log X)^{1-\varepsilon}$ will need to be able to exploit this feature to its advantage. This is not something that arguments in analytic number theory are designed to address!

The main new idea in the proof of Theorem 6 is an iterative scheme, factoring out from the Dirichlet polynomial

$$\sum_{X \le n \le 2X} \frac{f(n)}{n^{1+it}}$$

Dirichlet polynomials supported on the primes in various ranges. The argument is designed to react to the size of the Dirichlet polynomial in each range. If the Dirichlet polynomial exhibits cancellations we are done. If it does not we move to a subsequent range, but retain the information that the Dirichlet polynomial in the previous range was large. Without doing this we would not be able to succeed. For the reader interested in these details we refer to an exposition of Soundararajan [2017] or our original paper Matomäki and Radziwiłł [2016b].

Part II

While analytic number theorists have by now a coherent set of tools to tackle problems about mean-values,

$$\sum_{n \le X} a(n) \tag{15}$$

with $a(n)$ sequences of arithmetical interest, very little is known about *correlations*,

$$\sum_{n \le X} a(n)a(n+h) \tag{16}$$

with $h \neq 0$. To get a sense of the gap in the difficulty set $a(n) = \Lambda(n)$. Then (15) corresponds to the prime number theorem, while (16) is the Hardy-Littlewood 2-tuple conjecture.

If one writes,

$$\Lambda(n) = \sum_{n=ab} \mu(a) \log b \tag{17}$$

then the problem of estimating (16) reduces to that of understanding correlations of the Möbius function. For a few technical reasons we will be interested instead in correlations of the Liouville function, which differs from the Möbius function only on powers of primes. This makes in practice the two interchangeable. The Liouville function $\lambda(n)$ is defined as $\lambda(n) = (-1)^{\Omega(n)}$ where $\Omega(n)$ is the number of prime factors of n counted with multiplicity.

For our approach to succeed, we need to at the very least understand (16) with $a(n) = \mu(n)$ or $a(n) = \lambda(n)$. A conjecture of Chowla [1965] predicts that such sums always exhibit cancellations.

Conjecture 2 (Chowla)**.** *Let $a(n) = \mu(n)$ or $a(n) = \lambda(n)$. Then, for any distinct set of integers $h_1, \dots, h_\ell$,*

$$\sum_{n \le x} a(n + h_1) \dots a(n + h_\ell) = o(x).$$

as $x \to \infty$.

Instead of (17) one could use Linnik's identity,

$$\frac{\Lambda(n)}{\log n} = \sum_{k \ge 1} \frac{d_k^\star(n)}{k} \cdot (-1)^{k+1}$$

where $d_k^\star(n)$ counts the number of solutions to $n = n_1 \dots n_k$ with all $n_i > 1$. Then estimating (16) requires us to understand correlations of the kth divisor function.

Conjecture 3. *Let $h_1, \dots, h_\ell$ be a set of distinct integers and $k_1, \dots, k_\ell \ge 1$ integers. Then,*

$$\sum_{n \le x} d_{k_1}(n + h_1) \dots d_{k_\ell}(n + h_\ell) \sim C(\mathbf{k}, \mathbf{h}) \cdot x (\log x)^{k_1 + \dots + k_\ell - \ell}.$$

as $x \to \infty$, with $C(\mathbf{k}, \mathbf{h})$ a complicated constant depending on the tuples $\mathbf{k} = (k_1, \dots, k_\ell)$ and $\mathbf{h} = (h_1, \dots, h_\ell)$.

Unfortunately for us both conjectures are open. At first the second conjecture appears somewhat more approachable. For instance the problem of estimating,

$$\sum_n d(n) d(n + h) W\Big(\frac{n}{X}\Big) \tag{18}$$

with $W(\cdot)$ a smooth function is completely resolved. Following works of Kuznetsov we are able to write down an explicit formula for (18), with the error term involving L-functions associated to eigenfunctions of the hyperbolic Laplacian on $\mathrm{SL}_2(\mathbb{Z}) \backslash \mathbb{H}$ ((see Motohashi [1994] for details). The triangle inequality shows that the error term is of size $O(X^{1/2+\varepsilon})$, and moreover this is optimal. One would hope that a similar strategy will work for the estimation of correlations of the third divisor function, and that the error term should involve objects related to $\mathrm{SL}_3(\mathbb{Z})$. So far such attempts have proven unsuccessful Vinogradov and Tahtadžjan [1978], and not even an asymptotic formula is known the mean-value of $d_3(n) d_3(n + 1)$.

0.4 Logarithmic Chowla. However it is no longer clear that Conjecture 2 will be the first to fall. At present our understanding of such convolution sums has changed dramatically. For instance we highlight a recent result of Tao [2016b].

Theorem 7 (Tao). *We have,*

$$\sum_{n \le x} \frac{\lambda(n)\lambda(n+1)}{n} = o(\log x) \tag{19}$$

as $x \to \infty$.

This is refered to as *logarithmic Chowla*.

To see the relevance of Theorem 6 let us start from the very modest goal of obtaining any cancellations at all in correlations of the Liouville function. So let us suppose that

$$\sum_{n \sim X} \lambda(n)\lambda(n+1) = (1 + o(1))X$$

as $X \to \infty$, and see if we can disprove it. The above would mean that for most n, the sign of $\lambda(n)$ and $\lambda(n+1)$ is equal. Therefore there exists an H going to infinity very slowly, so that,

$$\Big| \sum_{x < n < x+H} \lambda(n) \Big| \sim H$$

in at least a positive proportion of the intervals $[x, x+H]$. Theorem 6 rules out such a possibility. In fact a quick consequence of Theorem 6 (see Matomäki and Radziwiłł [2016b, Corollary 2]) is that there exists $\delta > 0$ such that,

$$\Big| \sum_{n \le X} \lambda(n)\lambda(n+1) \Big| \le (1-\delta)X$$

this resolved an old folklore conjecture (see for instance Hildebrand [n.d.]) and opened the door for further progress on Conjecture 1.

The next natural step is to establish Chowla's conjecture "on average". In joint work with Tao Matomäki, Radziwiłł, and Tao [2015] we obtained such a result.

Theorem 8 (Matomäki-Radziwiłł-Tao). *We have,*

$$\sum_{|h| \le H} \Big| \sum_{n \le X} \lambda(n)\lambda(n+h) \Big| = o(HX). \tag{20}$$

as soon as $H \to \infty$ *with* $X \to \infty$.

This was the crucial arithmetic ingredient in Theorem 7. Let us quickly sketch the ideas that go into the proof of Theorem 8. The identity

$$\int_0^1 \Big(\int_{\mathbb{R}} \Big| \sum_{\substack{x \le n < x+H \\ X \le n \le 2X}} \lambda(n) e(n\alpha) \Big|^2 \Big)^2 d\alpha = H \sum_{|h| \le H} (H - |h|) \Big| \sum_{n \le X} \lambda(n)\lambda(n+h) \Big|^2$$

shows that Theorem 8 will follow from

$$\sup_{\alpha \in \mathbb{R}} \sum_{X \le x \le 2X} \Big| \sum_{x \le n \le x+H} \lambda(n) e(n\alpha) \Big| = o(XH) \tag{21}$$

as $X \to \infty$ and $H \to \infty$ with X, arbitrarily slowly. This is a short-interval analogue of a classical result of Davenport (itself a variant of Vinogradov's result for primes),

$$\sup_{\alpha} \Big| \sum_{X \le n \le 2X} \lambda(n) e(n\alpha) \Big| = o(X).$$

Similarly to the proof of Davenport's theorem, the proof of (21) splits depending on the diophantine nature of α. For simplicity let us imagine that α is fixed and we are aiming at obtaining cancellations in

$$\sum_{x \le n \le x+H} \lambda(n) e(n\alpha)$$

in almost all short intervals $X \le x \le 2X$ as $X \to \infty$, and $H \to \infty$ with X. The proof splits into two cases depending on whether $\alpha \in \mathbb{Q}$ or $\alpha \notin \mathbb{Q}$. When $\alpha \notin \mathbb{Q}$ the phase $e(n\alpha)$ oscilates rather randomly, and we succeed by using ideas of Daboussi and Delange [1974], which are a variant of Vinogradov's method. This requires $f(n)$ to be multiplicative only in a certain range $[P, Q]$. [2] On the other hand when $\alpha \in \mathbb{Q}$ the phase $e(n\alpha)$ is predictable and we need to obtain cancellations from $\lambda(n)$. The most extreme case corresponds to $\alpha = 0$. However this is exactly a consequence of Theorem 6! As one might expect the proof for rational α follows by generalizing Theorem 6 to the case of arithmetic progressions. Note that in this case we use the multiplicativity of $f(n)$ in many intervals, and thus this case is significantly more arithmetic.

The estimate (21) is the crucial arithmetic input in the proof of logarithmic Chowla (19). The other input is an ingenious use of entropy allowing to replace the event $\lambda(n+p)\mathbf{1}_{p|n}$ by $\lambda(n+p)/p$, and the creation of a bilinear structure which is possible thanks to the

[2] Daboussi-Delange's work is also often refered to as the Bourgain-Katai-Sarnak-Ziegler criterion (after Bourgain, Sarnak, and Ziegler [2013], Kátai [1986]), however the correct attribution is to Daboussi-Delange who were the first to obtain such a result

logarithmic weights. In fact rather than directly attacking the Chowla conjecture, Tao shows cancellations in

$$\sum_{n \le X} \lambda(n) \Big(\sum_{\substack{p \mid n \\ P \le p \le 2P}} \lambda(n+p) \Big). \tag{22}$$

A feature of the logarithmic weights is that if one can show that (22) is non-trivially small, for any choice of P, then this implies (19). Tao's *entropy decrement argument* shows the existence of a range P on which one can replace the event $\lambda(n+p)\mathbf{1}_{p \mid n}$ by $\lambda(n+p)/p$. Effectively this shows that (22) is close to

$$\frac{1}{P} \sum_{P \le p \le 2P} \sum_{n \le X} \lambda(n)\lambda(n+p)$$

This is now a ternary problem, that we can hope to address by the circle method. The main ingredient here is (21): The range of P can be quite small with respect to X, and one needs a form of (21) with H of size about P.

0.4.1 Odds and ends. Of course one wonders if logarithmic Chowla has implications for prime numbers. Unfortunately while until recently this was the heuristic expectations of most experts, it turns out that the error terms in (19) are too weak to give back anything about prime numbers. This appears to be a conceptual obstruction rather than a purely technical one. The proof of (19) hinges on Theorem 6 which in turn, in the first step, discards prime numbers from consideration. Thus it seems that these methods cannot be used for a direct attack on the twin prime conjecture.

Nonetheless the logarithmic Chowla conjecture had some dramatic consequence *outside* of number theory. It led for instance to the resolution of the Erdős Discrepancy Problem Tao [2016a] in combinatorics.

Concerning the cases of logarithmic Chowla with more shifts, there has been recent progress due to Tao-Teräväinen Tao and Teräväinen [2017]. It turns out that the case of an odd number of shifts is considerably simpler, and does not require any short interval results. That this is reasonable can be perhaps most easily seen in a result of Elliott according to which, there exists a $\delta > 0$ such that,

$$\Big| \sum_{n \le x} \lambda(n)\lambda(n+1)\lambda(n+2) \Big| \le (1-\delta)X$$

The proof of the above inequality is completely elementary (see Cassaigne, Ferenczi, Mauduit, Rivat, and Sárközy [1999]). This is in stark contrast with the binary case that requires Theorem 6.

0.5 Sarnak's conjecture. One can think of Chowla's conjecture probabilistically, as asserting that if one picks a typical integer n, and we are given $\lambda(n), \lambda(n+1), \ldots, \lambda(n+k-1)$ then we get no discernible advantage in predicting $\lambda(n+k)$.

Such a point of view is appealing from the point of view of information theory: We can think of the "signal" $\lambda = (\lambda(1), \lambda(2), \lambda(3), \ldots)$, and then ask how redundant this signal is? More precisely is knowing the neighborhood of a point $\lambda(n)$ enough to reconstruct $\lambda(n)$ at least with some positive probability? If the sequence $\lambda(n)$ is truly random, then the answer should be no. In particular $\lambda(n)$ should be "orthogonal" to all sequences that have a lot of redundancies, i.e those of entropy 0.

Let us then define what we mean by the entropy of a sequence. Let $f : \mathbb{N} \to \mathbb{C}$ be an arbitrary sequence, and consider the set $S_m = \{(f(n), f(n+1), \ldots, f(n+m-1)) : n > 0\} \subset \mathbb{C}^m$ of m-tuples in $\mathbb{C}^m$. Given ε let $B(m, \varepsilon)$ be the number of m-dimensional balls in $\mathbb{C}^m$ of radius ε that are needed to cover S_m. Then, the *topological entropy of f* is defined as

$$\sigma = \sup_{\varepsilon > 0} \Big(\limsup_{m \to \infty} \frac{1}{m} \log B(m, \varepsilon) \Big).$$

Conjecture 4 (Sarnak). *Let $f : \mathbb{N} \to \mathbb{C}$, have topological entropy zero. Then,*

$$\sum_{n \leq x} \lambda(n) f(n) = o(x).$$

Sarnak's conjecture is a natural generalization of Davenport's theorem. The latter corresponds to the case of $f(n) = e(\alpha n)$ which clearly has entropy 0 (simply cover the m-dimensional unit circle by a finite union of ε balls). Sarnak's conjecture also appears naturally in additive combinatorics. A fundamental step in the proof of the Green-Tao theorem is the orthogonality of the Möbius function to nilsystems Green and Tao [2012], this corresponds to a special case of Sarnak's conjecture.

As one might expect there is a tight link between the conjectures of Chowla and Sarnak. Chowla's conjecture implies Sarnak's conjecture. Moreover if one considers the logarithmic versions of the two conjectures, then they are equivalent Tao [2017] This leads one to believe that Chowla's conjecture and Sarnak's conjecture are equivalent, but this is so far unproven.

There is a large body of literature concerning Sarnak's conjecture (see Ferenczi, Kułaga-Przymus, and Lemańczyk [2017]). There are currently two main tools : When the the sequence f is sufficiently random, one uses a criterion steeming from the work of Daboussi-Delange, also known as the Bourgain-Katai-Sarnak-Ziegler criterion. This says that if, for all fixed primes p, q,

$$\sum_{n \leq x} f(pn) \overline{f(qn)} = o(X), \tag{23}$$

then Sarnak's conjecture holds. Verifying the above condition in practice might require a substantial amount of work depending on the f under consideration. There are also cases for which (23) fails but Sarnak conjecture is expected to hold nonetheless (the simplest example being an f that changes randomly signs on blocks that grow to infinity very slowly). In such a situation one frequently tries to find a way to use Theorem 6. In fact Theorem 6 can be rephrased in purely ergodic terms.

Due to the large number of works on the subject we cannot cover it in great depth, but refer to the recent survey Ferenczi, Kułaga-Przymus, and Lemańczyk [ibid.] for the state of the art.

0.6 The shfited convolution problem on average and L-functions. When the sequence $a(n)$ is of an *automorphic* origin, then the problem of estimating the correlations,

$$\sum_{n \le x} a(n)a(n+h)$$

is often-times refered to as the *shifted convolution problem*. The reason for the different terminology is that the problem occurs frequently and naturally when trying to either estimate moments of L-functions or obtain subconvex bounds.

The prototypical moment problem is the problem of estimating,

$$M_k(T) := \int_T^{2T} |\zeta(\tfrac{1}{2} + it)|^{2k} dt.$$

as $T \to \infty$ where $\zeta(s)$ is the Riemann zeta-function. It is conjectured that $M_k(T) \ll T^{1+\varepsilon}$ for all $k > 0$, and even more precisely that,

$$M_k(T) = TP_k(\log T) + O(T^{1-\delta_k}) \tag{24}$$

with P_k a polynomial of degree k^2 and $\delta_k > 0$ a positive exponent. This is a powerful conjecture that implies the bound $|\zeta(\frac{1}{2} + it)| \ll_\varepsilon 1 + |t|^\varepsilon$. Such a bound (known as the Lindelöf hypothesis) would be a great substitute for the Riemann Hypothesis in many applications.

Unfortunately (24) is known only for $k = 1$ and $k = 2$. While the case of $k = 1$ can be treated with harmonic analysis alone, the estimation for $k = 2$ depends crucially on having a power-saving in the shifted convolution problem. In fact one finds that,

$$\begin{aligned} \int_T^{2T} |\zeta(\tfrac{1}{2} + it)|^{2k} dt &\approx \\ &\approx TP_k(\log T) + \text{“}\frac{1}{T^{k/2-1}} \sum_{|h| \sim T^{k/2-1+\varepsilon}} e^{ih/T^{k/2-1}} \sum_{n \sim T^{k/2}} d_k(n)d_k(n+h)\text{”} \end{aligned} \tag{25}$$

where the meaning of $\sim$ is left vague on purpose. Notice that the error term is trivial when $k = 1$, and is barely non-trivial for $k = 2$. Once $k \geq 3$ the error term is unfortunately too big to be manageable by any known method, and $k > 4$ seems to present extraordinary challenges as even square-root cancellation is no longer sufficient (but see B. Conrey and Keating [2015] for possible ways to circumvent this). Many results on moments focus exactly on the threshold where the sum in the error term is barely non-trivial.

0.6.1 The shifted convolution problem for $d_k(n)$. The shape of the error term in (25) motivates us to understand,

$$\sum_{n \leq X} d_k(n) d_k(n+h)$$

on average over h, or more generally with $d_k(n)$ replaced by coefficients of a $GL(k)$ automorphic form. The literature contains average results when the number of shifts $|h| \leq H$ is substantial, i.e $H > X^{1/3+\varepsilon}$ (and recently for $H > X^{8/33+\varepsilon}$). However using methods related to Theorem 6 one can reduce the number of shifts to be as small as $(\log X)^{O(k \log k)}$.

Theorem 9 (Matomäki-Radziwiłł-Tao)**.** *Let $k \geq \ell \geq 2$ be real numbers. Then, for $H = (\log X)^{10000 k \log k}$,*

$$\sum_{|h| \leq H} \Big| \sum_{n \leq X} d_k(n) d_\ell(n+h) - X P_{k+\ell}(\log X) \Big| = o(HX (\log X)^{k+\ell-2})$$

where $P_{k+\ell}$ is a polynomial of degree $k + \ell - 2$.

Compared to our earlier work multiplicative functions such as $d_k(n)$ present new challenges because they are unbounded. But it is not the fact that they are unbounded in itself which is the main difficulty – it's rather the fact that because they are unbounded, the main contribution to

$$\sum_{n \leq X} d_k(n) d_k(n+h)$$

comes from a thin subset of integers having $(k + o(1)) \log \log x$ prime factors and on which $d_k(n)$ is unusually large. The density of such integers is about $(\log X)^{-k \log k + k - 1}$.

The idea behind the proof of Theorem 9 is to first restrict to a density one subset of integers $\mathfrak{N}$ on which we can construct efficient sieve majorants for $d_k(n)$. Subsequently after some harmonic analysis, we find that the crucial issue is to obtain a non-trivial estimate for

$$\sum_{j=1}^{J} \sum_{X \leq x \leq 2X} \Big| \sum_{\substack{x < n < x+H \\ n \in \mathfrak{N}}} d_k(n) e(n \alpha_j) \Big|$$

where α_j is a set of well-spaced points and J is arbitrary. When J is large one can simply appeal to the large sieve. On the other hand when J is bounded obtaining a non-trivial result amounts to a variant of Theorem 6 for unbounded multiplicative functions (similar variants for other sparse sets of integers were studied in Goudout [2017] and Teräväinen [2016]). It is the intermediate range between bounded and large J which turns out the most subtle. In this range through a use of duality we (essentially) replace $d_k(n)$ by the corresponding sieve majorant $\widetilde{d}_k(n)$, and we are reduced to needing cancellations in

$$\sum_{x<n<x+H} \widetilde{d}_k(n)e(n\alpha_j)$$

after a substantial amount of effort (this would be trivial if $H > X^{\varepsilon}$). We then estimate high moments of such a sum to conclude that it exhibits cancellations for most x.

In fact the argument works quite generally for estimating $d_k(n)b(n+h)$ with $b(n)$ any sequence for which efficient sieve majorants can be constructed. For instance, we obtain results for the higher-order Titchmarsh divisor problem,

$$\sum_{n\leq X} d_k(n)\Lambda(n+h)$$

with an average over $|h| \leq H$ of size at most $H = (\log X)^{10000k\log k}$. For individual h the problem is open for all $k > 2$.

0.6.2 Moments of L-functions. Coming back to the moment problem, our results (such as Theorem 9) say nothing new in the case of the Riemann zeta-function, but they are nonetheless useful in other families. For instance, a q-analogue of moments of the Riemann zeta-function, is the problem of estimating,

$$\sum_{q\leq Q}\sum_{\chi \pmod q}\int_{\mathbb{R}} |L(\tfrac{1}{2}+it,\chi)|^{2k}\cdot\Phi(t)dt \tag{26}$$

where χ is a sum over primitive characters, $L(s,\chi)$ is a Dirichlet L-function and $\Phi(t)$ is a fixed smooth function. This problem (but without the t averaging) is also related to understanding the distribution of $d_k(n)$ in arithmetic progressions.

The large sieve gives sharp upper bound for the above moment problem when $k \leq 4$. A few years ago, J. B. Conrey, Iwaniec, and Soundararajan [2013] devised a method to obtain asymptotic estimates in moments such as (26). They illustrated their method to obtain an asymptotic for (26) when $k = 3$. The case $k = 4$ represents the absolute limit of their method, and also the limit of what should be realistically feasible. In Chandee and Li [2014] it was addressed conditionally on the Generalized Riemann Hypothesis.

Roughly for the solution of the case $k = 4$ one needs to estimate on average a shifted convolution problem in a short interval. It turns out that this problem is amenable to our earlier methods and leads to the following Theorem (the write-up is currently in preparation).

Theorem 10 (Chandee-Li-Matomäki-Radziwiłł). *Let $\Phi(t)$ be a fixed smooth function. Then,*

$$\sum_{q \le Q} \sum_{\chi \pmod q} \int_{\mathbb{R}} |L(\tfrac{1}{2} + it, \chi)|^8 \cdot \Phi(t) dt \sim C \widehat{\Phi}(0) \cdot Q^2 (\log Q)^{16}.$$

as $Q \to \infty$, where the summation is over primitive characters and $C > 0$ is an absolute constant.

Previously the assumption of the Generalized Riemann Hypothesis was used to estimate non-trivially the shifted convolution problem that arises in this problem.

0.6.3 Gaps between multiplicative sequences. The applications of the shifted convolution problem are not restricted to problems related in one way or another to L-functions. A prominent example is Hooley's Hooley [1971] work on gaps between sums of two squares. Let $1 = s_1 < s_2 < \dots$ be the sequence of integers representable as sums of two squares. Then the average gap between $s_{n+1} - s_n$ for $s_n \le x$ is $\asymp \sqrt{\log x}$. Hooley investigated how often the gaps deviate from the mean. He proved that for $\gamma < 5/3$,

$$\sum_{s_n \le x} (s_{n+1} - s_n)^{\gamma} \asymp x (\log x)^{\frac{1}{2}(\gamma - 1)}. \tag{27}$$

The form of the expression is motivated by Erdős's conjecture that,

$$\sum_{p_n \le x} (p_{n+1} - p_n)^2 \asymp x \log x.$$

In (27) the lower bound is easy and follows from Holder's inequality against the case $\gamma = 1$. In principle (27) is conjectured to hold for all finite γ, but this is a very deep conjecture. It implies for instance that for any fixed $\varepsilon > 0$ in all intervals of the form $[x, x + x^{\varepsilon}]$ there is a sum of two squares.

The problem of showing (27) is equivalent to obtaining the following frequency bound,

$$\#\Big\{x \le X : \sum_{x < s_n \le x + H\sqrt{\log X}} 1 = 0\Big\} \ll \frac{X}{H^{\gamma - 1}} \tag{28}$$

uniformly in $1 \le H \le X$. It is the uniformity that is difficult to maintain in this problem.

Hooley's approach goes as follows: In order to obtain (28) he notices that for any weight w_n supported on sums of two squares, and for any constant A, the frequency (28) is

$$\ll \frac{1}{A^2} \sum_{x \le X} \Big| \sum_{x \le n \le x+H} w_n - A \Big|^2 \tag{29}$$

Moreover the estimation of (29) is feasible if we can get a good estimate for the correlations

$$\sum_{|h| \le H} \sum_{n \le X} w_n \overline{w_{n+h}}$$

that saves at least a power of H over the trivial bound.

When H is small, Hooley selects $w_n = r(n)\rho(n)$ where $r(n)$ is the number of representations of n as a sums of two squares, and $\rho(n)$ is a sieve weight dampening the size of $r(n)$ so that $r(n)\rho(n) \approx 1$ on average. When H is large (say $H > X^{\varepsilon}$) we can afford to loose powers of the logarithm, and it's enough to choose $w_n = r(n)$. Then we require a shifted convolution on average, with a power-saving. This is possible to obtain because $r(n)$ is roughly similar to the divisor function, and similar techniques that can be used to estimate,

$$\sum_{n \le x} d(n) d(n+h)$$

will also do the job for the $r(n)$ function.

The sequence of sums of two squares is a norm-form, since $a^2 + b^2$ is the norm of Gaussian integers. It is natural to ask if Hooley's result can be extended to norm forms of higher degree fields. Unfortunately we run right away in a serious difficulty : If K is a number field of degree 3 and $r_K(n)$ is the number of representations of n as a norm of an ideal in K, then there are no results with power-savings for the shifted convolution problem,

$$\sum_{|h| \le X^{\delta}} \sum_{n \le X} r_K(n) r_K(n+h)$$

In fact this is of a comparable difficulty as our Conjecture 2 in the case $\ell = 2$ and $k_1 = k_2 = 3$.

It turns out however that we can circumvent these difficulties by using techniques related to Theorem 6. The advantage of using these "more restrictive techniques" (after all we forfeit any possibility of power-savings) is that not only we can extend the result to norm-forms, but more generally to any "multiplicative sequence" (of which norm-forms are an example).

Corollary 2 (Matomäki-Radziwiłł). *Let $\mathfrak{P}$ be a set of primes of positive density δ. Let $\mathfrak{N}$ be the set of all square-free integers all of whose prime factors belong to $\mathfrak{P}$. Denote the elements of $\mathfrak{N}$ by $n_1 < n_2 < \dots$. Then, for all $\gamma < \frac{3}{2}$,*

$$\sum_{n_k \leq x} (n_{k+1} - n_k)^{\gamma} \asymp_{\mathfrak{P},\gamma} x(\log x)^{(1-\delta)(\gamma-1)}. \tag{30}$$

We think it is remarkable that the exponent $\frac{3}{2}$ does not depend on the density of $\mathfrak{P}$. Let us very quickly explain the kind of ideas that go into the proof of Corollary (2). The case of H small can be disposed by proving a variant of our Theorem 6 for multiplicative functions that are supported on a set of primes of density $0 < \delta < 1$.

When H is large, the techniques that go into Theorem 6 are not immediately applicable, and need to be modified. Let us highlight the spirit of these modifications in the case of norm forms. What makes the shifted convolution problem (30) difficult are certain specific sets of integers, for instance the integers n that factor into abc with a, b, c roughly of equal size. However if one restricts in the shifted convolution problem to integers of the form $n = abc$ with a, b, c in a certain special configuration then the problem can be solved with a power-saving in X. Unfortunately the set of integers in such a desirable configuration might be of density zero, but this is not a problem when H is large!

The reason is that in the regime H large we can afford to lose some powers of the logarithm, and by restricting to a density zero subset (instead of the full sequence) we are typically loosing at most powers of the logarithm. So we simply run Hooley's method on the subsequence of integers representable as norm forms and that factor as $n = abc$ with abc in certain desirable configurations. In reality the proof of Corollary 2 goes along completely different lines, however what we explained highlights the spirit of the ideas.

0.7 The structure of multiplicative functions. Our Theorem 6 is of course rather immediately applicable to the study of general multiplicative functions. For instance it immediately implies the following result.

Corollary 3. *Let $f : \mathbb{N} \to \mathbb{R}$ be a multiplicative function which is non-zero for positive proportion of natural numbers. Then f has a positive proportion of sign changes if and only if $f(n) < 0$ for some $n \in \mathbb{N}$.*

This was an improvement even in the case of $f(n) = \mu(n)$. This result has been the starting point of several "rigidity theorems" of Klurman [2017] and Klurman and Mangerel [2017]. We highlight a few of their results. For instance, in Klurman [2017] an old conjecture of Katai is resolved. This asserts that for f taking values on the unit disk,

$$\sum_{n \leq x} |f(n+1) - f(n)| = o(x) \tag{31}$$

if and only if, $f(n) = n^{it}$ for some $t \in \mathbb{R}$, or

$$\sum_{n \le x} |f(n)| = o(x).$$

Notice that if f is assumed to be real-valued and lying on the unit-disk (i.e $f(n) = \pm 1$) then Katai's conjecture is nothing more than a convoluted restatement of Corollary 3.

Moreover in Klurman and Mangerel [2017] appears a solution to an old conjecture of Chudakov. The conjecture states that if f is a multiplicative function such that $f(n)$ takes only finitely many values, and $f(p)$ is zero on only finitely many primes, and

$$\sum_{n \le x} f(n) = \alpha x + O(1)$$

then f is a Dirichlet character. In a similar vein one can think of the Erdős Discrepancy Problem as a rigidity theorem, since it implies that there are no completely multiplicative functions f with $\sum_{n \le x} f(n) = O(1)$ for all x.

References

J. Bourgain, P. Sarnak, and T. Ziegler (2013). "Disjointness of Moebius from horocycle flows". In: *From Fourier analysis and number theory to Radon transforms and geometry*. Vol. 28. Dev. Math. Springer, New York, pp. 67–83. MR: 2986954 (cit. on p. 352).

Julien Cassaigne, Sébastien Ferenczi, Christian Mauduit, Jöel Rivat, and András Sárközy (1999). "On finite pseudorandom binary sequences. III. The Liouville function. I". *Acta Arith.* 87.4, pp. 367–390. MR: 1671629 (cit. on p. 353).

Vorrapan Chandee and Xiannan Li (2014). "The eighth moment of Dirichlet L-functions". *Adv. Math.* 259, pp. 339–375. MR: 3197660 (cit. on p. 357).

S. Chowla (1965). *The Riemann hypothesis and Hilbert's tenth problem*. Mathematics and Its Applications, Vol. 4. Gordon and Breach Science Publishers, New York-London-Paris, pp. xv+119. MR: 0177943 (cit. on p. 350).

Brian Conrey and Jonathan P. Keating (2015). "Moments of zeta and correlations of divisor-sums: I". *Philos. Trans. Roy. Soc. A* 373.2040, pp. 20140313, 11. MR: 3338122 (cit. on p. 356).

J. Brian Conrey, Henryk Iwaniec, and Kannan Soundararajan (2013). "Critical zeros of Dirichlet L-functions". *J. Reine Angew. Math.* 681, pp. 175–198. MR: 3181494 (cit. on p. 357).

Hédi Daboussi and Hubert Delange (1974). "Quelques propriétés des fonctions multiplicatives de module au plus égal à 1". *C. R. Acad. Sci. Paris Sér. A* 278, pp. 657–660. MR: 0332702 (cit. on p. 352).

S. Ferenczi, J. Kułaga-Przymus, and M. Lemańczyk (Oct. 2017). "Sarnak's Conjecture – what's new". arXiv: 1710.04039 (cit. on pp. 354, 355).

Élie Goudout (2017). "Lois locales de la fonction ω dans presque tous les petits intervalles". *Proc. Lond. Math. Soc. (3)* 115.3, pp. 599–637. MR: 3694294 (cit. on p. 357).

Andrew Granville, Adam J Harper, and K. Soundararajan (June 2017). "A new proof of Halász's Theorem, and its consequences". arXiv: 1706.03749 (cit. on p. 345).

Ben Green and Terence Tao (2012). "The Möbius function is strongly orthogonal to nilsequences". *Ann. of Math. (2)* 175.2, pp. 541–566. MR: 2877066 (cit. on p. 354).

G. Halász and P. Turán (1969). "On the distribution of roots of Riemann zeta and allied functions. I". *J. Number Theory* 1, pp. 121–137. MR: 0236124 (cit. on p. 347).

Adolf Hildebrand (n.d.). review. no. 95d:11099 (cit. on p. 351).

Christopher Hooley (1971). "On the intervals between numbers that are sums of two squares". *Acta Math.* 127, pp. 279–297. MR: 0294281 (cit. on p. 358).

M. N. Huxley (1972). "On the difference between consecutive primes". *Invent. Math.* 15, pp. 164–170. MR: 0292774 (cit. on p. 341).

I. Kátai (1986). "A remark on a theorem of H. Daboussi". *Acta Math. Hungar.* 47.1-2, pp. 223–225. MR: 836415 (cit. on p. 352).

Oleksiy Klurman (2017). "Correlations of multiplicative functions and applications". *Compos. Math.* 153.8, pp. 1622–1657. MR: 3705270 (cit. on p. 360).

Oleksiy Klurman and Alexander P. Mangerel (July 2017). "Rigidity Theorems for Multiplicative Functions". arXiv: 1707.07817 (cit. on pp. 360, 361).

N. M. Korobov (1958). "Estimates of trigonometric sums and their applications". *Uspehi Mat. Nauk* 13.4 (82), pp. 185–192. MR: 0106205 (cit. on p. 342).

Kaisa Matomäki and Maksym Radziwiłł (2016a). "Multiplicative functions in short intervals". *Ann. of Math. (2)* 183.3, pp. 1015–1056. arXiv: 1502.02374. MR: 3488742 (cit. on p. 342).

– (2016b). "Multiplicative functions in short intervals". *Ann. of Math. (2)* 183.3, pp. 1015–1056. MR: 3488742 (cit. on pp. 347–349, 351).

Kaisa Matomäki, Maksym Radziwiłł, and Terence Tao (2015). "An averaged form of Chowla's conjecture". *Algebra Number Theory* 9.9, pp. 2167–2196. MR: 3435814 (cit. on p. 351).

H. L. Montgomery and R. C. Vaughan (1974). "Hilbert's inequality". *J. London Math. Soc. (2)* 8, pp. 73–82. MR: 0337775 (cit. on p. 342).

– (2001). "Mean values of multiplicative functions". *Period. Math. Hungar.* 43.1-2, pp. 199–214. MR: 1830577 (cit. on p. 345).

Yōichi Motohashi (1994). "The binary additive divisor problem". *Ann. Sci. École Norm. Sup. (4)* 27.5, pp. 529–572. MR: 1296556 (cit. on p. 350).

K. Ramachandra (1976). "Some problems of analytic number theory". *Acta Arith.* 31.4, pp. 313–324. MR: 0424723 (cit. on p. 341).

Atle Selberg (1943). "On the normal density of primes in small intervals, and the difference between consecutive primes". *Arch. Math. Naturvid.* 47.6, pp. 87–105. MR: 0012624 (cit. on p. 340).

Kannan Soundararajan (2017). "The Liouville function in short intervals". *Astérisque* 390. Séminaire Bourbaki. Vol. 2015/2016. Exposés 1104–1119, Exp. No. 1119, 453–479. MR: 3666035 (cit. on p. 349).

Terence Tao (2016a). "The Erdős discrepancy problem". *Discrete Anal.* Paper No. 1, 29. MR: 3533300 (cit. on p. 353).

– (2016b). "The logarithmically averaged Chowla and Elliott conjectures for two-point correlations". *Forum Math. Pi* 4, e8, 36. MR: 3569059 (cit. on p. 351).

– (2017). "Equivalence of the logarithmically averaged Chowla and Sarnak conjectures". In: *Number theory—Diophantine problems, uniform distribution and applications*. Springer, Cham, pp. 391–421. MR: 3676413 (cit. on p. 354).

Terence Tao and Joni Teräväinen (Aug. 2017). "The structure of logarithmically averaged correlations of multiplicative functions, with applications to the Chowla and Elliott conjectures". arXiv: 1708.02610 (cit. on p. 353).

Joni Teräväinen (2016). "Almost primes in almost all short intervals". *Math. Proc. Cambridge Philos. Soc.* 161.2, pp. 247–281. MR: 3530507 (cit. on p. 357).

R. C. Vaughan (1981). "An elementary method in prime number theory". In: *Recent progress in analytic number theory, Vol. 1 (Durham, 1979)*. Academic Press, London-New York, pp. 341–348. MR: 637357 (cit. on p. 343).

A. I. Vinogradov and L. A. Tahtadžjan (1978). "Theory of the Eisenstein series for the group $SL(3, \mathbf{R})$ and its application to a binary problem. I. Fourier expansion of the highest Eisenstein series". *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* 76. Analytic number theory and the theory of functions, pp. 5–52, 216. MR: 527787 (cit. on p. 350).

Received 2017-12-09.

Kaisa Matomäki
ksmato@utu.fi

Maksym Radziwiłł
maksym.radziwill@mcgill.ca