Analytic Number Theory: the Additive Perspective

Introduction to the Field of Research

Cédric Pilatte

Contents

1. Introduction

Analytic number theory may be defined as the branch of mathematics that uses analytical tools to tackle number-theoretical questions, such as solving an equation in the integers or studying the distribution of prime numbers. The diversity of the methods used in this area makes it a rich field that is highly interconnected with many other mathematical disciplines.

One of the most celebrated results in analytic number theory – and probably in the whole of mathematics – is the prime number theorem, which states that

$$|\{p \leq x : p \text{ prime}\}| \sim \frac{x}{\log x}$$

as $x \to \infty$ (here, $\log x$ always denotes the natural logarithm). It was proved independently by Jacques Hadamard and Charles Jean de la Vallée Poussin in 1896.

Many facts about prime numbers can be restated in terms of the Liouville function $\lambda : \mathbb{N} \to \{\pm 1\}$ defined by

$$\lambda(n) := (-1)^{\Omega(n)},$$

where $\Omega(n)$ is the number of prime factors of n, counted with multiplicity. For example, it can be proven in an elementary way that the prime number theorem is equivalent to

$$\frac{1}{x}\sum_{n \leq x} \lambda(n) \to 0$$

as $x \to \infty$ [14, p. 55]. In other words, roughly half of the integers $\leq x$ have an even number of prime factors, and roughly half have an odd number of prime factors.

Another example is given by the famous Riemann hypothesis. We give two of its many equivalent formulations below. Define $\mathrm{Li}(x) := \int_2^x \frac{1}{\log t}\,dt$, so that

$$\mathrm{Li}(x) = \frac{x}{\log x} + \frac{1!\,x}{(\log x)^2} + \frac{2!\,x}{(\log x)^3} + \ldots + \frac{(k-1)!\,x}{(\log x)^k} + O\left(\frac{x}{(\log x)^{k+1}}\right)$$

for any fixed $k \geq 1$. Then, the Riemann hypothesis is equivalent to the very strong error term

$$|\{p \leq x : p \text{ prime}\}| = \mathrm{Li}(x) + O\left(\sqrt{x}\log x\right) \tag{1}$$

in the prime number theorem. Of course, the Riemann hypothesis is still unsolved, and the best known result in that direction is the bound

$$|\{p \leq x : p \text{ prime}\}| = \mathrm{Li}(x) + O\left(\frac{x}{\exp((\log x)^{3/5+o(1)})}\right) \tag{2}$$

by Vinogradov and Korobov, which has an equivalent counterpart on averages of the Liouville function. Alternatively, the Riemann hypothesis is the statement that, for every $\varepsilon > 0$,

$$\sum_{n \leq x} \lambda(n) = O\left(x^{1/2+\varepsilon}\right).$$

The heuristic idea is that the Liouville function λ, whose first values are

$$+1, -1, -1, +1, -1, +1, -1, -1, +1, +1, -1, -1, -1, +1, +1, +1, -1, -1, -1, -1, +1, +1, -1, \ldots$$

should behave statistically like a random function from $\mathbb{N}$ to $\{-1, +1\}$. Thus, for a large random x, the values of the normalized sum $\frac{1}{\sqrt{x}}\sum_{n \leq x} \lambda(n)$ should approximately follow a standard normal distribution. By a "large random x", we mean that we fix some large X and choose x independently at random in the interval $(X, 2X]$. Then $\frac{1}{\sqrt{x}}\sum_{n \leq x} \lambda(n)$ is a random variable and we can study its distribution as $X \to \infty$.

According to this heuristic principle, we should also expect consecutive values of the Liouville function to be "independent", or at least "uncorrelated". If x is a large and n is an integer chosen uniformly at random in the interval $[1, x]$, then $\lambda(n)$ and $\lambda(n+1)$ are random variables with mean (approximately) zero and covariance (approximately)

$$\sum_{n \leq x} \lambda(n)\lambda(n+1).$$

Here, the trivial bound is $\left|\sum_{n \leq x} \lambda(n)\lambda(n+1)\right| \leq x$, and the Chowla conjecture asserts that this sum should be $o(x)$.

Conjecture 1.1 (Chowla's conjecture in degree 2)**.** *For any integer $h > 0$,*

$$\frac{1}{x}\sum_{n \leq x} \lambda(n)\lambda(n+h) \to 0$$

as $x \to \infty$.

This conjecture is still open. Some experts believe Conjecture 1.1 to be of the same order of difficulty as the twin prime conjecture, which states that there are infinitely many pairs of primes $(p, p+2)$. Indeed, a suitable quantitative refinement of Conjecture 1.1 would imply the twin prime conjecture.[1]

The general Chowla conjecture applies to an arbitrary number of shifted Liouville functions.

Conjecture 1.2 (Chowla's conjecture). *For any fixed $k \geq 1$ and nonnegative integers $h_1 < \ldots < h_k$,*

$$\frac{1}{x}\sum_{n\leq x}\lambda(n+h_1)\cdots\lambda(n+h_k)\to 0$$

as $x \to \infty$.

In Section 2, we present some of the methods used in analytic number theory, in order to give the reader a flavour of what this field is all about. We return to the Chowla conjecture in Section 3.

2. A Variety of Techniques

There is a myriad of commonly used techniques in analytic number theory. It would be an illusory task to make an exhaustive list of them, and we will only attempt to describe a few of the most important ones.

2.1. **Multiplicative Number Theory.** In the middle of the nineteenth century, Riemann realized that complex analysis could be used to study the distribution of prime numbers. Let us briefly explain this connection.

Define the Riemann zeta function

$$\zeta(s)=\sum_{n\geq 1}\frac{1}{n^s}$$

for $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$. More than a hundred years before Riemann, Euler had discovered (for real s) what is now called the Euler product formula

$$\zeta(s)=\prod_p(1-p^{-s})^{-1},$$

where the product is over all primes. The formula is obtained by expanding each factor on the right-hand side as a geometric series, and then multiplying out the infinite product. Taking a complex logarithm of both sides and derivating, this formula implies

$$\frac{\zeta'(s)}{\zeta(s)}=\sum_{n\geq 1}\frac{\Lambda(n)}{n^s},$$

where Λ is the von Mangoldt function

$$\Lambda(n)=\begin{cases}\log p & \text{if } n \text{ is a power of the prime } p,\\ 0 & \text{otherwise.}\end{cases}$$

[1]See Tao's blog https://terrytao.wordpress.com/2015/03/17/an-averaged-form-of-chowlas-conjecture.

The von Mangoldt function acts as a weighted indicator function of the primes (perfect powers p^k, $k \geq 2$, are very sparse among the integers and can be ignored for most purposes).

It is then possible to obtain information on the statistical behaviour of Λ through analytic properties of the Riemann zeta function. One (sophisticated) way to do so is via the Wiener-Ikehara tauberian theorem.

Theorem 2.1 (Wiener-Ikehara theorem). *Let $a(n)$ be non-negative real numbers, and suppose that the series*

$$\sum_{n\geq 1} \frac{a(n)}{n^s} \tag{3}$$

has a meromorphic extension to an open set containing $\{s \in \mathbb{C} : \mathrm{Re}(s) \geq b\}$, except possibly at $s = b$ where it has a simple pole with residue c. Then

$$\sum_{n\leq x} a(n) \sim \frac{c}{b} x^b$$

as $x \to \infty$.

Series as in (3) are known as *Dirichlet series*. Applying this result to the logarithmic derivative ζ'/ζ of the Riemann zeta function, we obtain the asymptotic

$$\sum_{n\leq x} \Lambda(n) \sim x.$$

After removing the negligible contribution of powers of primes p^k, $k \geq 2$, this becomes

$$\sum_{p\leq x} \log p \sim x,$$

where, as usual, the sum is over all primes at most x. Using the technique of summation by parts, it is easy to replace the weight $\log p$ by 1, and obtain

$$\sum_{p\leq x} 1 \sim \frac{x}{\log x},$$

which is the prime number theorem.

However, the core of the proof lies in proving that the hypotheses of Theorem 2.1 are satisfied, namely that

(1) ζ has a meromorphic extension to the left of the line $\mathrm{Re}(s) = 1$, and;
(2) ζ'/ζ has no pole on $\mathrm{Re}(s) \geq 1$ except at $s = 1$.

The first property is not too difficult to show; in fact, one can prove that ζ extends meromorphically to the whole complex plane. For property (2), the main thing to prove is that $\zeta(s)$ has no zero on the line $\mathrm{Re}(s) = 1$. This non-vanishing result is common to all complex-analytic proofs of the prime number theorem and is really at the heart of the proof.

Historically, it also was the most difficult part – today, there are some tricks to prove it with little effort.

The Wiener-Ikehara theorem is not the most standard way to obtain the prime number theorem, and has the drawback of only producing a *qualitative* result, meaning a main term but without interesting error term. If one wishes to obtain a stronger *quantitative* result, it is possible to use alternative analytical techniques such as the more elementary (truncated) Perron formula.

It was apparent in the previous discussion that zeroes of the Riemann zeta function play a fundamental role in the study of the distribution of prime numbers. The prime number theorem is essentially equivalent to the fact that ζ has no zero on the line $\mathrm{Re}(s) = 1$. The Vinogradov-Korobov bound (2) corresponds to a larger zero-free region, and the Riemann hypothesis (1) is the statement that ζ has no zero on $\mathrm{Re}(s) > 1/2$.

Generalizing the above strategy to more general complex-analytic functions, it is possible to obtain a lot of fundamental estimates in analytic number theory, such as substantial information on the distribution of primes in arithmetic progressions. In this case, one studies so-called *Dirichlet L-series*

$$\sum_{n\geq 1} \frac{\chi(n)}{n^s}$$

where $\chi : \mathbb{N} \to \mathbb{C}$ are multiplicative functions known as *Dirichlet characters*. A classic reference for multiplicative number theory is given by [2].

2.2. **The Circle Method.** In this section, we introduce the Hardy-Littlewood circle method, originally developed to estimate the number of ways of writing an integer as an unordered sum of smaller integers (with arbitrarily many terms). For a standard account of the Hardy-Littlewood circle method, we refer the reader to [15]. This technique, which is more than a hundred years old, has proven to be exceptionally fruitful and is still widely used today. For example, it was used in 1930 by Vinogradov to tackle the ternary Goldbach problem.[2]

Theorem 2.2 (Vinogradov's theorem)**.** *Every sufficiently large odd number can be written as the sum of three primes.*

The corresponding problem for sums of two primes is still open.

Conjecture 2.3 (Binary Goldbach conjecture)**.** *Every sufficiently large even number can be written as the sum of two primes.*

Let us give a brief sketch of how the circle method can be applied to answer the following question: are there infinitely many arithmetic progressions of length 3 in the set of prime numbers? That is, are there infinitely many triples of prime numbers of the form p, $p+r$, $p+2r$? This problem was solved in 1939 by van der Corput by means of the circle method.

Not only does the circle method show the existence of infinitely many three-term arithmetic progressions in the primes: it even gives a precise asymptotic formula! Van der Corput proved the remarkable asymptotic

$$(\mathfrak{S} + o(1))\frac{x^2}{(\log x)^3}$$

[2]In 2013, Helfgott removed the "sufficiently large" assumption and proved, by very careful estimates, that every odd number ≥ 7 is a sum of three primes.

for the number of three-term arithmetic progressions in the prime numbers up to x, where $\mathfrak{S}$ is the constant

$$\mathfrak{S} := \prod_{p\geq 3}\left(1 - \frac{p^2-4p+1}{(p-1)^4}\right) \approx 1.0481,$$

known as a *singular series*.

Instead of working with the primes directly, we use the more convenient von Mangoldt function Λ that already appeared in the previous section. After some elementary manipulations, the result that we want to prove is

$$\sum_{n,r\in\mathbb{Z}} f(n)f(n+r)f(n+2r) \sim \mathfrak{S}x^2,$$

where $f(n) = \Lambda(n)\mathbf{1}_{n\leq x}$.

The idea is to rewrite the left-hand side in the language of Fourier analysis. Write $e(\alpha) := e^{2\pi i\alpha}$. The Fourier transform of f is the function $S : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ given by

$$S(\alpha) = \sum_{n\leq x} \Lambda(n)e(-n\alpha),$$

and we have the identity

$$\sum_{n,r\in\mathbb{Z}} f(n)f(n+r)f(n+2r) = \int_0^1 S(\alpha)^2 S(-2\alpha)\, d\alpha.$$

The problem of estimating a sum has turned into that of estimating an integral over $[0,1]$. The crucial observation is that the function $S(\alpha)$ is small unless α is "close" to a rational number with a "small" denominator. The integral can thus be split in two parts:

- the set $\mathfrak{M}$ of real numbers $\alpha \in [0,1]$ close to a rational number with a small denominator;
- the complement $\mathfrak{m} = [0,1] \setminus \mathfrak{M}$.

Here $\mathfrak{M}$ is a union of small intervals around each rational number with a small denominator, called the *major arcs*, and $\mathfrak{m}$ is called the *minor arcs*.

As is typical, $S(\alpha)$ can be computed very precisely for α in the major arcs. If α is the rational a/q, where the denominator q is small, then

$$S(a/q) = \sum_{n\leq x} \Lambda(n)e(-na/q),$$

but $e(-na/q)$ takes only q distinct values, so we can group

$$S(a/q) = \sum_{0\leq r<q} e(-r/q) \sum_{\substack{n\leq x \\ na\equiv r \ (\mathrm{mod}\ q)}} \Lambda(n).$$

The inner sum can be estimated by the methods of multiplicative number theory, yielding a precise expression for $S(a/q)$. This expression remains accurate for α in a small neighbourhood of a/q. If one can obtain a suitable bound for the integral on the minor arcs, the claimed asymptotic follows.

There are of course many technical details that need to be filled in, such as specifying what a "small" or "close" mean in the previous paragraphs, or carrying out all the computations.

The definition of major arcs differs according to the problem that is being studied. For example, in [9], Maynard proves that there are infinitely many primes whose decimal expansion does not contain the digit a_0 (where $a_0 \in \{0, 1, \ldots, 9\}$ is arbitrary). Here, the main contribution comes from the $\alpha \in [0, 1]$ close to a rational number with a small denominator, and whose decimal expansion contains many 0's or 9's.

Proving that the integral, when restricted to the minor arcs, is negligible, is often the most difficult part of the proof. This is precisely why Conjecture 2.3 is still open: even if the integral on the minor arcs seems to be small, nobody knows how to prove it. To control the integral on minor arcs for his result on primes with restricted digits, Maynard [9] used *sieve theory*, a powerful collection of techniques that we introduce in the next section.

2.3. **Sieve Theory.** Sieve theory is a vast area of which we will only scratch the surface. In a nutshell, sieve theory is concerned with giving lower and upper bounds on the size of a set obtained as the result of some inclusion-exclusion procedure. Sieve theory is actually more general, and the interested reader is advised to check out the comprehensive monograph [3] by Friedlander and Iwaniec.

The primary example is when we have a set $A \subset \mathbb{N}$, and we want to know how many elements in A are not divisible by any $p \in P$, where P is a fixed set of primes. If A is a subset of $[1, x]$, and if P is the set of all primes $p \leq \sqrt{x}$, then these elements are exactly the prime numbers in A that are $> \sqrt{x}$.

Let us mention three of the highlights of sieve theory. The first is Chen's theorem, which comes remarkably close to the twin prime conjecture.

Theorem 2.4 (Chen's theorem (1973))**.** *There are infinitely many primes p such that $p+2$ is a product of at most two primes.*

In a different flavour, we have the Friedlander-Iwaniec theorem, which shows the existence of primes in a very thin arithmetically interesting set of integers.

Theorem 2.5 (Friedlander-Iwaniec (1997))**.** *There are infinitely many primes of the form $a^2 + b^4$, with a, b integers.*

Finally, here is Zhang's breakthrough on small gaps between primes.

Theorem 2.6 (Zhang (2014))**.** *There is some $C > 0$ such that the following holds. There are infinitely many pairs of distinct primes (p, q) with $|p - q| \leq C$.*

Further refinements of Zhang's ideas showed that $C = 246$ is achievable. Of course, the twin prime conjecture claims that $C = 2$ should work, but this seems to be out of reach.

Let us give a very basic example of a sieve, called the Legendre sieve. Let x be a large real number and let $A = \mathbb{N} \cap [1, x]$. Suppose that we wish to count the number of primes in A (without using the prime number theorem). To do so, we should remove from A the multiples of 2, of 3, of 5, and so on, up to $\sqrt{x}$ (except the primes $2, 3, 5, \ldots$ themselves).

We start by showing what we obtain when we remove the multiples of 2, 3 and 5. Let

$$A_d = \{n \leq x \mid d \mid n\}.$$

By the inclusion-exclusion principle, we can say that

$$|\{n \leq x : \forall p \in \{2,3,5\},\ p \nmid n\}| = |A| - |A_2| - |A_3| - |A_5| + |A_6| + |A_{10}| + |A_{15}| - |A_{30}|.$$

The right-hand side can be written more concisely as

$$|\{n \leq x : \forall p \in \{2,3,5\},\ p \nmid n\}| = \sum_{\substack{d \in \mathbb{N} \\ P^+(d) \leq 5}} \mu(d)|A_d|,$$

where $\mu(d) = \lambda(d)\mathbf{1}_{d \text{ is squarefree}}$ is called the *Möbius function* and $P^+(d)$ is the largest prime factor of d.

We will not quite be able to consider primes up to $\sqrt{x}$, so we introduce a parameter z that will be optimized later. By the inclusion-exclusion principle, we have

$$|\{n \leq x : \forall p \leq z \text{ (prime)},\ p \nmid n\}| = \sum_{\substack{d \in \mathbb{N} \\ P^+(d) \leq z}} \mu(d)|A_d|.$$

Since the number of positive integers $\leq x$ that are divisible by d is $|A_d| = \lfloor x/d \rfloor = x/d + O(1)$, we have

$$\begin{aligned}
|\{n \leq x : \forall p \leq z,\ p \nmid n\}| &= \sum_{\substack{d \in \mathbb{N} \\ P^+(d) \leq z}} \mu(d) \left(\frac{x}{d} + O(1) \right) \\
&= x \sum_{\substack{d \in \mathbb{N} \\ P^+(d) \leq z}} \frac{\mu(d)}{d} + O(2^z) \\
&= x \prod_{p \leq z} \left(1 - \frac{1}{p} \right) + O(2^z),
\end{aligned}$$

where p always denotes a prime number. The error term $O(2^z)$ is quickly very large: this is what prevents us from choosing $z = \sqrt{x}$. We choose instead $z = \frac{1}{10} \log x$ and obtain

$$|\{n \leq x : \forall p \leq \tfrac{1}{10} \log x,\ p \nmid n\}| = x \prod_{p \leq \frac{1}{10} \log x} \left(1 - \frac{1}{p} \right) + O(x^{1/5}).$$

Mertens' estimate (which is a lot simpler than the prime number theorem) implies that

$$\prod_{p \leq z} \left(1 - \frac{1}{p} \right) = O\left(\frac{1}{\log z} \right).$$

Putting everything together, we obtain the bound

$$|\{n \leq x : n \text{ is prime}\}| = O\left(\frac{x}{\log \log x} \right),$$

since any prime number $n \leq x$ is in the set $\{n \leq x : \forall p \leq \frac{1}{10}\log x,\ p \nmid n\}$, except for the primes $\leq \frac{1}{10}\log x$, but there are very few of them.

Thus, only using the inclusion-exclusion principle (and Mertens' estimate), we managed to prove that there are $O\left(x/\log\log x\right)$ prime numbers $\leq x$. Naturally, this is much weaker than the prime number theorem, but this example already opens the way to other more sophisticated and far more powerful methods.

2.4. **The Density Increment Strategy.** This technique goes by many names according to the context in which it is used: density/energy/quality increment strategy, rank/entropy decrement argument, compression argument, etc. The original idea is often attributed to Roth, even if the spirit of the technique is so general that a lot of different proofs can be rephrased in that language.

Combining the Fourier analytic arguments of Section 2.2 with the density increment strategy, Roth [11] was able to prove the following result.

Theorem 2.7 (Roth (1953)). *Let A be a subset of $\{1, 2, \ldots, N\}$ without three-term arithmetic progressions. Then*

$$|A| = O\left(\frac{N}{\log\log N}\right).$$

The proof of Roth's theorem uses the following dichotomy:

- either A has "many" three-term arithmetic progressions;
- or there is a smaller interval $I \subset \{1, 2, \ldots, N\}$ such that the density

  $$\frac{|A \cap I|}{|I|}$$

  is "significantly larger" than the original density $|A|/N$.

This dichotomy is proved by studying the Fourier transform of the indicator function of the set A. Roughly speaking, if all Fourier coefficients of the centered indicator function $\mathbf{1}_A - \frac{|A|}{N}$ are small, then a direct computation (as in the circle method) shows that A contains many arithmetic progressions. Otherwise, there is a large Fourier coefficient, and this coefficient can be used to construct an interval I where A has increased density.

If the first case holds, we are done. In the second case, we replace A with $A \cap I$ and $\{1, 2, \ldots, N\}$ with I. The dichotomy can again be applied to $A \cap I$. If $A \cap I$ has many three-term arithmetic progressions, then so does A and we are done. Otherwise, we find, yet again, a subset with increased relative density. We can repeat the argument; since the density never exceeds 1, this procedure must terminate after a bounded number of steps. If we specified what "many" and "significantly larger" mean, we would, after some simple computations, obtain Theorem 2.7.

Since the work of Roth, the density increment strategy has experienced significant growth and can now reasonably be considered a standard technique in the analytic number theorist's toolbox. For example, it was used in 2020 by Koukoulopoulos and Maynard to prove the Duffin-Schaeffer conjecture [6] on approximations of real numbers by rationals.

The upper bound on $|A|$ in Theorem 2.7 has been improved several times. For many years, it seemed impossible with the available techniques to prove anything better than

$$|A| = O\left(\frac{N}{\log N}\right).$$

In a recent breakthrough by Bloom and Sisask [1], this barrier was finally crossed.

Theorem 2.8 (Bloom-Sisask (2020))**.** *Let A be a subset of $\{1, 2, \ldots, N\}$ without three-term arithmetic progressions. Then*

$$|A| = O\left(\frac{N}{(\log N)^{1+c}}\right),$$

where c is some (very small, absolute) positive constant.

Despite its innocuous appearance, this theorem was the first sufficiently strong bound to imply the Erdös conjecture on arithmetic progressions of length three.

Corollary 2.9 (Erdös' conjecture)**.** *Let A be a subset of $\mathbb{N}$ such that $\sum_{n\in A} \frac{1}{n} = +\infty$. Then A contains infinitely many three-term arithmetic progressions.*

The same statement is believed to be true for longer progressions, but this seems completely out of reach with the current methods.

Using the techniques of Bloom and Sisask, the author proved in [10] a multidimensional variant of Theorem 2.8 and deduced the following conjecture of Shkredov and Solymosi.

Theorem 2.10 (P. (2021))**.** *Every subset $A \subset \mathbb{Z}^2$ such that $\sum_{a\in A} \frac{1}{\|a\|^2} = +\infty$ contains infinitely many isosceles right triangles.*

3. Towards Chowla's Conjecture

After this overview of some standard techniques in analytic theory, we return to the specific problem of Chowla's conjecture, Conjecture 1.2.

In 2015 Matomäki, Radziwiłł and Tao proved that Conjecture 1.2 holds *on average* over $h_1, \ldots, h_k$ (for every k)[8]. A crucial ingredient in their proof was the groundbreaking work by Matomäki and Radziwiłł [7] on sums of multiplicative functions over short intervals.

One year later, Tao proved a *logarithmic* version of Chowla's conjecture for $k = 2$ [12]. This means that, instead of the regular average $\frac{1}{x}\sum_{n\le x}$ in Conjecture 1.1, he considers the logarithmic average

$$\frac{1}{\log x}\sum_{n\le x}\frac{1}{n}\lambda(n)\lambda(n+h),$$

and showed that this average has limit zero as $x \to \infty$.

Even if this result is weaker than Conjecture 1.1, it is extremely interesting because it breaks the so-called *parity barrier*. Quoting Friedlander-Iwaniec [3, Section 16.4], the parity obstruction is the inability to "establish [...] the existence of integers with an even number of prime factors or of integers with an odd number of prime factors" in a general set $A \subset \mathbb{N}$ using the standard framework of sieve theory.

Tao's proof [12] used a very original entropy decrement argument (see Section 2.4) and obtained the explicit bound

$$\frac{1}{\log x}\sum_{n\leq x}\frac{1}{n}\lambda(n)\lambda(n+h) = O\left(\frac{1}{(\log\log\log\log x)^{1/5}}\right).$$

Later work of Tao and Terräväinen [13] improved this to

$$\frac{1}{\log x}\sum_{n\leq x}\frac{1}{n}\lambda(n)\lambda(n+h) = O\left(\frac{1}{(\log\log\log x)^{c}}\right)$$

for some constant $c > 0$. In 2021, Helfgott and Radziwiłł [5] used a completely different approach to produce the following quantitative improvement (for simplicity we fix $h = 1$).

Theorem 3.1 (Logarithmic Chowla conjecture in degree 2)**.** *As* $x \to \infty$*,*

$$\frac{1}{\log x}\sum_{n\leq x}\frac{1}{n}\lambda(n)\lambda(n+1) = O\left(\frac{1}{(\log\log x)^{1/2}}\right).$$

To obtain Theorem 3.1, the authors reduced the problem to bounding the eigenvalues of a certain linear operator A carrying deep arithmetical information. They then used the *high trace method*, which is well-known in certain areas such as random matrix theory. The idea is to bound the trace of a high power of A. Since A is hermitian, its eigenvalues are real and a bound on

$$\mathrm{Tr}(A^{2k})$$

induces a bound on the eigenvalues of A, as this trace is the sum of the $2k$ powers of these eigenvalues. The trace is much more amenable to explicit computations than the eigenvalues themselves. Using a convenient basis and rephrasing the problem in terms of walks on an arithmetic graph, the authors managed to obtain an optimal bound for the norm of A and deduce Theorem 3.1.

The previous sketch only briefly touched on the linear algebra part of the proof. The proof of Theorem 3.1 requires a lot more ingenious ideas. An expository paper explaining the work of Helfgott and Radziwiłł can also be found in [4].

The proof of Theorem 3.1 is currently being re-written in a more concise way by the author for his Master's thesis, under the supervision of Prof. James Maynard (Oxford). Further explorations will follow to see if the bound can be improved and to try to use the techniques introduced by Helfgott and Radziwiłł to other problems.

References

[1] Thomas Bloom and Olof Sisask, *Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions*, arXiv:2007.03528 (2020).

[2] Harold Davenport, *Multiplicative number theory*, vol. 74, Springer, 1980.

[3] John B. Friedlander and Henryk Iwaniec, *Opera de Cribro*, vol. 57, American Mathematical Society, 2010.

[4] Harald Andrés Helfgott, *Expansion, divisibility and parity: an explanation*, arXiv:2201.00799 (2022).

[5] Harald Andrés Helfgott and Maksym Radziwiłł, *Expansion, divisibility and parity*, arXiv:2103.06853 (2021).

[6] Dimitris Koukoulopoulos and James Maynard, *On the Duffin-Schaeffer conjecture*, Annals of Mathematics **192** (2020), no. 1, 251–307.

[7] Kaisa Matomäki and Maksym Radziwiłł, *Multiplicative functions in short intervals*, Annals of Mathematics (2016), 1015–1056.

[8] Kaisa Matomäki, Maksym Radziwiłł, and Terence Tao, *An averaged form of Chowla's conjecture*, Algebra Number Theory **9** (2015), no. 9, 2167–2196.

[9] James Maynard, *Primes with restricted digits*, Inventiones mathematicae **217** (2019), no. 1, 127–218.

[10] Cédric Pilatte, *New bound for Roth's theorem with generalized coefficients*, arXiv:2111.03838 (2021).

[11] Klaus F. Roth, *On certain sets of integers*, J. London Math. Soc **28** (1953), no. 104-109, 3.

[12] Terence Tao, *The logarithmically averaged Chowla and Elliott conjectures for two-point correlations*, **4** (2016).

[13] Terence Tao and Joni Teräväinen, *The structure of correlations of multiplicative functions at almost all scales, with applications to the Chowla and Elliott conjectures*, Algebra & Number Theory **13** (2019), no. 9, 2103–2150.

[14] Gerald Tenenbaum, *Introduction to analytic and probabilistic number theory*, vol. 46, Cambridge University Press, 1995.

[15] R. C. Vaughan, *The Hardy-Littlewood method*, vol. 80, Cambridge University Press, 1981.

Department of Mathematics and their Applications, École Normale Supérieure (ENS).
Email address: `cedric.pilatte@ens.fr`