Teoría probabilística de números

PROBABILISTIC NUMBER THEORY

EMMANUEL KOWALSKI

By this art you may contemplate
the variation of the 23 *letters*
J-L. Borges, "La biblioteca de Babel"

I could be bounded in a nutshell and
count myself a king of infinite space
W. Shakespeare, "Hamlet"

1. Introduction

What is probabilistic number theory? In the widest sense, one could say that it coincides more or less with analytic number theory (it is maybe for this reason that certain arithmeticians, lest their work be associated with analysis, prefer the term "arithmetic statistics"), but this survey will approach it from a more restrictive direction, in order to focus on certain specific aspects (chosen by the author due to his own personal bias): we will consider probabilistic number theory as *the study of the asymptotic behavior of sequences (or families) of arithmetically defined random variables*. This is also the spirit of the lecture notes [16], and readers may consider that these notes form the natural extension and continuation of this survey. Another excellent recent survey, covering an especially wide variety of the most recent interactions, is due to Perret-Gentil [20].

To illustrate our intended focus, we start by stating what is usually taken to be the first great result of probabilistic number theory, the Erdős–Kac Theorem (published in 1940, see [3]):

Theorem 1.1 (Erdős–Kac)**.** *For any integer $n \geqslant 1$, let $\omega(n)$ be the number of distinct prime divisors of n. As $\mathrm{N} \to +\infty$, the random variables*

$$n \mapsto \frac{\omega(n) - \log\log \mathrm{N}}{\sqrt{\log\log \mathrm{N}}}$$

on the finite probability space $\Omega_{\mathrm{N}} = \{1, \dots, \mathrm{N}\}$, with the uniform probability measure, converge in law to the standard gaussian random variable with expectation 0 *and variance* 1*.*

In concrete terms (and as it was first stated), this means that for any real numbers $a < b$, we have the following limit (where even the existence of the limit is by no means obvious):

$$\lim_{\mathrm{N}\to+\infty} \frac{1}{\mathrm{N}} \Big| \Big\{ n \leqslant \mathrm{N} \mid a < \frac{\omega(n) - \log\log \mathrm{N}}{\sqrt{\log\log \mathrm{N}}} < b \Big\} \Big| = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt.$$

Thus, the most fundamental of all probability distributions, the gaussian distribution, can be realized from extremely simple, and seemingly deterministic, objects: the number of prime divisors of "random" integers. And we see how considering the intervals $\{1, \dots, \mathrm{N}\}$

with $N \to +\infty$ is a natural way to speak of "random" integers – but it is not the only one, as we will discuss later (see Example 2.3).

Some basic probabilistic notation and terminology, and a few results, will be recalled in the Appendix, but the notion of convergence in law is so essential that we wish to state it here formally in a fairly general setting:

Definition 1.2. Let (X_n) be a sequence of random variables, defined on some probability spaces[1] and with values in a fixed metric space M. Let X be a random variable with values in M. The sequence (X_n) converges in law to X if, for any function $f\colon M \to \mathbf{C}$ that is continuous and bounded, we have

$$\lim_{n\to+\infty} \mathbf{E}(f(X_n)) = \mathbf{E}(f(X)).$$

Although the majority of results will involve only $M = \mathbf{R}$ or a finite-dimensional vector space, we will see in Section 4 that probabilistic number theory may also involve functional limit theorems, where the space M is an infinite-dimensional vector space.

In the next section, we will sketch a proof of Theorem 1.1, but only after explaining some older results which are also of great interest. In Section 3, we discuss one of the most important theorems of probabilistic number theory: Selberg's Theorem on the limiting distribution of $\log|\zeta(1/2+it)|$, where $\zeta(s)$ denotes the Riemann Zeta function. We will be able to motivate the result, but a full proof is quite delicate. Nevertheless, this is an area where much of the most recent and most exciting work has been done, with new connections to sophisticated concepts of "pure" probability theory, such as branching random walks or gaussian multiplicative chaos. Then in Section 4, we explain a rather different result, due to W. Sawin and the author, where the limit theorems involve random variables with values in the space $\mathcal{C}([0,1])$ of continuous functions on $[0,1]$, and where moreover the key arithmetic tools are especially deep – Deligne's version of the Riemann Hypothesis over finite fields [2] is the key to the proof.

Acknowledgements. Thanks to J. Fresán for suggesting to me to write this survey, and for translating it into Spanish.

The section on Selberg's Theorem is inspired by the lecture I gave in the Betty B. Seminar in March 2018, in preparation for the talk of A. Harper in the N. Bourbaki Seminar the following day; I thank N. and B. Bourbaki for the invitation to speak there.

2. Classical probabilistic number theory

2.1. **Schoenberg's Theorem.** Although the Erdős–Kac Theorem is, as we mentioned, usually taken as the starting point of probabilistic number theory, one can argue that earlier results of Schoenberg probably might well deserve this title rather more. In fact, one can go further and say that the philosophical origin of probabilistic number theory, is the following elementary fact, which not only lies at the core of both Schoenberg's results and the Erdős–Kac Theorem but also explains why interesting interactions between probability theory and number theory are possible.

Theorem 2.1. *For any positive integer q, let π_q be the random variable on Ω_N with values in $\mathbf{Z}/q\mathbf{Z}$ defined by $\pi_q(n) = n \pmod q$.*

[1] Which may depend on n.

Let $\mathcal{Q}$ be a finite set of coprime positive integers and let Q be the product of the elements of $\mathcal{Q}$. Then the family $(\pi_q)_{q\in\mathcal{Q}}$ converges in law as $\mathrm{N}\to+\infty$ to a family of independent random variables $(\mathrm{U}_q)_{q\in\mathcal{Q}}$, where U_q is uniformly distributed in $\mathbf{Z}/q\mathbf{Z}$.

In fact, for any function

$$f\colon \prod_{q\in\mathcal{Q}} \mathbf{Z}/q\mathbf{Z}\to\mathbf{C}$$

we have

$$\Big|\frac{1}{\mathrm{N}}\sum_{n\in\Omega_{\mathrm{N}}} f((\pi_q(n))_q) - \frac{1}{\mathrm{Q}}\sum_{x\in\prod_{q\in\mathcal{Q}}\mathbf{Z}/q\mathbf{Z}} f(x)\Big| \leqslant \frac{1}{\mathrm{N}}\sum_x |f(x)|.$$

Remark 2.2. In other words, this is a case of Definition 1.2 with

$$\mathrm{M}=\prod_{q\in\mathcal{Q}}\mathbf{Z}/q\mathbf{Z},\qquad \mathrm{X}_{\mathrm{N}}(n)=(\pi_q(n))_{q\in\mathcal{Q}},\qquad \mathrm{X}((a_q)_{q\in\mathcal{Q}})=(a_q)_{q\in\mathcal{Q}}.$$

It is more or less tautological using the definition of independent families (see Section A.1) that the components of X are indeed independent and uniformly distributed.

This is in some sense "obvious", and indeed it is used without comments or explanation in many treatments of classical probabilistic number theory. In fact, it follows (qualitatively) from the Chinese Remainder Theorem, which shows that the product of $\mathbf{Z}/q\mathbf{Z}$ for $q\in\mathcal{Q}$ can be identified with $\mathbf{Z}/\mathrm{Q}\mathbf{Z}$, together with the elementary limit

$$\lim_{\mathrm{N}\to+\infty}\frac{1}{\mathrm{N}}|\{n\in\Omega_{\mathrm{N}}\mid n\equiv a\,(\mathrm{mod}\,\mathrm{Q})\}|=\frac{1}{\mathrm{Q}}$$

for any a modulo Q.

However, almost any attempt to generalize non-trivially the Erdős–Kac Theorem is likely to face the fact that it is necessary to generalize Theorem 2.1, and that this might involve very deep mathematics indeed. We illustrate this with two examples:

Example 2.3. (1) Replace the integers from 1 to N by the values $p-1$, where p runs over primes from 1 to N (again with the uniform probability measure on this finite set of primes); then the analogue of Theorem 2.1 is immediately connected to questions about primes in arithmetic progressions which are closely related to the Generalized Riemann Hypothesis.

(2) Fix an integer $m\geqslant 2$. Replace the integers from 1 to N by the integers $\mathrm{Tr}(g)$, where g runs uniformly over all matrices in $\mathrm{SL}_m(\mathbf{Z})$ with all coefficients $\leqslant \mathrm{N}$. Then the analogue of Theorem 2.1 is intimately related to the issue of spectral gaps for congruence subgroups of $\mathrm{SL}_m(\mathbf{Z})$ (see for instance the discussion in [15]), and thus linked to topics like expander graphs, automorphic forms, Kazhdan's Property (T), etc.

Theorem 2.1 is often applied with $\mathcal{Q}$ a set of primes, so that the coprimality condition between distinct elements of $\mathcal{Q}$ is satisfied. What is crucial from the probabilistic point of view is that we see *sequences of independent random variables* appear. Since independence is one of the most fundamental of all concepts in probability theory (one of the first things that distinguishes probability from integration theory), it cannot be surprising to think that interesting interactions should exist between number theory and probability.

Using Theorem 2.1, we will be able now to motivate the following very nice statement due to Schoenberg [23]; in fact, a full proof doesn't require any more arithmetic information, and only quite basic probability theory, as our sketch will show.

We recall that Euler's function φ is defined so that $\varphi(n)$ is the number of invertible residue classes modulo n for all positive integers n. Equivalently, it is the number of integers a such that $0 \leqslant a < n$ and a is coprime to n, so that for instance $\varphi(n) = n - 1$ is equivalent to saying that n is prime.

The size of $\varphi(n)$ is always "relatively close" to n in some sense, but the ratio[2] $\varphi(n)/n$ does not have a limit when n is large, and indeed experiments show quickly that it varies quite erratically in $[0, 1]$ – if you write down a large integer (say with 50 digits) and try to guess what $\varphi(n)/n$ will be, you might well be very puzzled by what happens. This is explained by Schoenberg's Theorem.

Theorem 2.4 (Schoenberg)**.** *For* $\mathrm{N} \geqslant 1$*, let* F_N *be the random variable* $n \mapsto \varphi(n)/n$ *on* Ω_N*. As* $\mathrm{N} \to +\infty$*, the random variables* (F_N) *converge in law to a random variable* F*, which can be described as the infinite product over prime numbers*

$$\mathrm{F} = \prod_p \Big(1 - \frac{\mathrm{B}_p}{p}\Big) \tag{1}$$

where (B_p) *is a sequence of independent Bernoulli random variables, indexed by primes, such that for any* p*, we have*

$$\mathbf{P}(\mathrm{B}_p = 1) = \frac{1}{p}, \qquad \mathbf{P}(\mathrm{B}_p = 0) = 1 - \frac{1}{p}.$$

This infinite product converges almost surely.

Sketch of proof. The starting point is the well-known formula

$$\frac{\varphi(n)}{n} = \prod_{p|n} \Big(1 - \frac{1}{p}\Big) \tag{2}$$

for the normalized Euler function. This reflects the fact that $\varphi(n) = |(\mathbf{Z}/n\mathbf{Z})^\times|$ (the number of invertible elements in the ring $\mathbf{Z}/n\mathbf{Z}$), and that the Chinese Remainder Theorem provides an isomorphism of groups

$$(\mathbf{Z}/n\mathbf{Z})^\times \simeq \prod_{p|n} (\mathbf{Z}/p^{v_p(n)}\mathbf{Z})^\times$$

where $v_p(n)$ is the exponent of the prime p in the factorization of n; the formula then follows from the fact that

$$\varphi(p^v) = p^v - p^{v-1} \quad \text{so} \quad \frac{\varphi(p^v)}{p^v} = 1 - \frac{1}{p}$$

for any prime number p and any integer $v \geqslant 1$.

Now we rewrite (2) cleverly as an infinite product over primes

$$\frac{\varphi(n)}{n} = \prod_p \Big(1 - \frac{b_p(n)}{p}\Big) \tag{3}$$

where $b_p(n)$ is simply equal to 1 if p divides n and to 0 otherwise.

Next we observe that, if we restrict n to Ω_N, and view b_p as a random variable, then *by definition* and by Theorem 2.1, we deduce that these random variables converge as $\mathrm{N} \to +\infty$

[2] Which also has a probabilistic interpretation, as the probability, for the uniform measure, that a residue class nodulo n is invertible.

to the Bernoulli random variable B_p with "success probability" equal to $1/p$. (Intuitively, we are just saying that a "random integer" has probability $1/p$ of being divisible by p, which certainly makes sense!). Furthermore, applying again Theorem 2.1, it follows that for any finite family of distinct primes $(p_i)_{1\leqslant i\leqslant k}$, the tuples $(b_{p_1}(n),\dots,b_{p_k}(n))$ for $n\in\Omega_{\mathrm{N}}$ converge in law to the tuple $(\mathrm{B}_{p_1},\dots,\mathrm{B}_{p_k})$, where the Bernoulli random variables are *independent*.

At this point, if we look at (3), knowing that the family (b_p) restricted to Ω_{N} converges in law to (B_p) when $\mathrm{N}\to+\infty$, we are naturally led to expect that there should be a limiting random variable, given by the infinite product (1): we are just passing to the limit in each term of (3).

We sketch an elementary approach to obtain a rigorous proof. It is based on the probabilistic Lemma A.2 in the Appendix, and roughly speaking amounts to finding successive approximations of the random variables, for each N, which each converge in law, and are such that the difference in small on average. Here the natural choice of approximation is the finite product

$$\prod_{p\leqslant \mathrm{M}}\Big(1-\frac{b_p(n)}{p}\Big),$$

where $\mathrm{M}\geqslant 1$ is a parameter determining the quality of the approximation (also viewed as random variables on Ω_{N}). Since the product is finite, it is a straightforward consequence of Theorem 2.1 that these approximations, restricted to Ω_{N} converge in law as $\mathrm{N}\to\infty$ to

$$\prod_{p\leqslant \mathrm{M}}\Big(1-\frac{\mathrm{B}_p}{p}\Big).$$

To check the quality of the approximation, we note simply that

$$\begin{aligned}\Big|\frac{1}{\mathrm{N}}\Big(\sum_{n\leqslant \mathrm{N}}\frac{\varphi(n)}{n}-\sum_{n\leqslant \mathrm{N}}\prod_{p\leqslant \mathrm{M}}\Big(1-\frac{1}{p}\Big)\Big)\Big| &\leqslant \frac{1}{\mathrm{N}}\sum_{n\leqslant \mathrm{N}}\Big|\prod_{\substack{p\mid n\\ p>\mathrm{M}}}\Big(1-\frac{1}{p}\Big)-1\Big|\\ &\leqslant \frac{1}{\mathrm{N}}\sum_{n\leqslant \mathrm{N}}\sum_{d\in \mathrm{D}_n}\frac{1}{d}\end{aligned}$$

where D_n is the set of integers $d\geqslant 2$ dividing n which only have prime factors $p>\mathrm{M}$. In particular, all elements d of D_n satisfy $\mathrm{M}<d\leqslant \mathrm{N}$, and if we exchange the order of the sums over n and d, we obtain

$$\frac{1}{\mathrm{N}}\sum_{n\leqslant \mathrm{N}}\sum_{d\in \mathrm{D}_n}\frac{1}{d}\leqslant \sum_{\mathrm{M}<d\leqslant \mathrm{N}}\frac{1}{d}\times\frac{1}{\mathrm{N}}\sum_{\substack{n\leqslant \mathrm{N}\\ d\mid n}}1\leqslant\sum_{d>\mathrm{M}}\frac{1}{d^2}.$$

The right-hand side is a function $f(\mathrm{N},\mathrm{M})$ of N and M which tends to 0 as $\mathrm{M}\to+\infty$ uniformly with respect to N (since N doesn't appear). This means that the conditions of Lemma A.2 are satisfied.

Schoenberg's Theorem now follows, except for the statement of convergence almost surely of the infinite product (1); this is, however, a simple consequence of Kolmogorov's Three Series Theorem (for instance), see Theorem A.3. □

The random variable F that occurs in this theorem is quite interesting. First of all, it certainly doesn't seem to be a "standard" random variable (Gaussian, Poisson, exponential,

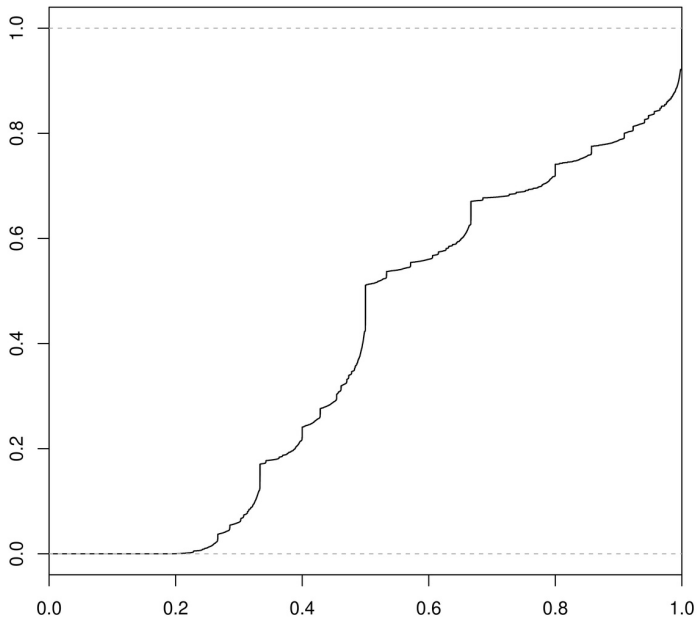

FIGURE 1. Empirical plot of the distribution function of $\varphi(n)/n$ for $n \leqslant 10^6$.

uniform, etc). In fact, it belongs to the class of real-valued random variables whose probability law are *singular* with respect to Lebesgue measure, a kind of "pathological" object that is usually encountered in the theory of integration, and constructed using Cantor-type sets. We thus see how wild and unpredictable number theory can be... Precisely, define a function f by $f(x) = \mathbf{P}(\mathrm{F} \leqslant x)$ for all $x \in \mathbf{R}$ (the function f is the *probability distribution function* of the random variable F). We then have the following:

Proposition 2.5 (Erdős)**.** *The function f is continuous on $\mathbf{R}$, strictly increasing on $[0, 1]$, differentiable almost everywhere on $\mathbf{R}$, with respect to Lebesgue measure, and moreover $f'(x) = 0$ for almost all x, with respect to Lebesgue measure.*

We illustrate this proposition (whose proof, due to Erdős, is quite elegant, see e.g. [16, Exercise 1.4.4]) with an approximation of the graph of the function $f(x)$ for $0 \leqslant x \leqslant 1$ (coming from the values $\mathrm{F}(n)$ for $n \leqslant 10^6$).

2.2. **The Erdős–Kac Theorem.** The method of proof of Schoenberg's Theorem can be extended to apply to many arithmetic functions replacing the Euler function. In fact, it is in general easier to deal with additive functions f, defined for positive integers n and which satisfy

$$f(nm) = f(n) + f(m)$$

when n and m are coprime. An example of such a function is $f(n) = \log(\varphi(n)/n)$, and there exist general distribution results for additive functions which recover Schoenberg's Theorem.

One of the simplest additive functions is the function ω that counts the number of distinct prime divisors of n. Understanding its distribution when n is taken uniformly at random in $\Omega_N = \{1, \dots, N\}$ is precisely what the Erdős–Kac Theorem does.

If we use the notation b_p of the previous section for the random variable on Ω_N that indicates whether n is divisible by p or not, we can write

$$\omega(n) = \sum_p b_p(n).$$

We might therefore expect an asymptotic behavior similar to that of the sum

$$\sum_p B_p.$$

We then see that the issue is more complicated than in Theorem 2.4: without some truncation, this sum diverges almost surely, by the (non-trivial direction of the) Borel–Cantelli Lemma and the fact that

$$\sum_p \mathbf{P}(B_p = 1) = \sum_p \frac{1}{p}$$

diverges (a well-known fact about the distribution of primes, going back to Euler, which we will also recall in the next section). But observing that any prime dividing $n \leqslant N$ is also $\leqslant N$, we can consider the finite sum

$$W_N = \sum_{p \leqslant N} B_p$$

and expect that this is a possible approximation to the distribution of the function $\omega(n)$ on Ω_N when N is large. This is indeed the case, but the approximation is not straightforward like convergence in law – indeed, the sequence (W_N) does not converge in law as $N \to +\infty$. However, because the (B_p) are independent Bernoulli random variables, it is a simple probability exercise to check that the renormalized random variables

$$\frac{W_N - \mathbf{E}(W_N)}{\sqrt{\mathbf{V}(W_N)}}$$

(which have expectation 0 and variance 1) converge in law to a standard gaussian. Here we have

$$\mathbf{E}(W_N) = \sum_{p \leqslant N} \frac{1}{p} \sim \log\log N$$

$$\mathbf{V}(W_N) = \sum_{p \leqslant N} \frac{1}{p}\Big(1 - \frac{1}{p}\Big) \sim \log\log N,$$

where the asymptotics for the sum of inverses of primes is known as *the Mertens formula* (see also (8) below for an explanation of where this comes from).

Although this is a somewhat unusual approach to the Erdős–Kac Theorem, it can be made rigorous, for instance by computing asymptocially the moments

$$\frac{1}{N} \sum_{n \leqslant N} \Big(\sum_{p \leqslant Q} b_p(n) \Big)^k$$

for all integers $k \geqslant 1$, and a suitable choice of $\mathrm{Q} < \mathrm{N}$, small enough to make the computation feasible, and large enough that the number of prime factors of $n \leqslant \mathrm{N}$ which are $> \mathrm{Q}$ is small enough to be negligible. When arguing in this manner, the key arithmetic fact is, once again, the elementary Theorem 2.1.

Another very probabilistic approach to the Erdős–Kac Theorem, based on Stein's method for Poisson and Gaussian approximation, is due to Harper [7]. One can also avoid the truncation step involving the parameter Q and compute asymptotically the characteristic functions of W_{N} and of $\omega(n)$ on Ω_{N}. This was first done by Rényi and Turán, and elaborating the precise form of the result, one is led to the so-called "mod-Poisson convergence" introduced by the author and Nikeghbali [17], which highlights intriguing connections with random permutations and random polynomials over finite fields. (An entertaining account of these is found in the nice survey [5] of A. Granville; his comic book with J. Granville [6] – somewhat flawed, in the opinion of this author – may also appeal to some readers.)

Example 2.6. Erdős observed the following very nice application of the fact that integers $n \leqslant \mathrm{N}$ have usually about $\log\log \mathrm{N}$ prime factors.[3] This is often called the "multiplication table problem". Consider the multiplication table of integers $1 \leqslant n \leqslant \mathrm{N}$, as can be found for $\mathrm{N} = 10$ or so in many schoolrooms. How many integers appear on the multiplication table? In other words, how many integers $m \leqslant \mathrm{N}^2$ can be expressed as a product $m = ab$ where $1 \leqslant a, b \leqslant \mathrm{N}$? If we call this number $m(\mathrm{N})$, then the first basic result is that we have

$$\lim_{\mathrm{N}\to+\infty} \frac{m(\mathrm{N})}{\mathrm{N}^2} = 0,$$

or in other words, "most" integers do not appear in the multiplication table. The reason, intuitively, is that since most integers $n \leqslant \mathrm{N}$ have about $\log\log \mathrm{N}$ prime factors, most products ab have about $2\log\log \mathrm{N}$ prime factors. This is however very atypical for integers $m \leqslant \mathrm{N}^2$, which again should have about $\log\log \mathrm{N}^2 \sim \log\log \mathrm{N}$ prime factors.

The (surprising) order of magnitude of $m(\mathrm{N})$ was determined by Ford [4], after works of a number of authors, especially Tenenbaum; his arguments rely on quite subtle probabilistic results.

3. Selberg's Theorem

We will discuss in this section the second most famous theorem in probabilistic number theory, and the starting point of what is today probably the deepest area of research in this area. This is a theorem of Selberg [24] concerning the limiting behavior of the logarithm of the Riemann zeta function on the so-called critical line.

3.1. The Riemann zeta function.

The Riemann zeta function is the meromorphic function on $\mathbf{C}$ which is obtained by analytic continuation from the function defined by the *Dirichlet series*

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

on the open set of complex numbers $s \in \mathbf{C}$ such that $\mathrm{Re}(s) > 1$, where the series converges absolutely and uniformly on compact subsets. This implies that ζ is holomorphic in this

[3] The required information can be proved more easily than the full Erdős–Kac Theorem.

domain. There are multiple approaches to obtain its analytic continuation and identify its poles (the classical book [25, Ch. II] of Titchmarsh lists seven).

One of the most intrinsic method (because of its links with modular forms, in this case theta functions) is to use the expression

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \frac{1}{s(s-1)} + \frac{1}{2}\int_1^{+\infty} (\theta(x)-1)(x^{s/2}+x^{(1-s)/2})\frac{dx}{x},$$

where

$$\theta(x) = \sum_{n\in\mathbf{Z}} e^{-\pi n^2 x}$$

for any $x > 0$. This formula follows relatively simply from the definition

$$\Gamma(s) = \int_0^{+\infty} e^{-t}t^s\frac{dt}{t}$$

of the Gamma function, combined with the Poisson summation formula. One sees easily (using the rapid decay of the theta function when $x \to +\infty$) that the right-hand side (hence the Riemann zeta function) is a meromorphic function on $\mathbf{C}$ with a unique simple pole at $s = 1$ with residue 1.

For the purpose of understanding Selberg's theorem, it is sufficient to know that $\zeta(s)$ can be defined for $\mathrm{Re}(s) > 0$ (except at $s = 1$), and one can establish this using very elementary means. For instance, the following "summation by parts" suffices: let $\{t\}$ denote the fractional part of a real number $t \geqslant 0$, so that $\{t\} = t - n$ where $n \geqslant 0$ is the largest integers $n \leqslant t$. For $\mathrm{Re}(s) > 1$ first, we compute

$$\begin{aligned}
\sum_{n\geqslant 1}\frac{1}{n^s} &= s\int_1^{+\infty}\Big(\sum_{1\leqslant n\leqslant t} 1\Big)t^{-s-1}dt \\
&= s\int_1^{+\infty}(t-\{t\})t^{-s-1}dt \\
&= s\int_1^{+\infty} t^{-s}dt - s\int_1^{+\infty}\{t\}t^{-s-1}dt \\
&= \frac{s}{s-1} + (\text{holomorphic function for } \mathrm{Re}(s) > 0),
\end{aligned}$$

which provides the desired analytic continuation.

The reason number theorists are interested in the properties of the Riemann zeta function is that it gives access to many properties of prime numbers. This is due to the remarkable *Euler product formula*, which states that for $\mathrm{Re}(s) > 1$, we have the relation

$$\zeta(s) = \prod_p (1-p^{-s})^{-1}, \tag{4}$$

where the infinite product over primes is absolutely convergent. This fact is rather elementary, and yet essential, so we explain the proof (where we will see that the arithmetic content of the formula is the existence and uniqueness of prime factorization): first, for any fixed real number $\mathrm{P} \geqslant 2$, we expand the geometric series for $(1-p^{-s})^{-1}$ and multiply the resulting

series over primes $p \leqslant \mathrm{P}$ to derive

$$\prod_{p \leqslant \mathrm{P}} (1 - p^{-s})^{-1} = \prod_{p \leqslant \mathrm{P}} \sum_{k \geqslant 0} p^{-ks} = \sum_{n \in \mathrm{N_P}} n^{-s},$$

where $\mathrm{N_P}$ is the set of integers $n \geqslant 1$ only divisible by primes $\leqslant \mathrm{P}$. Note that the uniqueness of prime factorization implies that each integer appears with multiplicity at most one, whereas the existence of prime factorization shows that every integer $n \leqslant \mathrm{P}$ belongs to $\mathrm{N_P}$. Letting $\mathrm{P} \to +\infty$, one then deduces the Euler formula for $\mathrm{Re}(s) > 1$.

Remark 3.1. The simplest applications of the Euler product are proofs that there are infinitely many primes:

- (Euler) Otherwise, $\zeta(\sigma)$ would be bounded for all $\sigma > 1$, which is not true.
- (Hacks) Otherwise, $\zeta(2) = \pi^2/6$ (Euler) would be a rational number, which is not true (Lambert).

3.2. **The Prime Number Theorem.** Riemann [22] discovered how to exploit systematically the Euler product in order to obtain asymptotic formulas for the number $\pi(x)$ of prime numbers $p \leqslant x$. Again, because of the importance of this idea, we discuss it briefly before coming to Selberg's Theorem.

The intuition can be presented relatively easily. First, it is better to consider the function

$$\psi(x) = \sum_{\substack{p \text{ premier}, \ k \geqslant 1 \\ p^k \leqslant x}} \log p = \sum_{n \leqslant x} \Lambda(n),$$

where Λ is the so-called von Mangoldt function, which is supported on (non-trivial) prime powers, and satisfies $\Lambda(p^k) = \log p$ for all $k \geqslant 1$. One can easily compare $\pi(x)$ and $\psi(x)$, but the function ψ is easier to handle, essentially because of the formula

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

which directly relates the von Mangoldt function with the oppposite of the logarithmic derivative of the Riemann zeta function (this follows from the Euler product by termwise differentiation). From the properties of the Riemann zeta function, we see that this is also a meromorphic function, and moreover its poles are all simple poles, and are located as follows:

- There is a unique simple pole with residue 1 at $s = 1$;
- There are simple poles at the points $s = \varrho$ for any $\varrho \in \mathbf{C}$ such that $\zeta(\varrho) = 0$, with residue equal to the opposite of the multiplicity of ϱ as a zero of the Riemann zeta function.

Using a computation based on the Perron formula, which is a form of the Fourier inversion formula, one can obtain (essentially as Riemann did) the formula

$$\psi(x) = \frac{1}{2i\pi} \int_{(3)} -\frac{\zeta'(s)}{\zeta(s)} x^s \frac{ds}{s},$$

where the integral is performed on the vertical line with real part 3, oriented from the negative imaginary part to the positive; here, the number 3 can be replaced by any number > 1.

Now suppose that we knew the existence of a real number $0 < c < 1$ with the property that all zeros of the zeta function have real part $< c$ (and moreover assume that we know

that the zeta function does not grow "too fast" when the imaginary part is large). Then by applying the Cauchy integral formula along a rectangle with sides parallel to the axes, with horizontal parts at heights $\pm \mathrm{T}$, and vertical parts having real parts c and 3, and letting $\mathrm{T} \to +\infty$, we obtain

$$\psi(x) = x + \frac{1}{2i\pi} \int_{(c)} -\frac{\zeta'(s)}{\zeta(s)} x^s \frac{ds}{s} = x + \mathrm{O}(x^c) \tag{5}$$

for any $x \geqslant 2$, at least if the integral on the line with real part c converges. Now we see that if $c < 1$, this formula determines the asymptotic behavior of $\psi(x)$. The *Prime Number Theorem* then follows easily: we have

$$\lim_{x \to +\infty} \frac{\pi(x)}{x/\log(x)} = 1, \tag{6}$$

but the formula (5) is much more precise.

Riemann then considers the possibility that all zeros within the region $0 < \mathrm{Re}(s) < 1$ have real part $1/2$,[4] which is the best possible result because there are indeed zeros with real part $1/2$ (Riemann himself had computed numerical approximations of the first ones). Then this sketch would establish the fact that

$$\pi(x) = \int_2^x \frac{dt}{\log t} + \mathrm{O}(x^{1/2+\varepsilon})$$

for any $x \geqslant 2$ and any $\varepsilon > 0$, where the implied constant depends only of ε.

However, this hypothesis of Riemann remains unproved, and it is also unknown if this straightforward approach is applicable: one doesn't know any allowable value $c < 1$. It was therefore a major progress when Hadamard and de la Vallée Poussin proved independently in 1896 that the Riemann zeta function does not vanish for $\mathrm{Re}(s) \geqslant 1$, and succeeded in deriving the validity of the Prime Number Theorem (6), although not in a precise form of the shape (5).

3.3. **Why should one consider the zeta function probabilistically?** One might at first consider a holomorphic or meromorphic function, such as the zeta function, to be a very regular and deterministic object. However, the behavior of the Riemann zeta function on the part of the complex plane which is most relevant for number theory (namely when $0 \leqslant \mathrm{Re}(s) \leqslant 1$) is extremely complex.

For instance, Riemann had already correctly stated that the number of zeros $\varrho = \beta + i\gamma$ of ζ such that $0 \leqslant \beta \leqslant 1$ and $|\gamma| \leqslant \mathrm{T}$ is, as $\mathrm{T} \to +\infty$, asymptotic to $\pi^{-1}\mathrm{T}\log(\mathrm{T})$. In particular, in a box of height 1 where the imaginary part belongs to $[\mathrm{T}, \mathrm{T}+1]$, the number of zeros is on average approximately equal to $\pi^{-1}\log(\mathrm{T})$, and thus increases as T increases. If we assume for the moment the Riemann Hypothesis in order to simplify the discussion, this means that the function $t \mapsto \zeta(1/2 + it)$ must be highly oscillatory. Since it is also known that it is unbounded (for instance, because it is known since the work of Hardy and Littlewood that

$$\frac{1}{2\mathrm{T}} \int_{-\mathrm{T}}^{\mathrm{T}} |\zeta(\tfrac{1}{2} + it)|^2 dt \sim \log \mathrm{T}$$

[4] "es ist sehr wahrscheinlich".

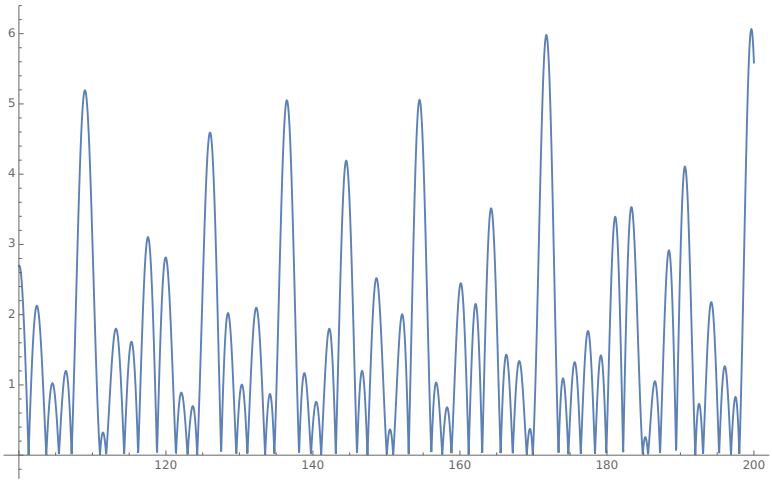

FIGURE 2. Plot of $|\zeta(\frac{1}{2}+it)|$ for $100 \leqslant t \leqslant 200$.

as $\mathrm{T} \to +\infty$), it follows that it is practically impossible to "guess" the value of $\zeta(1/2+it)$ when t is a large "random" real number, unless one actually computes the value approximately. Figure 2 illustrates this point: it plots the modulus $|\zeta(\frac{1}{2}+it)|$ for $t \in [100, 200]$.

It is then understandable, and one can highly recommend, to consider the variations of the values of the zeta function in a *probabilistic* manner, to attempt to understand its "average" behavior. This is also quite natural from the point of view of arithmetic applications, since it is rather rare that an individual value of ζ plays a deciding factor. It will usually be a question of some integral involving the zeta function, or of the location of a whole set of zeros, and not just one. Indeed, in many cases, one can work around the lack of proof of the Riemann Hypothesis by exploiting this feature. (The first spectacular instance of this is the proof by Hoheisel that there exists $c < 1$ such that there is always a prime between x and $x + x^c$ for all large enough values of x; a straighforward attempt would seem to require that all zeros of zeta have real part $\leqslant c$, but it was realized that "sufficiently few" potential exceptions could be accommodated for such a result; see, e.g. [11, Ch. 10]).

The most important probabilistic result concerning the Riemann zeta function is due to Selberg [24] in 1946.[5] For simplicity we only consider the modulus of zeta, although Selberg proved a similar result (of equal importance) for the argument.

Theorem 3.2 (Selberg)**.** *For any real number* $\mathrm{T} \geqslant 1$*, define a random variable on the interval* $[-\mathrm{T}, \mathrm{T}]$*, with the normalized Lebesgue measure* $\frac{1}{2\mathrm{T}}dt$*, by the formula*

$$t \mapsto \begin{cases} \log|\zeta(1/2+it)|/\sqrt{\frac{1}{2}\log\log \mathrm{T}} & \text{if } \zeta(1/2+it) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Then these random variables converge in law to a standard gaussian as $t \to +\infty$*.*

[5] It is not the first: here, the previous work of Bohr and Jessen and others on the values $\zeta(\sigma+it)$ for $\sigma > \frac{1}{2}$ takes precedence.

Concretely, denoting by λ the Lebesgue measure, this means that for all real numbers $a < b$, the following limit exists and takes the stated value:

$$\lim_{\mathrm{T}\to+\infty} \frac{1}{2\mathrm{T}}\lambda\Big(\Big\{t\in[-\mathrm{T},\mathrm{T}]\mid a<\frac{\log|\zeta(1/2+it)|}{\sqrt{\frac{1}{2}\log\log\mathrm{T}}}<b\Big\}\Big)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}dt. \tag{7}$$

Selberg proved this result by computing asymptotically the moments of the logarithm of the Riemann zeta function on the critical line, i.e., by showing that for any integer $k \geqslant 0$, we have

$$\lim_{\mathrm{T}\to+\infty}\frac{1}{2\mathrm{T}}\int_{-\mathrm{T}}^{\mathrm{T}}\Big(\frac{\log|\zeta(1/2+it)|}{\sqrt{\frac{1}{2}\log\log\mathrm{T}}}\Big)^k dt=\begin{cases}\frac{k!}{2^{k/2}(k/2)!} & \text{if } k \text{ is even,}\\ 0 & \text{if } k \text{ is odd,}\end{cases}$$

where the right-hand side is well-known to equal $\mathbf{E}(\mathrm{N}^k)$ for a standard gaussian random variable N, i.e.

$$\mathbf{E}(\mathrm{N}^k)=\frac{1}{\sqrt{2\pi}}\int_{\mathbf{R}} t^k e^{-t^2/2}dt.$$

Theorem 3.2 then follows from the *method of moments*. At that time, it doesn't seem that Selberg was aware of the probabilistic interpretation of his result.

Modern proofs (such as that of Radziwiłł and Soundararajan [21]) emphasize the probabilistic nature of the result by identifying an approximation of the zeta function by an object which is naturally close to a gaussian, as we will now explain intuitively. But it should be emphasized that this theorem is quite a bit deeper and more difficult to prove than the Erdős–Kac Theorem, for instance.

Theorem 3.2 is the fundamental statement that justifies the use of certain heuristic arguments concerning the distribution of values of the Riemann zeta function on the critical line. For instance, since the normalizing factor $\sqrt{\frac{1}{2}\log\log\mathrm{T}}$ tends to infinity as T does, we see that most values of $\zeta(1/2+it)$ are either very large, or very small: taking, for instance, $a=-1/10$ and $b=1/10$ in (7), it follows that for at least 92% of the real numbers $t\in[-\mathrm{T},\mathrm{T}]$, we have either

$$|\zeta(1/2+it)|\geqslant\exp(\tfrac{1}{10}\sqrt{\tfrac{1}{2}\log\log\mathrm{T}}),$$

or

$$|\zeta(1/2+it)|\leqslant\exp(-\tfrac{1}{10}\sqrt{\tfrac{1}{2}\log\log\mathrm{T}}).$$

In the first case, the zeta function is "very large", and in the second, it is "very small", since these values also tend to $+\infty$ or to 0 as $\mathrm{T}\to+\infty$.

3.4. **Justifiying Selberg's Theorem.** We can provide a heuristic for Selberg's Theorem which is very natural from the probabilistic point of view.

Step 1. Although the Euler product (4) diverges on the line $\mathrm{Re}(s)=\frac{1}{2}$, one can still hope to preserve an approximation of some kind of $\zeta(\frac{1}{2}+it)$ by a partial product

$$\zeta(1/2+it)\approx\prod_{p\leqslant\mathrm{P}}(1-p^{-1/2-it})^{-1},$$

for $|t|\leqslant\mathrm{T}$ and some suitable value of P; such an approximation is quite delicate, and can only hold in a probabilistic sense, and the value of P should be roughly such that $\log\mathrm{P}\sim\log\mathrm{T}$.

Step 2. Using the Taylor series of $\log(1-z)$ around $z=0$, we then expect an approximation of the kind

$$\log|\zeta(1/2+it)| \approx \mathrm{Re}\Big(\sum_{p\leqslant \mathrm{P}} p^{-1/2-it}\Big),$$

because the contributions of higher order (involving p^{-1-it}, $p^{-3/2-it}$, etc) are negligible; this is because the corresponding series are either absolutely convergent or almost so.

Step 3. We can then use the following result, which provides in this case the crucial connection between arithmetic and probability (it plays roughly the same role as Theorem 2.1 in the Erdős–Kac Theorem):

Theorem 3.3. *Let $\mathbf{T}$ be the infinite product of copies of the unit circle in $\mathbf{C}$, indexed by primes. For any real number $\mathrm{T}\geqslant 1$, define a random variable $\mathrm{X_T}$ on the interval $[-\mathrm{T},\mathrm{T}]$, given with the normalized Lebesgue measure $\frac{1}{2\mathrm{T}}\lambda$, taking values in $\mathbf{T}$, by the formula*

$$\mathrm{X_T}(t) = (2^{-it}, 3^{-it}, \dots) = (p^{-it})_p \in \mathbf{T}.$$

Then the random variables $\mathrm{X_T}$ converge in law as $\mathrm{T}\to+\infty$ to a random variable $\mathrm{U}=(\mathrm{U}_p)$ where the components U_p are all independent and uniformly distributed on the unit circle.

Remark. Another way of stating this is to observe that $\mathbf{T}$ is a compact abelian group, and then the limit in the theorem is also the probability Haar meassure on this group.

This result is again fairly elementary: first, one can check that it suffices to prove the formula

$$\lim_{\mathrm{T}\to+\infty} \frac{1}{2\mathrm{T}}\int_{-\mathrm{T}}^{\mathrm{T}} \varphi(\mathrm{X_T}(t))dt = \mathbf{E}(\varphi(\mathrm{U}))$$

when φ is a function on $\mathbf{T}$ of the type

$$\varphi((x_p)) = \prod_p x_p^{m_p}$$

where $m_p\in\mathbf{Z}$ for all primes, and all but finitely many of them are zero (this is a case of the "Weyl Criterion"). We then have $\varphi(\mathrm{X_T}(t)) = r^{-it}$, where r is the rational number

$$r = \prod_p p^{m_p},$$

and a straightforward integration shows that

$$\lim_{\mathrm{T}\to+\infty} \frac{1}{2\mathrm{T}}\int_{-\mathrm{T}}^{\mathrm{T}} \varphi(\mathrm{X_T}(t))dt = \begin{cases} 1 & \text{if } m_p=0 \text{ for all } p\\ 0 & \text{otherwise,}\end{cases}$$

because $r=1$ if and only if all m_p are zero, by the uniqueness of prime factorization, which turns out again to be the key arithmetic fact here. Another simple computation shows that this coincides indeed with $\mathbf{E}(\varphi(\mathrm{U}))$.

Step 4. Based on this result and on Step 2, it is natural to expect that, probabilistically speaking, $\log|\zeta(1/2+it)|$ will be, for $|t|\leqslant \mathrm{T}$, "close" to $\mathrm{Re}(\mathrm{Y_P})$, where

$$\mathrm{Y_P} = \sum_{p\leqslant \mathrm{P}} \frac{\mathrm{U}_p}{\sqrt{p}},$$

and the random variables (U_p) are, as in Step 3, independent and uniformly distributed on the unit circle.

Step 5. The question of the behavior of the random variable Y_P belongs to standard probability theory: the theory of sums of independent random variables, the realm of the Central Limit Theorem. The expectation of Y_P is zero, since each U_p has expectation zero. The variance $\sigma_P = \mathbf{E}(Y_P^2)$ of Y_P satisfies

$$\sigma_P^2 = \sum_{p \leqslant P} \frac{1}{p}$$

by independence. We have already stated that

$$\sigma_P^2 = \sum_{p \leqslant P} \frac{1}{p} \sim \log\log P \tag{8}$$

as $P \to +\infty$, since this was an important fact in the proof of the Erdős–Kac Theorem. We can now observe that this follows easily[6] from the the Prime Number Theorem[7].

Since the variance tends to $+\infty$ as $P \to +\infty$, a suitable version (e.g., Lyapunov's) of the Central Limit Theorem implies that Y_P/σ_P converges in law to a standard complex gaussian. Hence the real part of Y_P/σ_P converges in law to the real part of such a random variable, which is a centered (real) gaussian random variable with variance $1/2$. This means that $\mathrm{Re}(Y_P/\frac{1}{\sqrt{2}}\sigma_P)$ converges to a standard gaussian. Hence, assuming that the approximations we did in the previous steps can be justified, we obtain Selberg's Theorem.

3.5. **Other links between the zeta function and probability theory.** We will finish this section with a few additional illustrations of probabilistic aspects of the Riemann zeta function. We also wish to indicate that the Riemann zeta function is by now well-understood to be one example (the simplest) of an important class of similar functions known as L-functions, which are associated to a variety of important arithmetic objects (number fields, algebraic varieties, automorphic forms, etc), and whose properties, be they known or conjectural, encode some of the deepest and most fascinating aspects of modern number theory.

Bagchi's Theorem and Voronin's Theorem. Selberg's Theorem deals with the probabilistic study of the Riemann zeta function on the "critical line" $\mathrm{Re}(s) = \frac{1}{2}$; it is of course of some interest to study what happens in other regions, and Bagchi proved a rather beautiful statement (which, however, is much easier than Selberg's). Consider a closed disc $D \subset \mathbf{C}$ centered at $3/4$ and with radius $r < 1/4$, so that D is contained in the vertical strip $\{s \in \mathbf{C} \mid 1/2 < \mathrm{Re}(s) < 1\}$. Let H(D) denote the Banach space of complex-valued functions f defined on D which are holomorphic in the interior of D, with the norm defined by $\|f\| = \sup_{d \in D} |f(z)|$.

Now for $T \geqslant 1$, define an H(D)-valued random variable on $[-T, T]$, by sending $t \in [-T, T]$ to the function obtained by translating the zeta function by t, i.e., the function associated to t is the holomorphic function $s \mapsto \zeta(s + it)$ on D.

[6] By "summation by parts".

[7] Although this is a simpler result, which was proved by Mertens using elementary methods, well before the proof of the Prime Number Theorem.

Bagchi's Theorem is that these random variables converge in law as $T \to +\infty$ to the random Dirichlet series

$$\sum_{n \geqslant 1} \frac{U_n}{n^s} = \prod_p (1 - U_p p^{-s})^{-1},$$

whose coefficients U_n are defined by multiplicativity starting from a sequence (U_p) with the properties of Theorem 3.3, namely

$$U_n = \prod_p U_p^{n_p} \quad \text{for} \quad n = \prod_p p^{n_p}.$$

Moreover, by computing the support of this random Dirichlet series, Bagchi recovered the so-called "universality theorem" of Voronin: for any function $f \in H(D)$ which does not vanish in the interior of D, and for all $\varepsilon > 0$, we have

$$\liminf_{T \to +\infty} \frac{1}{2T} \lambda(\{t \in [-T, T] \mid \|\zeta(\cdot + it) - f(\cdot)\| < \varepsilon\}) > 0.$$

Conjectural links with Random Matrix Theory. Besides the probabilistic model suggested by Theorem 3.2 (namely, for $\zeta(\frac{1}{2} + it)$ when $|t| \leqslant T$, and at least for the modulus of $\zeta(s)$, a centered gaussian with variance $\frac{1}{2} \log \log T$), there is a remarkable body of evidence (including numerical evidence) that relates the distribution and properties $\zeta(\frac{1}{2}+it)$ to random unitary matrices of large size N, elements of the unitary group $U_N(\mathbf{C})$ distributed according to the probability Haar measure on this compact group, where $N \sim \log T$. The nature of this approximation (if it is true) remains very mysterious, but it leads for instance (due to work of Keating and Snaith) to very precise conjectures for the moments

$$\frac{1}{2T} \int_{-T}^{T} |\zeta(1/2 + it)|^k dt,$$

for $k > 0$ real (or even $k \in \mathbf{C}$ with $\mathrm{Re}(k) \geqslant 0$). The Bourbaki seminar of Ph. Michel [19] is an excellent survey of these conjectural relations.

Links with gaussian multiplicative chaos (and other probabilistic objects). Some of the deepest works of probabilistic number theory in recent years have been devoted to studies of finer aspects of the distribution of the Riemann Zeta function on the critical line. A particular focus has been a conjecture of Fedorov, Hiary and Keating that addresses the distribution of the maximum of $\zeta(1/2+it)$ when t varies over an interval of length 1 (and t is viewed as taken uniformly at random in $[-T, T]$ or $[T, 2T]$ with $T \to +\infty$). This leads to links with objects like log-correlated fields, branching random walks, or gaussian multiplicative chaos. We refer to the Bourbaki seminar survey of Harper [8] for a discussion of the work of Najnudel and Arguin–Belius–Bourgade–Radziwiłł–Soundararajan, and to Harper's recent preprint [9] for the latest developpments.

4. Kloosterman paths

Our last example of probabilistic number theory is a recent result due to W. Sawin and the author [18], which concerns very different arithmetic and probabilistic objects, and whose proof also involves fascinating connections with deep aspects of arithmetic geometry, in the form of Deligne's strongest form of the Riemann Hypothesis over finite fields [2], and its

subsequent and equally remarkable elaborations in the works of Katz. Moreover, this result leads to intriguing figures like those in Figures 3 and 4.

The arithmetic objects we consider are finite *exponential sums*, precisely the *Kloosterman sums* $\mathrm{Kl}(a,b;p)$ modulo primes. To define them, we first introduce the notation $e(z) = \exp(2i\pi z)$ for any complex number z. Fix a prime p, and integers a and b coprime to p. For any integer x coprime to p, denote by $\bar{x}$ the inverse of x modulo p, i.e., the residue class modulo p such that $x\bar{x} \equiv 1 \pmod{p}$. Observe that $e(\bar{x}/p)$ is well-defined, since changing the representative of $\bar{x}$ in $\mathbf{Z}$ changes the exponential by a factor $e(k) = 1$ for some integer $k \in \mathbf{Z}$. Then let

$$\mathrm{Kl}(a,b;p) = \frac{1}{\sqrt{p}} \sum_{1 \leqslant x \leqslant p-1} e\Big(\frac{ax+b\bar{x}}{p}\Big) \tag{9}$$

(where the presence of the normalizing factor $1/\sqrt{p}$ will soon be explained).

The importance of these exponential sums in analytic number theory (whether solving diophantine equations, or studying prime numbers in arithmetic progressions, or automorphic forms) cannot be overstated. We refer to [14, 10] for surveys that present many of the remarkable applications of these sums; we can however also quite simply motivate them (to some extent) by observing that the integral

$$\int_0^{+\infty} \exp(-ax - b/x)dx$$

which is a natural "continuous" analogue of (9), is none other than a Bessel function, whose importance in physics and applied mathematics is well-known.

The Kloosterman sum is a simple-looking object – a finite sum of roots of unity. Nevertheless, as Figure 3 shows, the summation process is extremely complicated: this shows, for $a = b = 1$ and $p = 10007$, the "path" taken when summing the terms

$$e\Big(\frac{x+\bar{x}}{10007}\Big)$$

(normalized by $1/\sqrt{10007}$) over x from $x = 1$ to $x = 10006$. Precisely, the vertices of the "polygon" are the successive partial sums

$$\frac{1}{\sqrt{10007}} \sum_{1 \leqslant x \leqslant y} e\Big(\frac{x+\bar{x}}{10007}\Big),$$

for $0 \leqslant y \leqslant 10006$, joined by line segments.

This picture obviously evokes some random happenings, and the paper [18] describes precisely in what sense this is indeed a random process. To state it, we first define for each p and a, b invertible modulo p a continuous function

$$\mathcal{K}_p(a,b)\colon [0,1] \to \mathbf{C}$$

such that, for all integers $0 \leqslant y \leqslant p$, we have

$$\mathcal{K}_p(a,b)\Big(\frac{y}{p}\Big) = \frac{1}{\sqrt{p}} \sum_{1 \leqslant x \leqslant y} e\Big(\frac{ax+b\bar{x}}{p}\Big),$$

and such that the values in an interval $y/p \leqslant t \leqslant (y+1)/p$ are obtained by linear interpolation. Thus the image of $\mathcal{K}_p(a,b)$ is the path represented on pictures like Figure 3. We view the map $(a,b) \mapsto \mathcal{K}_p(a,b)$ as a random variable, defined on the (finite) probability space

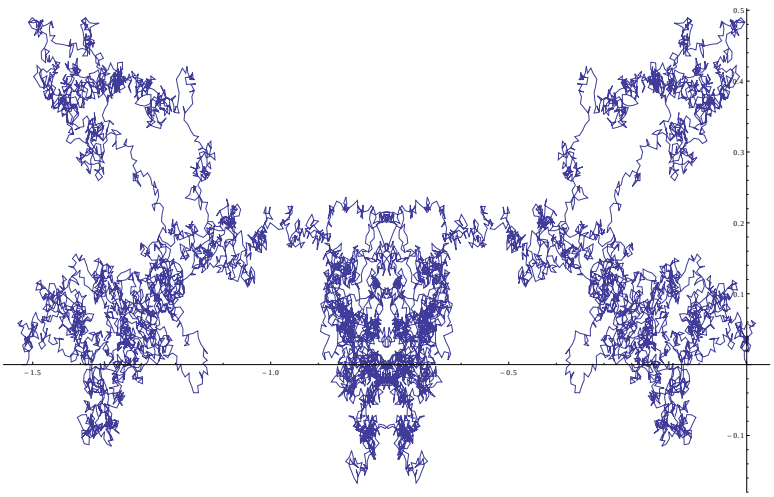

FIGURE 3. The Kloosterman path for $p = 10007$ and $a = b = 1$.

$\mathbf{F}_p^\times \times \mathbf{F}_p^\times$ (with uniform probability measure) and with values in the Banach space $\mathcal{C}([0,1])$ of continuous complex-valued functions on the interval $[0,1]$.

Theorem 4.1 (Kowalski–Sawin)**.** *Let $(\mathrm{S}_h)_{h\in\mathbf{Z}}$ be a sequence of independent random variables with values in $[-2,2]$ and distributed according to the so-called Sato–Tate measure*

$$\frac{1}{\pi}\sqrt{1-\frac{x^2}{4}}dx.$$

Define the random Fourier series

$$\mathrm{X}(t) = t\mathrm{S}_0 + \sum_{\substack{h\in\mathbf{Z}\\ h\neq 0}} \frac{e(ht)-1}{2i\pi h}\mathrm{S}_h, \tag{10}$$

where the sum over h is understood in the sense of the limit as $\mathrm{H}\to+\infty$ of the symmetric partial sums over $|h|\leqslant \mathrm{H}$.

Then X defines a $\mathcal{C}([0,1])$-valued random variable[8] *and the random variables $\mathcal{K}_p$ converge in law to X as $p\to+\infty$.*

To illustrate the behavior of the random Fourier series X, Figure 4 presents an approximation to a sample of this series (obtained by computing the partial sum over $|h|\leqslant 10000$, with random coefficients simulating the sequence (S_h)).

As we have done before, we will attempt to motivate this result. We begin by the explanation of the normalizing factor: it is already a deep result, proved by A. Weil in 1948 as consequence of the Riemann Hypothesis for curves over finite fields, that individual Kloosterman sums satisfy $|\mathrm{Kl}(a,b;p)|\leqslant 2$ (furthermore, they are real-valued, which is easy to prove by computing the complex conjugate), and that one cannot replace the constant 2 by any function of p that tends to 0. Thus the Kloosterman sums may also be studied probabilistically.

[8] I.e., the series converges almost surely uniformly on $[0,1]$.

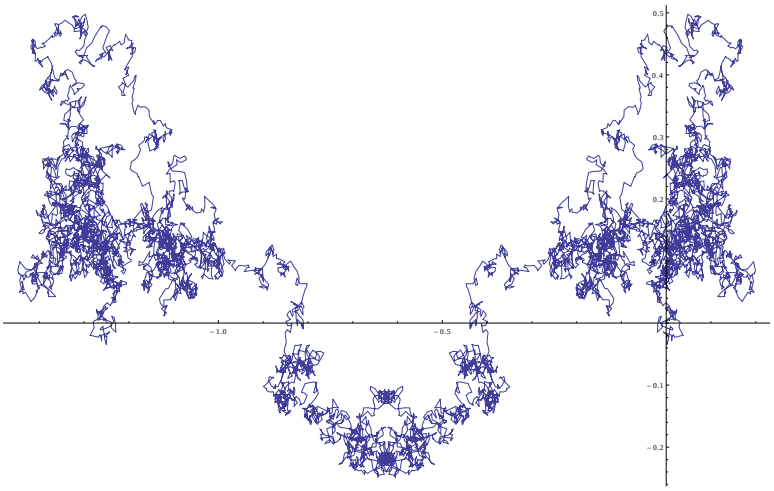

FIGURE 4. A sample of the random Fourier series.

We look at the partial sum over $1 \leqslant x \leqslant y$, and apply a standard principle of harmonic analysis, the *completion method*. This means that we expand the characteristic function φ_y of the summation interval in discrete Fourier series modulo p:

$$\varphi_y(x) = \sum_{-p/2<h\leqslant p/2} \alpha_y(h;p) e\Big(\frac{hx}{p}\Big),$$

where

$$\alpha_y(h;p) = \frac{1}{p} \sum_{1\leqslant x\leqslant y} e\Big(-\frac{hx}{p}\Big),$$

and then input this expression in the partial sum, obtaining

$$\begin{aligned} \frac{1}{\sqrt{p}} \sum_{1\leqslant x\leqslant y} e\Big(\frac{ax+b\bar{x}}{p}\Big) &= \frac{1}{\sqrt{p}} \sum_{-p/2<h\leqslant p/2} \alpha_y(h;p) \sum_{1\leqslant x\leqslant p-1} e\Big(\frac{(a+h)x+b\bar{x}}{p}\Big) \\ &= \sum_{-p/2<h\leqslant p/2} \alpha_y(h;p)\, \mathrm{Kl}(a+h,b;p). \end{aligned} \tag{11}$$

By a simple interpretation as Riemann sums, we see easily that $\alpha_y(0;p) = y/p$ and

$$\lim_{p\to+\infty} \alpha_y(h;p) = \int_0^{y/p} e(ht)dt = \frac{e(hy/p)-1}{2i\pi h}$$

for $h \neq 0$. This means that the right-hand side of (11) is reminiscent of the right-hand side of (10), evaluated at $t = y/p$, with the random variables S_h replaced by the random variables on $\mathbf{F}_p^\times \times \mathbf{F}_p^\times$ defined by

$$(a,b) \mapsto \mathrm{Kl}(a+h,b;p).$$

Theorem 4.1 is then quite understandable from a result which is essentially a consequence of the deep work of Katz [12] on the distribution of Kloosterman sums, and which we state somewhat imprecisely: the family $(\mathrm{Kl}(a+h,b;p))_h$ "converges in law" to a family of independent Sato–Tate distributed random variables. Precisely, the fact that, for fixed h, the random variables $\mathrm{Kl}(a+h,b;p)$ converge in law to the Sato–Tate measure was proved

by Katz [12], and in another work [13], he developed the basic principles to derive that the different shifted Kloosterman sums are asymptotically independent.

5. Conclusion

We hope that this survey will have illustrated the beautiful connections that exist between number theory and probability theory. As we mentioned at the beginning, this is currently a subject in great expansion, and it is especially remarkable to see how the links between the two subjects are becoming tighter and the interactions richer. One may expect that a lot of beautiful mathematics will arise from these links in the coming years!

Appendix A. Probabilistic notions

A.1. **Probability spaces, random variables, independence.** A *probability space* is a triple $(\Omega, \Sigma, \mathbf{P})$ where Ω is a set, Σ is a σ-algebra on Ω and $\mathbf{P}$ is a probability measure on Σ (i.e., a positive measure such that $\mathbf{P}(\Omega) = 1$).

Given a topological space T, a random variable X on Ω with values in T is a map $\mathrm{X}\colon \Omega \to \mathrm{T}$ which is measurable, when T is equiped with the Borel σ-algebra: for any open set U in T, the inverse image $\mathrm{X}^{-1}(\mathrm{U})$ is in Σ. The *law* of X is the probability measure μ_{X} on T image of $\mathbf{P}$ by X. We typically write $\mathbf{P}(\mathrm{X} \in \mathrm{A})$ for $\mu_{\mathrm{X}}(\mathrm{A})$. If a set A satisfies $\mathbf{P}(\mathrm{X} \in \mathrm{A}) = 1$, then one says that the event that A represents holds *almost surely*.

We denote by $\mathbf{E}(\mathrm{X})$ the expectation and $\mathbf{V}(\mathrm{X})$ the variance for a complex-valued random variable, when it exists:

$$\mathbf{E}(\mathrm{X}) = \int_{\Omega} \mathrm{X} d\,\mathbf{P}, \qquad \mathbf{V}(\mathrm{X}) = \mathbf{E}(|\mathrm{X} - \mathbf{E}(\mathrm{X})|^2).$$

The precise nature of the "sample space" Ω is frequently left unspecified – what matters is that one can define suitable families of random variables with various properties, especially independence, whose meaning we now recall.

Definition A.1. Given an arbitrary, possibly infinite, set I, a family $(\mathrm{X}_i)_{i\in \mathrm{I}}$ of T-valued random variables is independent if for any finite subset $\mathrm{J} \subset \mathrm{I}$, and any open sets $\mathrm{U}_j \subset \mathrm{T}$ for $j \in \mathrm{J}$, we have

$$\mathbf{P}(\mathrm{X}_j \in \mathrm{A}_j \text{ for } j \in \mathrm{J}) = \prod_{j\in \mathrm{J}} \mathbf{P}(\mathrm{X}_j \in \mathrm{A}_j).$$

A basic existence theorem that suffices for many applications of probability theory is then that given any sequence (μ_n) of probability measures on $\mathbf{R}$, say, there exists a probability space $(\Omega, \Sigma, \mathbf{P})$ and a family (X_n) of real-valued random variables on Ω such that the law of X_n is μ_n and the (X_n) are independent.

A.2. **Special random variables.** Let p be a real number such that $0 \leqslant p \leqslant 1$. A *Bernoulli random variable* with success probability p is a real-valued random variable such that

$$\mathbf{P}(\mathrm{B} = 1) = p, \qquad \mathbf{P}(\mathrm{B} = 0) = 1 - \frac{1}{p}$$

(in other words, the map $\mathrm{B}\colon \Omega \to \{0, 1\}$ is the characteristic function of a set which has measure p).

A real-valued random variable X is a *gaussian random variable* with expectation $m \in \mathbf{R}$ and variance $\sigma^2 > 0$ if

$$\mathbf{P}(\mathrm{X} \in \mathrm{U}) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathrm{A}} e^{-(x-m)^2/2\sigma^2} dx$$

for all open sets U. If $m = 0$, then we say that X is *centered*.

A.3. **Some useful statements.** We use the following lemma in Section 2.1.

Lemma A.2. *Let* M *be a finite-dimensional Banach space. Let* $(\mathrm{X}_n)_{n\geqslant 1}$ *and* $(\mathrm{X}_{n,m})_{n\geqslant m\geqslant 1}$ *be* M*-valued random variables. Define* $\mathrm{E}_{n,m} = \mathrm{X}_n - \mathrm{X}_{n,m}$*. Assume that*

(1) *For each* $m \geqslant 1$*, the random variables* $(\mathrm{X}_{n,m})_{n\geqslant m}$ *converge in law to a random variable* Y_m*.*

(2) *There exists a function* $f(n, m)$ *such that*

$$\mathbf{E}(\|\mathrm{E}_{n,m}\|) \leqslant f(n, m)$$

and $f(n, m) \to 0$ *as* m *tends to* $+\infty$ *uniformly for* $n \geqslant m$*.*

Then the sequences (X_n) *and* (Y_m) *converge in law as* $n \to +\infty$*, and have the same limit distribution.*

See [1, Th. 3.2] (which assumes that one knows beforehand that (Y_m) converges in law) or [16, Prop. B.4.4] for the proof.

We also referred to Kolmogorov's Three Series Theorem (see, e.g. [1, Th. 22.8]):

Theorem A.3. *Let* $(\mathrm{X}_n)n \geqslant 0$ *be a sequence of* independent *real-valued random variables on* Ω*. Define*

$$\widetilde{\mathrm{X}}_n = \begin{cases} \mathrm{X}_n & \text{if } |\mathrm{X}_n| \leqslant 1 \\ 0 & \text{otherwise.} \end{cases}$$

The series

$$\sum_{n\geqslant 0} \mathrm{X}_n$$

converges almost surely if and only if the three series

$$\sum_{n\geqslant 0} \mathbf{E}(\widetilde{\mathrm{X}}_n), \qquad \sum_{n\geqslant 0} \mathbf{V}(\widetilde{\mathrm{X}}_n^2), \qquad \sum_{n\geqslant 0} \mathbf{P}(|\mathrm{X}_n| > 1)$$

are convergent.

References

[1] P. Billingsley: *Convergence of probability measures*, 2nd edition, Wiley, 1999.

[2] P. Deligne: *La conjecture de Weil, II*, Publ. Math. IHÉS 52 (1980), 137–252.

[3] P. Erdős and M. Kac: *The Gaussian law of errors in the theory of additive number theoretic functions*, Amer. J. Math. 62 (1940), 738–742.

[4] K. Ford: *The distribution of integers with a divisor in a given interval*, Annals of Math. 168 (2008), 367–433.

[5] A. Granville: *The anatomy of integers and permutations*, preprint (2008), http://www.dms.umontreal.ca/~andrew/PDF/Anatomy.pdf

[6] A. Granville and J. Granville: *Prime suspects*, Princeton Univ. Press, 2019; illustrated by R.J. Lewis.

[7] A. Harper: *Two new proofs of the Erdős–Kac Theorem, with bound on the rate of convergence, by Stein's method for distributional approximations*, Math. Proc. Camb. Phil. Soc. 147 (2009), 95–114.

[8] A. Harper: *The Riemann zeta function in short intervals*, Séminaire N. Bourbaki, 71ème année, exposé 1159, March 2019.
[9] A. Harper: *On the partition function of the Riemann zeta function, and the Fyodorov–Hiary–Keating conjecture*, preprint (2019); arXiv:1906.05783.
[10] D.R. Heath–Brown: *Arithmetic applications of Kloosterman sums*, Nieuw Archief voor Wiskunde 5 (2000), 380–384.
[11] H. Iwaniec and E. Kowalski: *Analytic number theory*, Colloquium Publ. 53, A.M.S 2004.
[12] N.M. Katz: *Gauss sums, Kloosterman sums and monodromy groups*, Annals of Math. Studies 116, Princeton Univ. Press (1988).
[13] N.M. Katz: *Exponential sums and differential equations*, Annals of Math. Studies 124, Princeton Univ. Press (1990).
[14] E. Kowalski: *Poincaré and analytic number theory*, in "The scientific legacy of Poincaré", edited by É. Charpentier, É. Ghys and A. Lesne, A.M.S, 2010.
[15] E. Kowalski: *Crible en expansion*, Séminaire Bourbaki, exposé 1028, November 2010, Astérisque 348, Soc. Math. France (2012), 17–64; English version at https://www.math.ethz.ch/~kowalski/sieve-expansion-bourbaki.pdf
[16] E. Kowalski: *Arithmetic randonnée: an introduction to probabilistic number theory*, lecture notes www.math.ethz.ch/~kowalski/probabilistic-number-theory.pdf; Cambridge Studies in Advanced Mathematics, to appear.
[17] E. Kowalski and A. Nikeghbali: *Mod-Poisson convergence in probability and number theory*, International Mathematics Research Notices 2010; doi:10.1093/imrn/rnq019
[18] E. Kowalski and W. Sawin: *Kloosterman paths and the shape of exponential sums*, Compositio Math. 152 (2016), 1489–1516.
[19] Ph. Michel: *Répartition des zéros des fonctions* L *et matrices aléatoires*, Séminaire N. Bourbaki, 53ème année, exposé 887, March 2001; in Astérisque 282 (2002), 211–248.
[20] C. Perret-Gentil: *Some recent interactions of probability and number theory*, in Newsletter of the European Mathematical Society, March 2019.
[21] M. Radziwiłł and K. Soundararajan: *Selberg's central limit theorem for* $\log|\zeta(1/2+it)|$, L'enseignement Mathématique 63 (2017), 1–19.
[22] Riemann, B.: *Über die Anzahl der Primzahlen unter einer gegebenen Grösse*, Monatsber. Berlin. Akad. 1859, 671–680, or in *Gesammelte mathematische Werke, wissenschaftlicher Nachlass und Nachtrage*, edited by R. Narasimhan, Springer-Verlag, Berlin, 1990, 177–185.
[23] I. Schoenberg: *Über die asymptotische Verteilung reeller Zahlen mod* 1, Math. Z. 28 (1928), 171–199.
[24] A. Selberg: *Contributions to the theory of the Riemann zeta function*, Arch. Math. Naturvid. 48 (1946), 89–155; or in *Collected works*, I.
[25] E.C. Titchmarsh: *The theory of the Riemann zeta function*, 2nd edition, Oxford Univ. Press, 1986.

ETH Zürich – DMATH, Rämistrasse 101, 8092 Zürich, Suisse
Email address: kowalski@math.ethz.ch