The Riemann zeta function and probability theory

Phillipe Biane

February 13, 2015

Abstract

Translation of “La Fonction zéta de Riemann et les probabilités” by Phillipe Biane. La fonction zêta, pp. 165–193 .

1 Introduction

In recent years, particularly under the influence of the physicist and number theorist F.J. Dyson, probabilistic ideas have penetrated the field of number theory, particularly the study of the Riemann zeta function. Without attempting to provide a comprehensive overview of the relationship between probability and number theory, I will try to explain two examples of how these seemingly distant areas are closely related.

The first example we consider is the theory of random matrices and its applications to the study of zeros of the Riemann zeta function. The origin of number theorists' interest in random matrix theory can be traced to the work of H.L. Montgomery on the distribution of the spacings between zeros of the zeta function. Let $\rho = 1/2 + i\gamma$ denote the nontrivial zeros of the zeta function (we ignore the trivial zeros -2, -4,...). We assume the Riemann hypothesis is true: i.e, γ are real. The zeros are distributed symmetrically with respect to 1/2 (this follows from the functional equation for ζ) and the number of zeros whose imaginary part is in the interval [0, T] is asymptotic to $\frac{T}{2\pi}\log T$ as $T \to \infty$. Therefore, the average density of zeros in the interval $[0,T]$ is $\frac{1}{2\pi}\log T$, and the average spacing between two consecutive zeros is $\frac{2}{\log T}$. Montgomery was interested in the asymptotic distribution of the difference $\gamma - \gamma'$ where (γ, γ') denotes the set of pairs of zeros in the interval $[0,T]$. These differences are rescaled by $\log T$ to make the average spacing 1, and we focus on $N_{a,b}(T)$, the number of pairs (γ, γ') such that $\gamma - \gamma' \in [\frac{2\pi a}{\log T}, \frac{2\pi b}{\log T}]$. Based on calculations that will be discussed in Section 2.1, Montgomery conjectured the following asymptotic behavior

$$N_{a,b}(T) \sim \frac{T}{2\pi} \log T \left(\int_a^b 1 - \left(\frac{\sin u}{\pi u} \right)^2 du + o(1) \right), \quad 0 < a < b, \tag{1}$$

as $T \to \infty$

Montgomery was at Princeton at the same time as Dyson when he made this conjecture. In a conversation with Dyson, he was astonished to learn that the above asymptotic behavior is the same as that of the differences between the eigenvalues of a Gaussian Hermitian matrix, a result well-known to theoretical physicists. Their motivation for this problem and some subtle calculations will be explained in Section 2.2. The coincidence between Montgomery's conjecture and the physicists results on random matrices cast new on Polya and Hilbert's suggestion that the numbers γ should be the eigenvalues of a self-adjoint operator on a Hilbert space. The existence of such an operator, which would imply in particular the validity of the Riemann hypothesis, is still speculative. Nevertheless, this possibility motivated A. Odlyzko to experimentally test Montgomery's conjecture. In numerical calculations to be discussed in Section 2.3, he verified that the zeros of the zeta functions conform to the predictions of the random matrix model with high precision. We will also see how the heuristic of random matrix theory led J. Keating and N. Snaith to offer a remarkable conjecture for the asymptotic behavior of the moments of the zeta function on the critical line, a problem which dates back to the work of Hardy and Littlewood.

Despite the convincing experimental confirmation, the relationship between probability and zeta function mentioned above remain largely conjectural. This is why I have also chosen to discuss other connections. These are probably more anecdotal in terms of the Riemann zeta function, but they involve Brownian motion, probably the most important object, and certainly the most studied, in modern probability theory. We will see that the Riemann ξ function, which expresses the functional equation of the zeta function in a symmetric manner, is the Mellin transform of a probability measure that appears in the study of Brownian motion, or more specifically in the theory of Brownian excursions. This discussion gives us the opportunity to present the basics of excursion theory. We also give a probabilistic interpretation of the functional equation and explain how probabilistic reasoning leads to a natural renormalization of the series $\sum_{n=1}^{\infty}(-1)^n n^{-s}$, which converges in the whole complex plane.

2 Correlations in the zeros of the Riemann zeta function and random matrices

2.1 Montgomery's conjecture on the pair correlation

The starting point of Montgomery's work [9] is the study of the asymptotic behavior of the Fourier transform of the distribution of the differences between the γ. Montogomery considers the quantity

$$F(\alpha) = \left(\frac{T}{2\pi}\log T\right)^{-1} \sum_{0\leq\gamma,\gamma'\leq T} T^{-i\alpha(\gamma-\gamma')}\frac{4}{4+(\gamma-\gamma')^2}. \tag{2}$$

The factor $4/(4+(\gamma-\gamma')^2)$, which is used to mitigate the contribution of large deviations, is the Fourier transform of the distribution of the numbers $(\gamma-\gamma')\log T/2\pi$. If we assume the Riemann hypothesis, F takes real values and is an even function of α. Always assuming that the Riemann hypothesis is true, Montgomery showed that for every $\alpha[0,1)$ we have

$$F(\alpha) = (1+o(1))\,T^{-2\alpha}\log T + \alpha + o(1); \quad T\to\infty, \tag{3}$$

the error term being uniform for $\alpha\in[0,1-\varepsilon)$, for any $\varepsilon>0$. The proof of this result, is too technical to be presented in detail here. It is based on an "explicit formula" that links the zeros of the zeta function and the prime numbers under the assumption that the Riemann hypothesis is true. For $t\in\mathbb{R}$ and $x\geq 1$, we have

$$2\sum_{0\leq\gamma\leq T}\frac{x^{i\gamma}}{1+(t-\gamma)^2} = \tag{4}$$

$$= -x^{1/2}\left(\sum_{n\leq x}\Lambda(n)\left(\frac{x}{n}\right)^{-1/2+it} + \sum_{n>x}\Lambda(n)\left(\frac{x}{n}\right)^{3/2+it}\right)$$

$$+x^{-1+it}\left(\log\left(|t|+2\right)+O(1)\right) + O\left(\frac{x^{1/2}}{|t|+2}\right) \tag{5}$$

where $\Lambda(n)$ is the arithmetic function that takes the value $\log p$ if n is divisible by the prime p, and 0 if not. Let $G(t,x)$ denote the left side of equation (4). It is easily verfied that

$$\int_0^T |G(t,T^\alpha)|^2\,dt = F(\alpha)T\log T + O(\log^3 T),$$

and (3) then becomes a delicate estimate of $\int_0^T |D(t, T^\alpha)|^2 \, dt$ where $D(t, x)$ denotes the right hand side of (4). This estimate is only possible if $\alpha \in [0, 1)$, but heuristic arguments of Montgomery (still under the Riemann hypothesis) suggest that $F(\alpha) = 1 + o(1)$ for $\alpha \geq 1$ uniformly on any compact set. This determines the asymptotic behavior of F on all of $\mathbb{R}$. The Fourier inversion formula and (2.1) imply

$$\sum_{\gamma,\gamma'} r\left((\gamma - \gamma')\,\frac{\log T}{2\pi}\right)\frac{4}{4 + (\gamma - \gamma')^2} = \frac{T}{2\pi}\log T \int_{-\infty}^{\infty} F(\alpha)\hat{r}(\alpha)\, d\alpha, \tag{6}$$

where $\hat{r}$ denotes the Fourier transform of r

$$\hat{r}(\alpha) = \int_{-\infty}^{\infty} e^{-2\pi i \alpha x} r(x)\, dx.$$

If we apply (6) to the function r defined by

$$r(u) = \begin{matrix} 1, & u \in [a, b], \\ 0, & \text{otherwise}, \end{matrix},$$

for fixed $0 < a < b$, we obtain

$$\begin{aligned} \int_{-\infty}^{\infty} \hat{r}(\alpha) T^{-2|\alpha|} \log T \, d\alpha &= \int_{-\infty}^{\infty} \hat{r}\left(\frac{\alpha}{\log T}\right) e^{-2|\alpha|}\, d\alpha \\ &= \hat{r}(0) + o(1) = b - a + o(1), \\ \int_{-\infty}^{\infty} \hat{r}(\alpha)\, d\alpha &= r(0) = 0. \end{aligned}$$

An application of Plancherel's formula then yields,

$$\int_{-\infty}^{\infty} \hat{r}(\alpha) \min(1 - |\alpha|, 0)\, d\alpha = \int_{-\infty}^{\infty} r(x) \left(\frac{\sin \pi x}{\pi x}\right)^2 dx = \int_a^b \left(\frac{\sin \pi x}{\pi x}\right)^2 dx.$$

We combine these calculations to obtain the estimate

$$F(\alpha) = 1 + o(1) T^{-2|\alpha|} \log T + 1 - \min(1 - |\alpha|, 0) + o(1).$$

We thus find equation (1).

2.2 GUE

Quantum theory implies that the energy levels of an atomic system are the eigenvalues of a Hermitian operator in Hilbert space, known as the Hamiltonian of the system. When the atomic system contains many elementary particles, there is a profusion of energy levels and the Hamiltonian is too complex to be diagonalized numerically. In this context, the physicist E. Wigner suggested that the energy levels of such a Hamiltonian can be modeled by the eigenvalues of a random Hermitian matrix. Wigner's hope was the statistical properties of energy levels, such as the distribution of their spacings, coincide with those of random matrices. Wigner's intuition proved well-founded and there is a good match between experiment and the predictions of the random matrix model in many quantum systems. See, for example, the introduction to the book of Mehta [7].

I will now explain how to describe the statistical structure of eigenvalues of a large random matrix. The term GUE - an acronym for Gaussian Unitary Ensemble - designates the space of $N \times N$ Hermitian matrices, $\mathcal{H}_N$, equipped with the standard Gaussian measure with density $(2)^{-N^2/2}\exp(-\mathrm{tr}(M^2)/2)$ with respect to the Lebesgue measure on $\mathcal{H}_N$. We now describe the law of the eigenvalues of a random matrix sampled from GUE. Any Hermitian matrix M_0 can be written in the form $U_0 X_0 U_0^*$ where U_0 is a unitary matrix and X_0 is a diagonal matrix of eigenvalues of M_0. Consider the map $(X, S) \mapsto U_0 e^{iS} X e^{-iS} U_0$ where X ranges over real diagonal matrices and S ranges over the Hermitian matrices. The differential at (X_0, S_0) of this map is $(X, S) \mapsto X + U_0[X_0, S]U_0^*$ where $[X, Y] = XY - YX$ denotes the commutator of two matrices. If M_0 is generic (all its eigenvalues are distinct), the kernel of the differential is the subspace of pairs $(0, S)$ where S is diagonal. It follows from the implicit function theorem that $(X, S) \mapsto U_0 e^{iS} X e^{-iS} U_0^*$, restricted to the subspace of S whose diagonal coefficients are zero, is a diffeomorphism between a neighborhood of $(x_0, 0)$ and a neighborhood of M_0. By identifying a pair (X, S) with the matrix $X + S$, we can calculate the eigenvalues of this change of variables, and find that the Jacobian at $(X_0, 0)$ is $\prod_{i<j}(x_i - x_j)^2$, where x_i are the eigenvalues of X_0 (and M_0). Let $x_1(M), \ldots, x_N(M)$ denote the eigenvalues associated with a matrix M. For any symmetric function of N variables $f(x_1, ..., X_N)$ we use the change of variables of the formula, to find

$$\int_{\mathcal{H}_N} f\left(x_1(M), \ldots, x_N(M)\right) \frac{e^{-\mathrm{tr}(M^2)/2}}{(2\pi)^{N^2/2}}\, dM$$
$$= \frac{1}{Z_N} \int_{\mathbb{R}^N} f(x_1, \ldots, x_N) \prod_{1 \le i < j \le N} (x_i - x_j)^2 e^{-\sum_{n=1}^N x_n^2}\, dx_1 \ldots dx_N,$$

where Z_N is a normalization constant that we calculate below. The density of the law of the eigenvalues of M is

$$P^{(N)}(x_1,\dots,x_N) = \frac{1}{Z_N} \prod_{1\le i<j\le N} (x_i - x_j)^2 e^{-\sum_{n=1}^N x_n^2}. \tag{7}$$

Knowledge of this density allows us to calculate the expectation of random variables of the type

$$N_f = \sum_{(i_1,\dots,i_n)\in[1,N]^n,\ i_j \,\text{distinct}} f(x_{i_1},\dots,x_{i_n})$$

for each bounded Borel function f. We have

$$\mathbb{E}(N_f) = \int_{\mathbb{R}^n} f(x_1,\dots,x_n)\, R_n^{(N)}(x_1,\dots,x_n)\, dx_1\dots dx_n, \tag{8}$$

where

$$\begin{aligned} &R_n^{(N)}(x_1,\dots,x_n) \\ &\quad = \frac{N!}{(N-n)!} \int_{\mathbb{R}^{N-n}} P^{(N)}(x_1,\dots,x_n,x_{n+1},\dots,x_N)\, dx_{n+1}\dots dx_N. \end{aligned}$$

In order to calculate $R_n^{(N)}$ we rewrite $P^{(N)}$ using the Vandermonde determinant

$$\prod_{i>j}(x_i - x_j) = \det\left[x_j^{i-1}\right]_{1\le i,j\le N}.$$

We take linear combinations of the columns of the matrix $\left[x_j^{i-1}\right]$ to find

$$\prod_{i>j}(x_i - x_j) = \det\left[P_{i-1}(x_j)\right]_{1\le i,j\le N}$$

for all monic polynomials P_i of degree i. We apply this identity to the Hermite polynomials defined by the recurrence relation

$$P_{n+1} = xP_n + P_{n-1}.$$

These polynomials satisfy the orthogonality relation [4]

$$\int_{\mathbb{R}} P_m(x)P_n(x)\frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx = \delta_{nm} n!$$

The normalized Hermite wave-functions defined below form an orthogonal basis for $L^2(\mathbb{R})$:

$$\varphi_n(x) = \frac{1}{\sqrt{n!}} P_n(x) \frac{e^{-x^2/4}}{(2\pi)^{1/4}}.$$

The density $P^{(N)}$ is proportional to $\left(\det\left[\varphi_{i-1}(x_j)\right]_{1\le i,j\le N}\right)^2$. To determine the constant of proportionality we evaluate the integral

$$\begin{aligned}
&\int_{\mathbb{R}^N} \left(\det\left[\varphi_{i-1}(x_j)\right]_{1\le i,j\le N}\right)^2 \, dx_1 \dots dx_N \\
&\quad = \int_{\mathbb{R}^N} \sum_{\sigma,\tau\in\Sigma_n} \varepsilon(\sigma)\varepsilon(\tau) \prod_{i=1}^N \prod_{j=1}^N \varphi_{i-1}(x_{\sigma_i})\varphi_{j-1}(x_{\tau_j})\, dx_1 \dots dx_N.
\end{aligned}$$

Since the functions φ_i are orthogonal, the only terms of the sum that give a non-zero contribution are those for which $\sigma = \tau$. Each such term gives a unti contribution. Thus, the above integral is $N!$ and

$$\begin{aligned}
P^{(N)}(x_1, \dots, x_N) \qquad &= \frac{1}{N!}\left(\det\left[\varphi_{i-1}(x_j)\right]_{1\le i,j\le N}\right)^2 \\
&= \frac{1}{N!}\det\left[K^{(N)}(x_i, x_j)\right]_{1\le i,j\le N}
\end{aligned}$$

where

$$K^{(N)}(x,y) = \sum_{k=1}^N \varphi_{k-1}(x_i)\varphi_{k-1}(x_j).$$

In light of the orthogonality of the φ_k we have

$$\int_{\mathbb{R}} K^{(N)}(x,x) = N; \quad \int_{\mathbb{R}} K^{(N)}(x,z)K^{(N)}(z,y)\, dz = K^{(N)}(x,y). \tag{9}$$

We also deduce that

$$R^{(N)}(x_1, \dots, x_N) = \det\left[K^{(N)}(x_i, x_j)\right]_{1\le i,j\le N}. \tag{10}$$

In fact, we will reason by induction that $R_n^{(N)}$ may be expressed as such a determinant for all n. Assuming this holds for $n+1$ we have

$$\begin{aligned}
R_n^{(N)}(x_1, \dots, x_n) &= \frac{1}{N-n}\int_{\mathbb{R}} R_{n+1}^{(N)}(x_1, \dots, x_n, x_{n+1})\, dx_{n+1} \\
&= \frac{1}{N-n}\int_{\mathbb{R}} \det\left[K^{(N)}(x_i, x_j)\right]_{1\le i,j\le n+1} dx_{n+1} \\
&= \frac{1}{N-n}\sum_{\sigma\in\Sigma_{n+1}} \varepsilon(\sigma)\int_{\mathbb{R}} K^{(N)}(x_1, x_{\sigma_1}) \dots K^{(N)}(x_{n+1}, x_{\sigma_{n+1}})\, dx_{n+1}.
\end{aligned}$$

If $\sigma_{n+1} = n+1$ in this sum, then the first equality in (9) implies

$$\begin{aligned}\int_{\mathbb{R}} K^{(N)}(x_1, x_{\sigma_1}) \cdots K^{(N)}(x_{n+1}, x_{\sigma_{n+1}})\, dx_{n+1} \\ = N K^{(N)}(x_1, x_{\sigma_1}) \cdots K^{(N)}(x_n, x_{\sigma_n}).\end{aligned} \tag{11}$$

If not, there exists $j \le n$ and $k \le n$ such that $\sigma_j = n+1$ and $\sigma_{n+1} = k$. We then use the second equality in (9) to find

$$\begin{aligned}\int_{\mathbb{R}} K^{(N)}(x_1, x_{\sigma_1}) \cdots K^{(N)}(x_{n+1}, x_{\sigma_{n+1}})\, dx_{n+1} \\ = K^{(N)}(x_1, x_{\sigma'_1}) \cdots K^{(N)}(x_n, x_{\sigma'_n}).\end{aligned} \tag{12}$$

where σ' is a permutation of $\{1, \ldots, n\}$ such that $\sigma'_j = k$ and $\sigma'_i = \sigma_i$ if $i \ne j$. Each permutation $\sigma' \in \Sigma_n$ may come from n permutations $\sigma \in \Sigma_{n+1}$. Thus, using equations (11) and (12) we have

$$\int_{\mathbb{R}} \det \left[K^{(N)}(x_i, x_j)\right]_{1 \le i,j \le n+1} dx_{n+1} = (N-n) \det \left[K^{(N)}(x_i, x_j)\right]_{1 \le i,j \le n}.$$

This explicit formula allows us to determine the asymptotic behavior of many statistics of the eigenvalues of GUE matrices. We will illustrate this by calculating the pair correlation. For this, we will need the asymptotic behavior of the kernel $K^{(N)}$. These asymptotics are calculated using the Christoffel-Darboux formula, which is easy to establish by induction from the recurrence relation for P_n.

$$\sum_{k=1}^{N} \varphi_{k-1}(x)\varphi_{k-1}(y) = \sqrt{N} \frac{\varphi_N(x)\varphi_{N-1}(y) - \varphi_N(y)\varphi_{N-1}(x)}{x-y}.$$

The Plancherel-Rotach formula for the asymptotics of Hermite functions (see [13]) implies that

$$\lim_{N\to\infty} \frac{1}{N} R_1^{(N)}(x\sqrt{N}) = \frac{1}{2\pi}\sqrt{4-x^2} \quad \text{if} \quad x \in [-2,2],$$

and for all continuous functions f with compact support

$$\begin{aligned}&\mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{N} f\left(\frac{x_k}{\sqrt{N}}\right)\right] \\ &\quad = \frac{1}{N}\int_{\mathbb{R}} f\left(\frac{x}{\sqrt{N}}\right) R_1^{(N)}(x)\, dx \overset{N\to\infty}{\longrightarrow} \frac{1}{2\pi}\int_{-2}^{2} f(x)\sqrt{4-x^2}\, dx.\end{aligned}$$

The limiting distribution of the eigenvalues is called the Wigner semicircle law. If we choose a small interval $(-\varepsilon_N, \varepsilon_N)$ around 0 such that $\varepsilon_N \to 0$ but $\varepsilon_N \sqrt{N} \to \infty$ the average number of eigenvalues in this interval will be of the order of $2\varepsilon_N \sqrt{N}/\pi$ and the average spacing between two of these eigenvalues will be $\pi/\sqrt{N}$. Again using the asymptotics of Hermite polynomials we arrive at the formula

$$K^{(N)}\left(\pi\xi/\sqrt{N}, \pi\eta/\sqrt{N}\right) \overset{N\to\infty}{\longrightarrow} \frac{\sin \pi(\xi-\eta)}{\pi(\xi-\eta)}.$$

This allows us to deduce that

$$\frac{\pi}{2\varepsilon_N\sqrt{N}}\mathbb{E}\left[\mathrm{card}\left\{(i,j)\,\middle|\,|x_i|,|x_j|<\varepsilon_N, i\neq j, \frac{a\pi}{\sqrt{N}} \le x_i - x_j \le \frac{b\pi}{\sqrt{N}}\right\}\right]$$
$$\overset{N\to\infty}{\longrightarrow} \int_a^b \left(1-\left(\frac{\sin \pi x}{\pi x}\right)^2\right) dx. \tag{13}$$

This formula is very different from that obtained by choosing N points X_1, $\ldots$, X_N randomly uniformly on $[0, 1]$, wit average spacing $\sim 1/N$. For such a choice, we see that

$$\frac{1}{N}\mathbb{E}\left[\mathrm{card}\left\{(i,j)\,\middle|\,\frac{a}{N} \le x_i - x_j \le \frac{b}{N}\right\}\right] \overset{N\to\infty}{\longrightarrow} b-a. \tag{14}$$

The fact that the function $1-\left(\frac{\sin \pi x}{\pi x}\right)^2$ vanishes at 0 reflects the fact that the eigenvalues of a random matrix (or the zeros of the zeta function) tend to repel. In the case of random matrices we can understand this phenomenon qualitatively. For a generic matrix, all the eigenvalues are distinct, and the size of its orbit, i.e., matrices of the form UMU^*, is equal to $N^2 - N$. By contrast, if two eigenvalues are equal then the dimension is only $N^2 - N - 2$. It is this leap in two dimensions which explains the fact that the "correlation function" vanishes at zero.

2.3 Experimental verifications and new conjectures

Montgomery's conjecture motivated A. Odlyzko [10] to numerically compute many zeros of the zeta function with positive imaginary part on the critical line. In 1987, he obtained the 10^5 zeros between the 10^{12}+1-st and $10^{12}+10^5$-th where the zeros are ordered according to their (positive) imaginary part. He found good agrement between numerical data and conjecture. Since 1987 computational power has greatly increased, and the most recent data concerns a million zeros around 10^{22}-th zero. For example, the $10^{22}+1$-st zero is

$$1/2 + i1370919909931995308226.68016095\ldots$$

and the lowest significant digits of the following three zeros are

$$\begin{array}{l} 8226.77659152 \\ 8226.94593324 \\ 8227.16707942. \end{array}$$

Thus, Montgomery's guess and Odlyzko's numerical verification add weight to Hilbert and Polya's conjecture that the γ are the eigenvalues of a Hermitian operator. While not providing a clear notion of the origin of this operator, they do help to identify its form a little better. In fact we can consider Gaussian random matrices presenting the different symmetries of Hermitian matrices, such as real symmetric matrices (GOE for Gaussian Orthogonal Ensemble), or symplectic matrices (GSE = Gaussian Symplectic Ensemble). Calculations similar to those of the preceding paragraph allow us to determine the pair correlation function (the function $1-\left(\frac{\sin \pi x}{\pi x}\right)^2$ for GUE), and other eigenvalue statistics. These statistics differ from that of GUE. Thus, the numerics with GUE suggest that the operator of Polya and Hilbert, if it exists, should be hermitian, not orthogonal or symplectic.

Montgomery's results have been extended to other L-functions and research in this area is currently very active. However, time and skill do not allow me address this subject, for which I refer to a recent Bourbaki seminar by P. Michel [8].

To conclude this first part we'll look at another problem concerning the zeta function, that of the asymptotic behavior of its moments on the critical line. The problem is to estimate

$$\int_0^T \left|\zeta(\frac{1}{2}+it\right|^{2k} dt,$$

as $T \to \infty$. This question is motivated by applications to number theory, in particular to the distribution of primes. The first results in this direction date back to Hardy and Littlewood. They showed that

$$\int_0^T \left|\zeta(\frac{1}{2}+it\right|^{2} dt \sim T\log T, \quad T \to \infty,$$

and Ingham showed that

$$\int_0^T \left|\zeta(\frac{1}{2}+it\right|^{4} dt \sim \frac{1}{2\pi^2} T\left(\log T\right)^4, \quad T \to \infty.$$

Nothing else has been shown for higher moments. There are conjectures for the asymptotics of the moments $\int_0^T \left|\zeta(\frac{1}{2}+it\right|^6 \, dt$ (due to Conrey and Ghosh) and for $\int_0^T \left|\zeta(\frac{1}{2}+it\right|^6 \, dt$ (by Conrey and Gonek) of the form

$$\int_0^T \left|\zeta(\frac{1}{2}+it\right|^{2k} dt, \sim a_k b_k T \left(\log T\right)^{k^2}, \tag{15}$$

with

$$a_k = \prod_{p\in\mathcal{P}} \left(1-\frac{1}{p}\right)^{k^2} \left(\sum_{m=0}^{\infty} \left(\frac{\Gamma(m+k)}{m!\,\Gamma(k)}\right)^2 p^{-m}\right).$$

The product is taken over the set of prime numbers $\mathcal{P}$. Also,

$$b_3 = \frac{42}{9!}, \quad b_4 = \frac{24024}{16!}.$$

The hardest part in the statement of the conjecture is to obtain the coefficients b_3 and b_4 through lengthy calculations. Using the GUE model, Keating and Snaith[6] have formulated a conjecture valid for all k, including the nature of the terms a_k and b_k. Their idea is that in the product $a_k b_k$ the first term, a_k is a specific contribution to the zeta function that involves the primes explicitly. But the second, b_k, is a universal term determined by the fluctuations of the zeta zeros, and it should be the same for any function whose zeros have the same fluctuations. In order to calculate this term, they replace $\zeta(s)$ by the characteristic polynomial of a random GUE matrix, and calculate an appropriate function.

I will not dwell on the origin of the idea of universality, which comes from the study of critical phenomena in statistical mechanics, since this far exceeds the scope of this presentation. But it is remarkable that this idea leads quickly to the same factors as those calculated by the methods of number theory. In the calculation of Keating and Snaith, we consider a random unitary matrix, chosen with the Haar measure on the group $U(N)$. (This model is called CUE (circular unitary ensemble), and it is very similar to GUE. Results similar to Keating and Snaith were also provided a little later by Brezin and Hikami [2]). We then calculate the average $\mathbb{E}|\det(1-e^{i\theta}U)|^{2k}$ where U is a CUE matrix and $e^{i\theta}$ a complex number of modulus 1. In fact we see immediately that this quantity does not depend on $e^{i\theta}$, and an explicit calculation gives

$$\mathbb{E}\left|\det(1-e^{i\theta}U)\right|^{2k} \overset{N\to\infty}{\longrightarrow} N^{k^2} \prod_{j=0}^{k-1} \frac{j!}{(j+k)!},$$

which allows us to conjecture that

$$b_k = \prod_{j=0}^{k-1} \frac{j!}{(j+k)!}.$$

It is easy to check that these values agree with those above for b_3 and b_4. Keating and Snaith conjecture that these are the general forms for all k.

3 The Riemann zeta function and Brownian motion

3.1 Some formulas from analysis

Recall that the Riemann zeta function is defined by the formula

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \Re s > 1.$$

It admits a meromorphic extenstion to the complex plane with a simple pole at zero. The functional equation for the ζ function can be written in a symmetric form by introducing the entire function

$$\xi(s) = \frac{1}{2} s(s-1)\pi^{-s/2}\Gamma(\frac{s}{2})\zeta(s), \quad \Re s > 1,$$

which satisfies the functional equation

$$\xi(s) = \xi(1-s), \quad s \in \mathbb{C}. \tag{16}$$

One way to establish this result is to note that

$$\xi(s) = \frac{1}{2}\int_0^{\infty} t^s \Psi(t)\, dt, \quad s \in \mathbb{C}, \tag{17}$$

where

$$\Psi(y) = 4y \sum_{n=1}^{\infty} \left(2\pi^2 n^4 y^2 - 3\pi n^2\right) e^{-\pi n^2 y^2}. \tag{18}$$

The function Ψ may be expressed with the help of Jacobi's θ function

$$\theta(t) = \sum_{n=-\infty}^{\infty} e^{-\pi n^2 t},$$

using the formula

$$\Psi(y) = 6y\theta'(y^2) + 4y^3\theta''(y^2).$$

The functional equation for the Jacobi theta function

$$\theta(t) = \frac{1}{\sqrt{t}}\theta(\frac{1}{t}), \quad t > 0$$

implies that Ψ satisfies

$$\Psi(y) = \frac{1}{y^3}\Psi(\frac{1}{y}), \quad y > 0. \tag{19}$$

This allows us to analytically continue the zeta function and deduce equation (16).

The starting point of the developments that follow is that Ψ is positive on the half-line $\mathbb{R}_+$ and has integral 1, thus it is the density of a probability measure on the half-line. Indeed, the formula (18) shows that $\Psi(y) > 0$ for $y > 1$, because it is a sum of positive terms, and the functional equation (19) implies positivity for $y < 1$. The graph of this density is indicated in Figure 1, and the distribution function

$$F_\Psi(y) = \int_0^x \Psi(y)\,dy = 1 + 2\sum_{n=1}^{\infty}\left(1 - 2\pi n^2 y^2\right) e^{-\pi n^2 y^2} \tag{20}$$

$$= \frac{4\pi}{y^3}\sum_{n=1}^{\infty} n^2 e^{-\pi n^2/y^2}. \tag{21}$$

is obtained by integrating equation (18) or (19) term-by-term.

It is a remarkable fact that in spite of the complex appearance of formula (18), this probability density arises in a natural way from the theory of random walks and Brownian motion. A complete review of the probabilistic interpretations of (18) in terms of Brownian motion may be found in [1, 14]. It is not possible in the context of this presentation to discuss all these results, but I will mention the interpretation which I think is the most accessible. This involves the excursions of Brownian motion away from zero. We give an elementary presentation of these ideas using the coin toss in Section 3.2. In Section 3.4 this will allow us to show that the functional equation for the zeta function is equivalent to the equality in law of two random variables defined using Brownian motion and related processes. Finally in Section 3.5 we will see that the random variable whose law has density Ψ admits a decomposition that can renormalize the sum $\sum_n n^{-s}$.

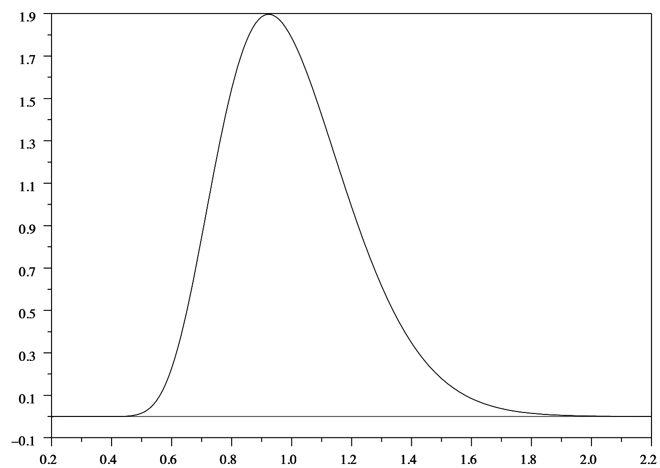

Figure 1: The distribution function of equation (20)

The origin of the relationship between zeta function and Brownian motion may be found in the fact that the theta functions of Jacobi, which are closely related to the Riemann zeta function, arise in the solutions of the heat equation. On the other hand, we know that heat flow and Brownian motion are two physical phenomena whose underlying mathematical structure is the same. Thus, we can imagine in this way that the Riemann zeta function must appear in the theory of Brownian motion. These general considerations, however, tell us nothing precise about the exact nature of these relationships. In particular the fact that the zeta function (or more accurately the probability density Ψ) appear in natural problems is remarkable.

3.2 The game of heads and tails

Two players compete in a game of heads and tails. It is assumed that the payoff for each win is a unit, and we are interested in the winnings of one of the players. We can represent this gain after n steps by a sum $S_n = X_1 + \ldots + X_n$, where X_i represents the result of the ith game. The X_i are independent random variables that satisfy $P(X_i = \pm 1) = 1/2$. We assume that the fortunes of the two players are endless, and that the game never stops. A classical theorem of Polya asserts that with probability 1, the gain of the players will be 0 for infinitely many values of n, i.e. both players will return to their initial (equal) fortunes infinitely often. We will establish this result below by elementary considerations.

Let $T_1, T_2, \ldots T_n$, denote the times of successive returns to zero, i.e., $T_0 = 0$ and $T_j = \inf\{n > T_{j-1},\, S_n = 0\}$ for $j > 0$. After each return to zero, the gain behaves like a simple random walk, that is to say the family of random variables $(S_{T_j+n}, n \geq 0)$ has the same law as the family $(S_n, n \geq 0)$. Moreover it is independent of the random variables $(S_n \mathbf{1}_{n \leq T_j}, n \geq 0)$. This is a consequence of the strong Markov property of the random walk, which can be verified by conditioning with respect to the value of T_j. The times $T_0, T_1, T_2, \ldots, T_n$, therefore form an increasing sequence, whose increments $(T_i - T_{i-1}; i \geq 1)$ are independent and have the same law as T_1. We can calculate the probability that the first return to 0 occurs at time $2n$ (it is clear that the return time cannot be an odd number). As we shall see below,

$$P(T_1 = 2n) = \frac{(2n-2)!}{2^{2n-1} n!(n-1)!}. \tag{22}$$

Similarly, the maximum difference in fortunes has a simple law. If we put $M_j = \max\{|S_n|, T_{j-1} \leq n \leq T_j\}$ then M_j are iid random variables with the law

$$P(M_j = r) = \frac{1}{r} - \frac{1}{r+1}, \quad r = 1, 2, \ldots$$

The length of most of the time intervals between any two successive returns to 0 is small. Thus the maximum difference in gain between the two players is also small. Nevertheless, sometimes this interval is very long, and the maximum difference in gain is important. To quantify this, we will calculate the probability that the return time is equal to $2n$ and the maximum difference in earnings is equal to m for given n and m (see Diaconis and Smith [12], and also [3, 11] for similar calculations). It is convenient to represent the sequence $(S_k, k \geq 0)$ by the graph obtained by linearly interpolating between integer times as shown in Figure 2. We consider such a graph restricted to the time interval $k \in [0, n]$. Each graph corresponds to the realization of a unique sequence $(X_1, \ldots, X_n) \in \{+1, -1\}^n$, therefore the probability of the event it represents is 2^{-n}.

Consider now the event that the first return time equals $2n$, denoted $\{T_1 = 2n\}$. A sequence $(X_1, \ldots, X_{2n})$ realizes this event if and only if the sequence of partial sums $(S_k; 1 \leq k \leq 2n)$ satisfies $S_{2n} = 0$ and $S_k \neq 0$ for $1 \leq k \leq 2n-1$. The calculation of the number of such sequences is a classic exercise in the use of the reflection principle of Désiré André. It suffices to count the number of sequences that are strictly positive for $k \in [1, 2n-1]$ and to multiply this number by 2. Each such positive sequence has $S_1 = 1$ and $S_{2n-1} = 1$. We first count the sequences with $S_0 = S_{2n} = 0$ and

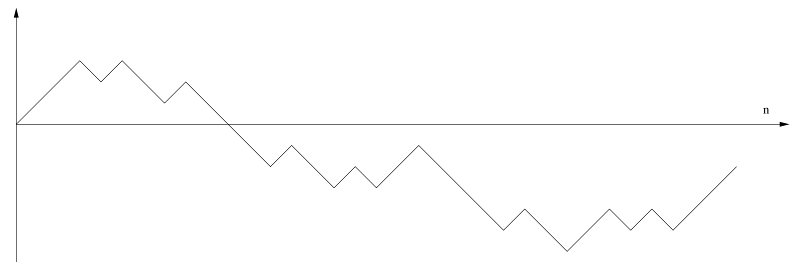

Figure 2: The random walk

$S_1 = S_{2n-1} = 1$. These sequences correspond to variables $X_1, \ldots, X_{2n}$ with $X_1 = 1$ and $X_{2n} = -1$ such that there are an equal number of $+1$'s and -1's amongst the $X_2, \ldots, X_{2n-1}$. The number of such sequences is given by the binomial coefficient

$$\binom{2n-2}{n-1} = \frac{(2n-2)!}{(n-1)!(n-1)!}.$$

From this set, we must remove the sequences that vanish for at least one $k \in [2, 2n-2]$. Let $(S_1, \ldots, S_{2n})$ be such a sequence. Then there exists a smallest integer $k_0 \in [2, 2n-1]$ such that $S_{k_0} = 0$. We define a new sequence S'_k with $S'_k = S_k$ for $k \leq k_0$, and $S_k = -S_k$ for $k_0 < k \leq 2n$. The graph of the sequence S' is obtained by reflecting the graph of the sequence S in the axis $y = 0$, after the first passage time at 0, see Figure 3. Conversely if a sequence S' satifies $S'_1 = 1$, $S'_{2n-1} = -1$ and $S'_{2n} = 0$, then it necessarily vanishes for some $k \in [2, 2n-2]$, and it can be reflected in the first moment after it enters zero to obtain a sequence S such that $S_{2n-1} = 1$, $S_{2n} = 0$ and S_k vanishes at some k between 2 and $2n-2$. The sequence S' corresponds to a sequence $(X'_i; 2 \leq i \leq 2n-1)$ for which the number of $+1$'s is $n-2$ and the number of -1's is n. Thus, the total number of such sequences is $\binom{2n-2}{n}$. As a consequence, the number of sequences S that do not vanish at any point between 1 and $2n-1$ is

$$2 \times \left(\frac{(2n-2)!}{(n-1)!(n-1)!} - \frac{(2n-2)!}{n!(n-2)!} \right) = 2\frac{(2n-2)!}{n!(n-1)!}$$

and we recover formula (22) upto the probability of each path. In particular, we see that $\sum_{n=1}^{\infty} P(T_1 = 2n) = 1$, thus $T_1 < \infty$ almost surely. Applying the Markov property, we see that for all j we have $T_j < \infty$ almost surely, and therefore S_n returns to zero infinitely often with probability 1.

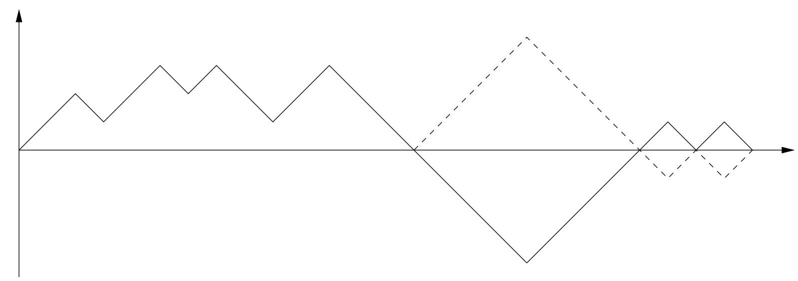

Figure 3: The random walk reflected at its first passage to zero

We now calculate the probability $P(T_1 = 2n; M_1 \leq m)$ that the return time is equal to $2n$ and that the maximum value $M_1 = \max_{1\leq k\leq 2n} |S_k| \leq m$. Upto a factor of 2, this reduces to a calculation of the number of sequences $(S_k; 0 \leq k \leq 2n)$ such that

$$S_0 = 0, \quad S_{2n} = 0, \quad \text{and} \quad 0 < S_k \leq m, \quad \text{for} \quad 1 \leq k \leq 2n-1. \tag{23}$$

We will perform this calculation by two different methods. The first is a generalization of the preceding calculation. Let $s_{2n}(k)$ denote the number of sequences S_k such that $S_1 = 1$ and $S_{2n-1} = k$. We have

$$s_{2n}(k) = \binom{2n-2}{n-l}, \quad \text{if} \quad k = 2l-1 \in [-2n+3, 2n-1],$$

and $s_{2n}(k) = 0$ otherwise. We start by counting the sequences that return to 1 at time $2n-1$[1] without other conditions. We then subtract the number of those that pass through 0, and those that pass through $m+1$. To calculate the number of the latter, we perform a reflection about the line $y = m+1$, and we find with reasoning as above, that we obtain all sequences satisfying $S_{2n-1} = 2m+1$, the number of which is $s_{2n}(2m+1)$. In doing so, we double counted sequences whose minimum is ≤ 0 and whose maximum is $\geq m+1$. It is therefore necessary to add these sequences. We count them by subjecting the path to two reflections at the times when it reaches 0 and $m+1$ respectively. We obtain a sequence such that $S_{2n-1} = 2m+3$ or $S_{2n-1} = -2$, depending on whether 0 or $m+1$ is reached first. Their number is $s_{2n}(2m+3) + s_{2m}(-2m-1)$. We see that this time we have removed too many sequences and we must correct this calculation again. Finally, by an

[1]check math and translation here

application of the inclusion-exclusion principle, we see that the number of sequences we seek is given by

$$\sum_{g\in G} \det(g) s_{2n}(g(1)),$$

where G is the group of isometries of $\mathbb{R}$ spanned by the reflections $x \mapsto -x$ and $x \mapsto 2m+2-x$, that is

$$\sum_{k\in Z} s_{2n}\left(1+(2m+2)k\right) - s_{2n}\left(-1+(2m+2)k\right).$$

Thus the desired probability is

$$\begin{aligned} &P(T_1 = 2n; M_1 \leq m) \\ &\quad = 2^{-(2n-1)} \sum_{k\in\mathbb{Z}} \binom{2n-2}{n-1-k(m+1)} - \binom{2n-2}{n-2-(k+1)(m+1)} \end{aligned} \tag{24}$$

with the convention $\binom{a}{b} = 0$ if $b < 0$ or if $a < b$. In particular, this sum contains only a finite number of non-zero terms.

We now indicate another method to calculate the same number. Consider the $m \times m$ matrix

$$\Gamma = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

We split $\Gamma = \Gamma_+ + \Gamma_-$ where Γ_+ is lower triangular and Γ_- is upper triangular. In the canonical basis of $\mathbb{R}^m$ we then find $\Gamma_+(e_i) = e_{i+1}$ for $i \leq m-1$ and $\Gamma_+(e_m) = 0$. Similarly, $\Gamma_-(e_1) = 0$ and $\Gamma_-(e_i) = e_{i-1}$ for $2 \leq i \leq m$. Consider the inner product $\langle \Gamma^{2n-2}(e_1), e_1 \rangle$. We develop the product to find

$$\begin{aligned} \langle \Gamma^{2n-2}(e_1), e_1 \rangle &= \langle \left(\Gamma_+ + \Gamma_-\right)^{2n-2}(e_1), e_1 \rangle \\ &= \left\langle \sum_{(\varepsilon_1,\dots,\varepsilon_{2n-2})\in\{\pm\}^{2n-2}} \Gamma_{\varepsilon_{2n-2}} \Gamma_{\varepsilon_{2n-3}} \dots \Gamma_{\varepsilon_1}(e_1), e_1 \right\rangle, \end{aligned}$$

where the sum runs over all 2^{2n-2} possible combinations of the symbols. Since the operators $\Gamma_\pm$ transform the elements of the canonical basis into other elements of this basis, in order that a term in the sum be nonzero, it is necessary and sufficient that $\Gamma_{\varepsilon_{2n-2}}\Gamma_{\varepsilon_{2n-3}}\ldots\Gamma_{\varepsilon_1}(e_1)=e_1$. In this case, we have $\Gamma_{\varepsilon_k}\Gamma_{\varepsilon_{k-1}}\ldots\Gamma_{\varepsilon_1}(e_1)=e_{S_{k+1}}$ where $(S_k:1\leq k\leq 2n-1)$ is a sequence that satisfies the conditions (23). Thus, $\langle\Gamma^{2n-2}(e_1),e_1\rangle$ is equal to the number of such sequences. This term may be calculated by diagonalizing the matrix Γ. The characteristic polynomial may be computed by a recurrence in m. We expand $\det(\lambda I_m-\Gamma)=P_m(\lambda)$ with respect to the last column, to obtain the relation $P_m(\lambda)=\lambda P_{m-1}(\lambda)+P_{m-2}(\lambda)$. This recurrence relation with initial conditions $P_1(\lambda)=\lambda$ and $P_2(\lambda)=\lambda^2-1$ yields P_m in terms of Chebyshev polynomials of the second kind, and we find

$$P_m\left(2\cos\theta\right)=\frac{\sin\left((m+1)\theta\right)}{\sin\theta},\quad 0<\theta<\pi.$$

In particular, the roots of P_m are the numbers $(2\cos\left(\frac{k\pi}{m+1}\right),1\leq k\leq m)$. The eigenvectors are computed as follows. The eigenvector $(x_1,\ldots,x_m)$ corresponding to the eigenvalue λ satisfies $x_{l-1}+x_{l+1}=\lambda x_l$. We set $\lambda=2\cos\left(k\pi/(m+1)\right)$ to find

$$x_l=P_{l-1}(\lambda)x_1=\frac{\sin\left(\frac{kl\pi}{m+1}\right)}{\sin\left(\frac{k\pi}{m+1}\right)}x_1.$$

Finally, we denote the orthonormal eigenvectors by $(w_1,\ldots,w_m)$, and we have

$$w_k=\frac{2}{\sqrt{m}}\left(\sin\left(\frac{k\pi}{m+1}\right),\sin\left(\frac{2k\pi}{m+1}\right),\ldots\sin\left(\frac{mk\pi}{m+1}\right)\right).$$

We have thus found

$$\begin{aligned}\langle\Gamma^{2n-2}(e_1),e_1\rangle&=\langle\Gamma^{2n-2}\left(\sum_{k=1}^{m}\langle e_1,w_k\rangle w_k\right),\sum_{l=1}^{m}\langle e_1,w_l\rangle w_l\rangle\\&=\sum_{k=1}^{m}\langle e_1,w_k\rangle^2\lambda_k^{2n-2}\\&=\sum_{k=1}^{m}\frac{4}{m}\sin^2\left(\frac{k\pi}{m+1}\right)\left(2\cos\frac{k\pi}{m+1}\right)^{2n-2}.\end{aligned}$$

The computed probability is therefore

$$P(T_1=2n;M_1\leq m)=\frac{2}{m}\sum_{k=1}^{m}\sin^2\left(\frac{k\pi}{m+1}\right)\left(\cos\frac{k\pi}{m+1}\right)^{2n-2}.\qquad(25)$$

The conditional distribution of the maximum deviation, knowing that the return time is equal to $2n$, is given by the distribution function

$$F_n(x) = P\left(M_1 \le m \,|\, T_1 = 2n\right) = \frac{P\left(M_1 \le m; T_1 = 2n\right)}{P(T_1 = 2n)}.$$

The zeta function, or more precisely the density Ψ, appears when we take the limit $n \to \infty$. Specifically, we will calculate the conditional distribution of $M_1/\sqrt{\pi n}$ given that $T_1 = 2n$. We use the first expression in (24) and Stirling's formula to obtain

$$\lim_{n\to\infty} P\left(M_1 \le y\sqrt{\pi n} \,|\, T_1 = 2n\right) = 1 + 2\sum_{n=1}^{\infty}\left(1 - 2\pi n^2 y^2\right) e^{-\pi n^2 y^2},$$

while the second expression in (25) yields

$$\lim_{n\to\infty} P\left(M_1 \le y\sqrt{\pi n} \,|\, T_1 = 2n\right) = \frac{4\pi}{y^3}\sum_{n=1}^{\infty} n^2 e^{-\pi n^2/y^2}.$$

We have thus established by this elementary method, the equalities (20) and (21). We will now interpret the passage to the limit with the help of Brownian motion.

3.3 Brownian motion

We first recall a classical result of probability theory. The random variable S_n has a binomial distribution, and the de Moivre-Laplace theorem asserts that the law of the random amount variable $S_n/\sqrt{n}$ converges to the centered normal distribution, with density of $e^{-x^2/2}/\sqrt{2\pi}$ with respect to the Lebesgue measure on $\mathbb{R}$. The convergence is in the sense of weak convergence of probability measures on $\mathbb{R}$ and for any interval $[a, b] \subset \mathbb{R}$ we have

$$\lim_{n\to\infty} P(S_n/\sqrt{n} \in [a, b]) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}\,dx.$$

Similarly it is easily seen that for any real number $t > 0$, the law of $S_{[nt]}/n$, where $[\cdot]$ denotes the integer part, converges to the normal distribution with variance t, whose density is $e^{-x^2/2t}/\sqrt{2\pi t}$. Finally, thanks to independence of increments $(Sn; n \ge 0)$ we see that for any sequence of times $t_1 < t_2 < \ldots < t_k$, the family of random variables $(\frac{S_{[nt_1]}}{\sqrt{n}}, \frac{S_{[nt_2]}-S_{[nt_1]}}{\sqrt{n}}, \ldots, \frac{S_{[nt_k]}-S_{[nt_{k-1}]}}{\sqrt{n}})$ converges in distribution to a family of independent normal variables with variances t_1, $t_2 - t_1$, …, $t_k - t_{k-1}$. Brownian motion is a stochastic process i.e. a family of random variables $(X_t, t \in \mathbb{R}_+)$ indexed by time t such that for all k and any k-tuple of times $t_1 < t_2 < \ldots < t_k$, the random variables $X_{t_1} - X_0$, $X_{t_2} - X_{t_1}$, …, $X_{t_k} - X_{t_{k-1}}$ are independent normal random variables with mean zero and variance t_1, $t_2 - t_1$, …, $t_k - t_{k-1}$. In other words, the finite-dimensional marginal distributions of the family $(X_t, t \in \mathbb{R}_+)$ are determined by the formula

$$\mathbb{E}\left(f(X_{t_1}, X_{t_2}, \ldots, X_{t_k})\right) = \int_{\mathbb{R}^n} f(x_1, x_2, \ldots, x_n) p_{t_1}(0, x_1) p_{t_2 - t_1}(x_1, x_2) \ldots p_{t_n - t_{n-1}}(x_{n-1}, x_n)\, dx \tag{26}$$

for each $t_1 < t_2 < \ldots < t_n$ and each Borel function f on $\mathbb{R}^n$, where the transition density p_t is given by

$$p_t(x, y) = \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2t}. \tag{27}$$

A fundamental property of the Brownian motion, obtained by first Wiener, is that its paths are almost surely continuous. If one considers the space of continuous functions from $[0, \infty)$ to $\mathbb{R}$ with the topology of uniform convergence and the associated Borel structure, then the Wiener measure on this space is a probability measure such that under this measure, the coordinate maps $X_t : C([0, \infty), \mathbb{R}) \to \mathbb{R}$, $\omega \mapsto \omega(t)$ satisfy the above conditions. It can be shown that the continuous stochastic process $S_t^{(n)}; t \geq 0)$ obtained by linearly interpolating the graph of the random walk $(S_n; n \geq 0)$ and renormalizing $S_t^{(n)} = S_{[nt]}/\sqrt{n}$ converges in the space $C([0, \infty), \mathbb{R})$ to the Wiener measure. This means that for any continuous function Φ on $C([0, \infty), \mathbb{R})$, we have $\mathbb{E}\left[\Phi\left(S^{(n)}\right)\right] \to \mathbb{E}_{\mathcal{W}}\left[\Phi(\omega)\right]$, where $\mathbb{E}_{\mathcal{W}}$ denotes expectation with respect to Wiener measure. It is important to have this information to calculate the laws of certain functionals of Brownian motion by means of the approximation by random walks. Note that the fact that the variables X_i are Bernoulli variables (ie, take only two values) is not of great importance, the expected result of approximation is still true under the assumption that $X)i$ are identically distributed, have mean zero and variance equal to 1. This result is known as Donsker's invariance principle. A good introduction to Brownian motion may be found in the book of Karatzas and Shreve [5]. One can visualize the continuity of the paths, and the approximation by the random walks through a computer simulation. Here is a simple program in Scilab which traces the trajectory of $S^{(n)}$ for t [0,1], which was used to plot the graph in 4.

```
xbasc();
plotframe([0 -2 1 2],[1,4,1,0]);
A=[0,1];B=[0,0]; plot2d(A,B,1,000); // Definition of the axes;
N=10000; // This is the number of steps.
rand(normal); X=0;Y=0;SX=0;SY=0;
for i=2 :N+1
U=X;X=X+1/N; V=Y;
if rand(1)>0 then Y=Y+1/(sqrt(N)) ;
else
Y=Y-1/(sqrt(N)) ;// Calculate increments.
end
SX=[U,X]; SY=[V,Y]; plot2d(SX,SY,1,000);
end
```

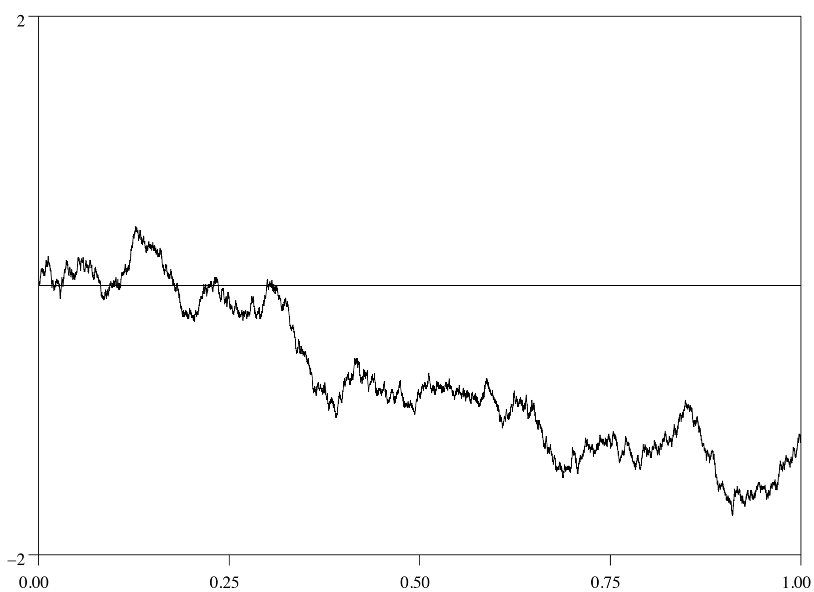

Figure 4: Numerical simulation of Brownian motion

In addition to Brownian motion there are several closely related stochastic processes. First, the Bessel process of dimension 3, which is the Euclidean norm of a three-dimensional Brownian motion, that is

$$R_t = \sqrt{(B_t^{(1)})^2 + (B_t^{(2)})^2 + (B_t^{(3)})^2}$$

where $B^{(1)}$, $B^{(2)}$, and $B^{(3)}$ are independent Brownian motions. The law of its finite-dimensional marginals is given as in (26) by

$$\mathbb{E}\left(f(X_{t_1}, X_{t_2}, \dots, X_{t_k})\right) = \tag{28}$$
$$\int_{\mathbb{R}^n} f(x_1, x_2, \dots, x_n) q_{t_1}(0, x_1) q_{t_2-t_1}(x_1, x_2) \dots q_{t_n-t_{n-1}}(x_{n-1}, x_n)\, dx$$

where the transition density q_t is given by

$$q_t(x,y) = \frac{y}{x}\left(p_t(x,y) - p_t(x,-y)\right) \equiv \frac{y}{x} p_t^0(x,y). \tag{29}$$

where p_t is defined in (27). We can show that with probability 1, $R_t \to \infty$ as $t \to \infty$ and $R_t > 0$ for all $t > 0$. Moreover, using the Feynman-Kac formula we can find the law of the first instant where $R_t = x$. If we define $T_x = \inf\{t \,|R_t \geq x\}$ then we have

$$\mathbb{E}\left[e^{-\lambda T_x}\right] = \frac{x\sqrt{\lambda}}{\sinh x\sqrt{\lambda}}. \tag{30}$$

This formula will play an important role in the sequel.

We will also need the Bessel bridge of dimension three, also called Brownian excursion, which is obtained by conditioning the Bessel process to reach 0 at time T. As Bessel process of dimension 3 never returns to 0, we must define this conditional event of probability 0 carefully. But it is possible and the Bessel bridge of dimension 3 is a stochastic process $(e(t); t[0,T])$ indexed by time $t \in [0,T]$, with law

$$n_T\left[f(e_{t_1}, e_{t_2}, \dots, e_{t_n})\right] = \int_{\mathbb{R}^n} f(x_1, x_2, \dots, x_n) q_{t_1}(0, x_1) q_{t_2-t_1}(x_1, x_2)$$
$$\dots q_{t_n-t_{n-1}}(x_{n-1}, x_n) 2(2\pi t^3)^{1/2} x_n^{-2} q_{T-t_n}(x_n, 0)\, dx, \tag{31}$$

for $t_1 < t_2 < \dots < t_n < T$. This process can be obtained as the limit of the random walk $(S^{(n)}, n \geq 0)$ conditioned to return to 0 at time N. More precisely, we have convergence of finite-dimensional marginals

$$\lim_{n\to\infty} \mathbb{E}\left[f\left(\frac{|S_{nt_1}|}{\sqrt{n}}, \dots, \frac{|S_{nt_n}|}{\sqrt{n}}\right) |\, T_1 = [Tn]\right] = n_T\left[f(e_{t_1}, e_{t_2}, \dots, e_{t_n})\right]. \tag{32}$$

and we can demonstrate the convergence in distribution of the process obtained by linear interpolation on the space of continuous functions. Moreover, with probability 1, we have $e(0) = e(T) = 0$ and $e(t) > 0$ for $0 < t < T$. So we see that the law with probability density Ψ is the law of the maximum of the Bessel bridge in dimension three multiplied by $\sqrt{2/\pi}$. That is, we have

$$n_1\left(\sqrt{2/\pi}\max_{t\in[0,1]} e(t)\in dx\right) = \Psi(x)\,dx, \quad x>0. \tag{33}$$

The Brownian motion, Bessel processes and bridge possess a critical property of scale invariance. For $\lambda > 0$, the transformation $\omega \mapsto \omega_\lambda$ where

$$\omega_\lambda(t) = \lambda^{-1/2}\omega(\lambda t), \tag{34}$$

leaves the laws of Brownian motion and the Bessel process invariant, and transforms the law n_T into the law $n_{T/\lambda}$.

3.4 The functional equation for the zeta function and Ito's measure

In this section we interpret the functional equation of the Riemann zeta function as the equality in distribution of two random variables. In order to see this, we introduce the Ito measure of Brownian excursions. We consider the space of continuous excursions $\omega : \mathbb{R} \to \mathbb{R}_+$, such that $\omega(0) = 0$, and there exists $T(\omega) > 0$ such that $\omega(t) > 0$ for $0 < t < T(\omega)$ and $\omega(t) = 0$ for $t \geq T(\omega)$. The law of the process obtained by linear interpolation from $|S_n|$, $0 \leq n \leq T_1$ for time upto T_1, and extended by zero after the time T_1, is a probability measure on this space. The Ito measure is the scaling limit, as $\lambda \to \infty$ of measures λP^λ where P^λ denotes the law of rescaled processes $\lambda{-}1/2|S_{\lambda t}|\mathbf{1}_{t\leq T_1/\lambda}$. This measure, denoted n_+ has infinite total mass. However, it can be expressed in terms of its finite-dimensional marginals by

$$\begin{aligned} n_+\left[f(\omega_{t_1},\omega_{t_2},\ldots,\omega_{t_n})\right] = \\ = \int_{\mathbb{R}_+^n} 2x_1\frac{x_1 e^{-x_1^2/2t}}{\sqrt{2\pi t^3}} p^0_{t_2-t_1}(x_1,x_2)\ldots p^0_{t_n-t_{n-1}}(x_{n-1},x_n)\,dx, \end{aligned} \tag{35}$$

where p^0_t is defined in (29). We can also describe this measure with the help of laws n_T defined in (31) by the formula

$$n_+ = \int_0^\infty n_T\frac{2dT}{\sqrt{2\pi T^3}} \tag{36}$$

following the limit theorem (3.17) and (3.7), by applying the Stirling formula

$$P(T_1 = 2n) = \frac{(2n-2)!}{2^{2n-1}n!(n-1)!} \sim \frac{1}{2\sqrt{\pi n^3}}.$$

Formula (36) means that the law of the time to return to 0 under n_+ is $2dT/\sqrt{2T^3}$, and conditionally on the return time, the excursion process is a Bessel bridge of dimension 3.

The Ito measure is fundamental to an understanding of the behavior of Brownian motion outside the times it vanishes. The set $Z_\omega = \{t \in [0,\infty)\,|\omega(t) = 0\}$ of zeros of Brownian motion is closed by the continuity of trajectories. With probability one, it is a perfect set (with no isolated points) of zero Lebesgue measure. In particular, it is uncountable, and we cannot define an increasing sequence of times that denumerate the returns to zero, unlike the discrete set $T_1, \ldots, T_n, \ldots$ for the random walk. Nevertheless, since the complement of Z_ω is open, it has a countable infinity of connected components, and excursions of Brownian motion are by definition the pieces of the path corresponding to these connected components. Next, following P. Lévy we can introduce the local time of the Brownian motion B by

$$L_t = \lim_{\varepsilon \to 0^+} \frac{1}{2\varepsilon} \int_0^t \mathbf{1}_{|B_s| \leq \varepsilon}\, ds,$$

and the inverse function $T_s = \inf\{t | L_t \geq s\}$ is the continuous analogue of the discrete sequence T_1,T_2, $\ldots$ for the random walk. For any time s we have $B_{T_s} = 0$ almost surely, and the excursions of Brownian motion in the interval $[0, T_s]$ form a Poisson point process with intensity n_+. This means the following. For each family A_1, $\ldots$, A_k of disjoint Borel subsets of the space of continuous functions, let $N(A_j)$ denote the number of excursions of Brownian motion in the interval $[0, T_s]$ which lie in A_j. Then the $N(A_j)$ are independent, Poisson random variables with parameter $n_+(A_j)$.

Now let us return to our initial goal of a probabilistic explanation for the functional equation for the zeta function. The following explanation of the Ito measure is due to David Williams. We consider two independent Bessel(3) processes (cf. (28)), $(R^1_t : t > 0)$ and $(R^2_t : t > 0)$. Given $x > 0$, let T^1_x and T^2_x denote the first passage times of R^1_t and R^2_t respectively, and consider the process defined by

$$e(t) = \begin{cases} R^1_t, & 0 \leq t \leq T^1_x, \\ R^2_{T^1_x + T^2_x - t}, & T^1_x \leq t \leq T^1_x + T^2_x \\ 0, & t > T^1_x + T^2_x. \end{cases}$$

We denote by n^x the law of this process, obtained by gluing together “back to back” two copies of the Bessel process, stopped at its first passage time at x. The Williams decomposition of Ito's measure asserts that

$$n_+ = \int_0^\infty n^x \frac{2\, dx}{x^2}. \tag{37}$$

In other words, under the Ito measure, the law of the maximum of the excusrion is $2dx/x^2$ and conditionally on the maximum $M = x$, the law of the excursion e is n^x. In particular, conditionally on the value of the maximum, the law of the return time to 0 is that of the sum of two hitting time of x by two independent Bessel(3) processes. Consider now the law of the pair (M, V) under the measure n_+, where M is the maximum excursion ω and V is the return time to 0, i.e. $V = \inf\{t > 0 | \omega(t) = 0\}$. In accordance with Ito's description (36) and using the scale invariance (34), we can write

$$(M, V) \stackrel{\text{law}}{=} (\sqrt{V} m, V)$$

where m denotes the maximum of a Bessel bridge under the law n_1, independent of V. Similarly, Williams decomposition gives

$$(M, V) \stackrel{\text{law}}{=} \big(M, M^2(T^1 + T^2)\big)$$

where T^1 and T^2 are the hitting times at 1 by two Bessel(3) processes, independent of one another and M. The relation

$$(\sqrt{V} m, V) \stackrel{\text{law}}{=} \big(M, M^2(T^1 + T^2)\big)$$

leads to a relationship between the laws of random variables m and T^1+T^2. But we must be careful, because the hasty conclusion that m has the same law as $(T_1 + T_2)^{-1/2}$ (obtained by considering the law of $M/\sqrt{V}$) is false! Indeed, the distribution of (M, V) is a measure of infinite total mass and its image under the map $(M, V) \mapsto M/\sqrt{V}$ is a degenerate measure, that is $+\infty$ on any set of Lebesgue measure > 0. The true relationship between the laws of m and T_1+T_2 is as follows: for any Borel function $f : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ we have

$$n_1\left[f(m^2)\right] = \sqrt{\frac{\pi}{2}} \mathbb{E}\left[\sqrt{T^1 + T^2} f\left(\frac{1}{T^1 + T^2}\right)\right]. \tag{38}$$

To verify the above equality, we write the following equalities for a positive function g, successively applying the identities (36) and (37), and the properties of the scale change (34).

$$
\begin{aligned}
n_+\left(f(M^2/V)g(V)\right) &= \int_0^\infty n_v\left(f(M^2/v)g(v)\right)\frac{2\,dv}{\sqrt{2\pi v^3}} \\
&= n_1(f(m^2))\int_0^\infty g(v)\frac{2\,dv}{\sqrt{2\pi v^3}} \\
&= n^x\left(f(x^2/(T_x^1+T_x^2))\right) g(T_x^1+T_x^2)\frac{2dx}{x^2} \\
&= \mathbb{E}\left[f\left(\frac{1}{T^1+T^2}\right)\int_0^\infty \left(g\left(x^2(T^1+T^2)\right)\right)\frac{2dx}{x^2}\right] \qquad (39)\\
&= \mathbb{E}\left[\sqrt{T^1+T^2}f\left(\frac{1}{T^1+T^2}\right)\right]\int_0^\infty g(v)\frac{dv}{v^{3/2}}. \qquad (40)
\end{aligned}
$$

The equality of lines 2 and 5 gives the result.

You can also write the relation (38) in terms of the densities of the laws of m and $\sqrt{T^1+T^2}$. If we call these densities Ψ_1 and Ψ_2 respectively, then we have the relation

$$\Psi_1(x) = \sqrt{\frac{\pi}{2}}x^{-3}\Psi_2(x^{-1}). \qquad (41)$$

In the above discussion we have not used an explicit knowledge of the laws of m and T_1+T_2. Thus. we can consider the relation (41) as a consequence of the scaling invariance of Brownian motion and the Ito excursion measure. If we now recall that the law of $\frac{\sqrt{\pi/2}}{m}$ has been calculated previously in (33), and its density is Ψ, which satisfies (19) we observe the curious identity

$$m^2 \stackrel{\text{law}}{=} \frac{\pi}{2}\left(T^1+T^2\right). \qquad (42)$$

The two identities (41) and (42) are equivalent, in view of (33) to the functional equation (19). Finally, since (41) is an immediate consequence of the scale invariance of Brownian motion, we see that it is the identity (42) that should be considered as the probabilistic basis for the functional equation of the Riemann zeta function. I know of no direct demonstration of this identity, which does not involve an explicit calculation of the laws in question. It would be very interesting to have a purely combinatorial demonstration of this identity, through manipulation of the paths of Brownian motion.

3.5 An approximation of the zeta function

We conclude this little voyage to the land of probabilities by an unexpected application of the ideas developed so far. We use simple probabilistic considerations to obtain a renormalization of the series $\sum_n (-1)^n/n^s$ that converges in the full complex plane to the entire function $(2^{1-s}-1)\zeta(s)$. More precisely, we will determine the coefficients $(a_{n,N}; 0 \le n \le N)$ such that for every n we have $\lim_{N\to\infty} a_{n,N} = (-1)^n$, and the partial sums $\sum_{n=1}^N a_{n,N}/n^s$ converge uniformly for s in compact sets to the entire function $(2^{1-s}-1)\zeta(s)$. Recall that this entire function is the sum of the series $\sum_{n=1}^\infty (-1)^n/n^s$, convergent for $\Re(s) > 1$. We can choose the coefficients $a_{n,N}$ in order to fix the value of the sum $\sum_{n=1}^N a_{n,N}/n^s$ at N values of s. It is natural to choose for these N values $s = 0, -2, -4, \ldots - 2(N-1)$ where the zeta function vanishes. It is not difficult to see that this implies

$$a_{n,N} = (-1)^n \frac{(N!)^2}{(N-n)!(N+n)!},$$

and we have as well

$$\lim_{N\to\infty} a_{n,N} = (-1)^n.$$

We now relate this renormalization of the series $\sum_n (-1)^n/n^s$ to the preceding considerations. First, the non-convergence of the series $\sum_{n=1}^\infty n^{-s}$ for $\Re s < 1$, based on the Ψ, relies on the fact that the series (18) does not converge uniformly on $\mathbb{R}_+$. In fact it is easy to see that

$$\min_{y\in[0,\varepsilon]} 4y \sum_{n=1}^N \left(2\pi^2 n^4 y^2 - 3\pi n^2\right) e^{-\pi n^2 y^2} \stackrel{N\to\infty}{\longrightarrow} -\infty,$$

for every $\varepsilon > 0$ (the convergence is also uniform). In particular, the partial sum is not positive. We will therefore seek a probabilistic approximation by simpler random variables of the function Ψ approaching a random variable with density Ψ. For this recall the Laplace transform, which can deduced from (30) and (42), or calculated directly by integrating (18) term by term

$$\int_0^\infty e^{-\lambda y^2} \Psi(y)\, dy = \left(\frac{\sqrt{\pi\lambda}}{\sinh\sqrt{\pi\lambda}}\right)^2.$$

Euler's formula

$$\frac{\pi x}{\sinh \pi x} = \prod_{n=1}^\infty \left(1 + \frac{x^2}{n^2}\right)^{-1},$$

and the elementary formula

$$\mathbb{E}\left[e^{-\kappa E}\right] = \int_0^\infty e^{-t} e^{-\kappa t}\, dt = (1+\kappa)^{-1},$$

for the Laplace transform of a standard, exponential random variable with law $P(E > t) = e^{-t}$ show that we have the equality in law

$$X^2 = \pi \sum_{n=1}^{\infty} \frac{E_n + E'_n}{n^2}.$$

Here X is a random variable with density Ψ and E_n and E'_n are independent, standard exponential random variables. It is thus natural to try to approximate the variable X^2 by the partial sums of $\pi \sum_{n=1}^{\infty} \frac{E_n+E'_n}{n^2}$. This leads to an approximation of the function $(1-s)\zeta(s)$ convergent in the whole complex plane. However, the calculations are more complicated than in the simple case that we will consider, which is the random variable

$$Y = \sum_{n=1}^{\infty} \frac{E_n}{n^2}$$

which satisfies

$$\mathbb{E}\left[Y^s\right] = s(1 - 2^{1-s})\Gamma(s/2)\zeta(s)$$

and which can be approximated by the partial sums $\sum_{n=1}^{N} E_n/n^2$. By breaking the product $\prod_{n=1}^{N} \left(1 + \frac{x^2}{n^2}\right)^{-1}$ into simple rational fractions we obtain the formula

$$\mathbb{E}\left[\left(\sum_{n=1}^{N} \frac{E_n}{n^2}\right)^{s/2}\right] = -s\Gamma(s/2) \sum_{n=1}^{N} \frac{a_{n,N}}{n^s}.$$

It is then easy to deduce using Hölder's inequality that

$$\sum_{n=1}^{N} \frac{a_{n,N}}{n^s} \to \left(2^{1-s} - 1\right) \zeta(s),$$

unifomrly over compact subsets of $\mathbb{C}$.

References

[1] P. Biane, J. Pitman, and M. Yor, *Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions*, Bulletin of the American Mathematical Society, 38 (2001), pp. 435–465.

[2] E. Brézin and S. Hikami, *Characteristic polynomials of random matrices*, Communications in Mathematical Physics, 214 (2000), pp. 111–135.

[3] N. de Bruijn, D. Knuth, and S. Rice, *The average height of planted plane trees*, Graph Theory and Computing/Ed. R.C. Read, (1972), p. 15.

[4] P. Deift, *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, vol. 3, American Mathematical Soc., 2000.

[5] I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*, vol. 113, springer, 1991.

[6] J. P. Keating and N. C. Snaith, *Random matrix theory and* $\zeta(1/2+it)$, Communications in Mathematical Physics, 214 (2000), pp. 57–89.

[7] M. L. Mehta, *Random matrices*, vol. 142, Academic press, 2004.

[8] P. Michel, *Répartition des zéros des fonctions-L et matrices aléatoires. Séminaire Bourbaki*, 2001.

[9] H. L. Montgomery, *The pair correlation of zeros of the zeta function*, in Proc. Symp. Pure Math, vol. 24, 1973, pp. 181–193.

[10] A. Odlyzko, *On the distribution of spacings between zeros of the zeta function*, Mathematics of Computation, 48 (1987), pp. 273–308.

[11] G. Polya, *Elementarer beweis einer thetaformel*, Collected papers of George Polya, (1974), pp. 157–161.

[12] L. Smith and P. Diaconis, *Honest Bernoulli excursions*, Journal of applied probability, (1988), pp. 464–477.

[13] G. Szegö, *Orthogonal Polynomials*, vol. 23, American Mathematical Soc., 1939.

[14] D. Williams, *Brownian motion and the Riemann zeta-function*, Disorder in physical systems, (1990), pp. 361–372.