10.1098/rspa.2000.0628

THE ROYAL SOCIETY

Random matrix theory and the derivative of the Riemann zeta function

By C. P. Hughes[1,2], J. P. Keating[1,2] and Neil O'Connell[2]

[1] *School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK*

[2] *BRIMS, Hewlett-Packard Laboratories, Filton Road, Stoke Gifford, Bristol BS34 8QZ, UK*

Received 21 February 2000; accepted 6 June 2000

Random matrix theory is used to model the asymptotics of the discrete moments of the derivative of the Riemann zeta function, $\zeta(s)$, evaluated at the complex zeros $\frac{1}{2}+\mathrm{i}\gamma_n$. We also discuss the probability distribution of $\ln|\zeta'(1/2+\mathrm{i}\gamma_n)|$, proving the central limit theorem for the corresponding random matrix distribution and analysing its large deviations.

Keywords: discrete moments; random matrix theory; Riemann zeta function

1. Introduction

Let $\zeta(s)$ be Riemann's zeta function, with (assuming the Riemann hypothesis) its complex zeros denoted by $\frac{1}{2}+\mathrm{i}\gamma_n$ with γ_n increasing with n, and $\gamma_1 = 14.13\ldots$. The purpose of this paper is to develop the random matrix model of Keating & Snaith (2000) in order to study the discrete moments of $\zeta'(s)$,

$$J_k(T) = \frac{1}{N(T)} \sum_{0<\gamma_n \leqslant T} |\zeta'(\tfrac{1}{2}+\mathrm{i}\gamma_n)|^{2k}, \tag{1.1}$$

where

$$N(T) = \sum_{0<\gamma_n \leqslant T} 1 \tag{1.2}$$

$$= \frac{T}{2\pi}\ln\frac{T}{2\pi\mathrm{e}} + O(\ln T). \tag{1.3}$$

$J_k(T)$ is clearly defined for all $k \geqslant 0$, and, on the additional assumption that all the zeros are simple, for all $k<0$. It has previously been studied by Gonek (1984, 1989, 1999) and Hejhal (1989), and is discussed in Odlyzko (1992, § 2.12) and Titchmarsh (1986, § 14).

The model proposed by Keating & Snaith (2000) is the characteristic polynomial of an $N \times N$ unitary matrix U with eigenangles θ_n,

$$Z(\theta) = \det(I - U\mathrm{e}^{-\mathrm{i}\theta}) \tag{1.4}$$

$$= \prod_{n=1}^{N} (1 - \mathrm{e}^{\mathrm{i}(\theta_n - \theta)}), \tag{1.5}$$

which can be considered as a continuous family of random variables (parametrized by θ), the probability state space being the group $U(N)$ of all $N \times N$ unitary matrices with Haar measure (the probability density being denoted by $\mathrm{d}\mu_N$). In the physics literature, this state space is referred to as the circular unitary ensemble, or CUE (see, for example, Mehta 1991).

They found that equating the mean densities of zeros and eigenangles, that is setting

$$N = \ln(T/2\pi), \tag{1.6}$$

the CUE statistics of $Z(\theta)$ model the local statistics of $\zeta(s)$ well. For example, the value distribution of $\ln|\zeta(1/2 + \mathrm{i}t)|$ high up the critical line is correctly predicted. Coram & Diaconis (1999) have subsequently verified that making the identification (1.6) leads to close agreement for other statistical measures. Also in favour of the model is the fact that theorems in restricted ranges (Montgomery 1973; Rudnick & Sarnak 1996), numerical evidence (Odlyzko 1992), and heuristic calculations (Bogomolny & Keating 1995, 1996) support the conjecture that the n-point correlation function of the Riemann zeros is asymptotically the same as the n-point correlation function of CUE eigenvalues for all n.

However, it appears that global statistics, like moments of $|\zeta(1/2 + \mathrm{i}t)|$ (rather than of $\ln|\zeta(1/2+\mathrm{i}t)|$) are not modelled precisely by random matrix theory (RMT). Indeed, we have the following conjecture.

Conjecture 1.1 (Keating & Snaith 2000). *For $k > -1/2$ fixed,*

$$\frac{1}{T}\int_0^T |\zeta(\tfrac{1}{2} + \mathrm{i}t)|^{2k}\,\mathrm{d}t \sim a(k)f(k)\left(\ln\frac{T}{2\pi}\right)^{k^2}, \tag{1.7}$$

as $T \to \infty$, where

$$a(k) = \prod_{\substack{p\\ \text{prime}}} \left(1 - \frac{1}{p}\right)^{k^2} \sum_{m=0}^{\infty} \left(\frac{\Gamma(m+k)}{m!\,\Gamma(k)}\right)^2 p^{-m} \tag{1.8}$$

is the zeta-function-specific (non-universal) part, and

$$f(k) = \lim_{N\to\infty} N^{-k^2} M_N(2k) \tag{1.9}$$

$$= \frac{G^2(1+k)}{G(1+2k)} \tag{1.10}$$

is the random-matrix (universal) part. Here, for $\mathrm{Re}(k) > -1/2$,

$$M_N(2k) = \int_{U(N)} |Z(\theta)|^{2k}\,\mathrm{d}\mu_N \tag{1.11}$$

$$= \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+2k)}{(\Gamma(j+k))^2} \tag{1.12}$$

is independent of θ, and $G(k)$ is the Barnes G-function (see the appendix).

This is in line with previous results for other statistics, where long-range deviations from RMT have also been related to the primes (Berry 1988; Berry & Keating 1999).

In the present paper we consider

$$\int_{U(N)} \frac{1}{N}\sum_{n=1}^{N} |Z'(\theta_n)|^{2k}\,\mathrm{d}\mu_N, \tag{1.13}$$

in the hope that it gives information about the universal part of $J_k(T)$.

We prove the following theorem.

Theorem 1.2. *For* $\mathrm{Re}(k) > -3/2$ *and bounded,*

$$\int_{U(N)} \frac{1}{N}\sum_{n=1}^{N} |Z'(\theta_n)|^{2k}\,\mathrm{d}\mu_N = \frac{G^2(k+2)}{G(2k+3)}\frac{G(N+2k+2)G(N)}{NG^2(N+k+1)} \tag{1.14}$$

$$\sim \frac{G^2(k+2)}{G(2k+3)} N^{k(k+2)}, \quad \text{as } N\to\infty. \tag{1.15}$$

Heuristic arguments then lead us to the following conjecture.

Conjecture 1.3. *For* $k > -3/2$ *and bounded,*

$$J_k(T) \sim \frac{G^2(k+2)}{G(2k+3)} a(k)\left(\ln\frac{T}{2\pi}\right)^{k(k+2)}, \tag{1.16}$$

as $T\to\infty$*, where* $a(k)$ *is given in* (1.8).

This conjecture agrees with all previously known results about $J_k(T)$, which are reviewed in § 2 *b* (i).

Some work has been done on the limiting distribution of $\ln|\zeta'(\gamma_n)|$. In particular, we note the following theorem.

Theorem 1.4 (Hejhal 1989). *If one assumes the Riemann hypothesis (RH) and the existence of an* α *such that*

$$\limsup_{T\to\infty} \frac{1}{N(2T)-N(T)}\left|\left\{n : T \leqslant \gamma_n \leqslant 2T,\ 0 \leqslant \gamma_{n+1}-\gamma_n \leqslant \frac{c}{\ln T}\right\}\right| \leqslant Mc^{\alpha} \tag{1.17}$$

holds uniformly for $0 < c < 1$*, with* M *a suitable constant, then, for* $a < b$*,*

$$\lim_{T\to\infty} \frac{1}{N(2T)-N(T)} \times \left|\left\{n : T \leqslant \gamma_n \leqslant 2T,\ \frac{1}{\sqrt{\frac{1}{2}\ln\ln T}} \ln\left|\frac{\zeta'(1/2+\mathrm{i}\gamma_n)}{(1/2\pi)\ln(\gamma_n/2\pi)}\right| \in (a,b)\right\}\right| = \frac{1}{\sqrt{2\pi}}\int_a^b \mathrm{e}^{-x^2/2}\,\mathrm{d}x. \tag{1.18}$$

In the same direction, we have the following theorem.

Theorem 1.5. *For* $a < b$*,*

$$\lim_{N\to\infty} \mathrm{P}\left\{ \frac{\ln\left|\frac{Z'(\theta_1)}{N\exp(\gamma-1)}\right|}{\sqrt{\frac{1}{2}(\ln N + 3 + \gamma - \pi^2/2)}} \in (a,b)\right\} = \frac{1}{\sqrt{2\pi}}\int_a^b \mathrm{e}^{-x^2/2}\,\mathrm{d}x. \tag{1.19}$$

In §§ 3 *b* and 3 *c* we study the asymptotics of the tails of this distribution, when scaled by a factor much greater than $\sqrt{\ln N}$, in § 3 *b* using large deviation theory, and in § 3 *c* using more refined asymptotic methods.

Throughout this paper we use the notation $f \ll g$ to denote $f/g \to 0$ and $f \gg g$ to denote $f/g \to \infty$. We write $f \asymp g$ whenever $f = O(g)$ and $g = O(f)$.

2. The discrete moments

(*a*) *The random matrix moments*

Proof of theorem 1.2. Differentiating $Z(\theta)$, we get

$$Z'(\theta) = \mathrm{i} \sum_{j=1}^{N} \mathrm{e}^{\mathrm{i}(\theta_j - \theta)} \prod_{\substack{m=1 \\ m \neq j}}^{N} (1 - \mathrm{e}^{\mathrm{i}(\theta_m - \theta)}), \tag{2.1}$$

and, therefore,

$$|Z'(\theta_n)| = \prod_{\substack{m=1 \\ m \neq n}}^{N} |\mathrm{e}^{\mathrm{i}\theta_m} - \mathrm{e}^{\mathrm{i}\theta_n}|. \tag{2.2}$$

The Haar probability density of $U(N)$ equals (Mehta 1991; Weyl 1946)

$$\mathrm{d}\mu_N = \frac{1}{N!(2\pi)^N} \prod_{1 \leqslant j < k \leqslant N} |\mathrm{e}^{\mathrm{i}\theta_j} - \mathrm{e}^{\mathrm{i}\theta_k}|^2 \prod_{p=1}^{N} \mathrm{d}\theta_p, \tag{2.3}$$

and so we may evaluate

$$\int_{U(N)} \frac{1}{N} \sum_{n=1}^{N} |Z'(\theta_n)|^{2k} \, \mathrm{d}\mu_N \tag{2.4}$$

as the N-fold integral

$$\int \cdots \int_{-\pi}^{\pi} \frac{1}{N!(2\pi)^N} \prod_{1 \leqslant j < k \leqslant N} |\mathrm{e}^{\mathrm{i}\theta_j} - \mathrm{e}^{\mathrm{i}\theta_k}|^2 \frac{1}{N} \sum_{n=1}^{N} \prod_{\substack{m=1 \\ m \neq n}}^{N} |\mathrm{e}^{\mathrm{i}\theta_m} - \mathrm{e}^{\mathrm{i}\theta_n}|^{2k} \prod_{p=1}^{N} \mathrm{d}\theta_p. \tag{2.5}$$

Due to the symmetry in the angles θ_n (the ones being summed over), we see that

$$\int_{U(N)} \frac{1}{N} \sum_{n=1}^{N} |Z'(\theta_n)|^{2k} \, \mathrm{d}\mu_N = \int_{U(N)} |Z'(\theta_N)|^{2k} \, \mathrm{d}\mu_N, \tag{2.6}$$

and so (2.5) equals

$$\int \cdots \int_{-\pi}^{\pi} \frac{1}{N!(2\pi)^N} \prod_{1 \leqslant j < k \leqslant N} |\mathrm{e}^{\mathrm{i}\theta_j} - \mathrm{e}^{\mathrm{i}\theta_k}|^2 \prod_{m=1}^{N-1} |\mathrm{e}^{\mathrm{i}\theta_m} - \mathrm{e}^{\mathrm{i}\theta_N}|^{2k} \prod_{p=1}^{N} \mathrm{d}\theta_p. \tag{2.7}$$

Putting all the terms from the first product with a factor $|\mathrm{e}^{\mathrm{i}\theta_j}-\mathrm{e}^{\mathrm{i}\theta_N}|^2$ in them into the second product gives

$$\int\cdots\int_{-\pi}^{\pi}\frac{1}{N!(2\pi)^N}\prod_{1\leqslant j<k\leqslant (N-1)}|\mathrm{e}^{\mathrm{i}\theta_j}-\mathrm{e}^{\mathrm{i}\theta_k}|^2\prod_{m=1}^{N-1}|\mathrm{e}^{\mathrm{i}\theta_m}-\mathrm{e}^{\mathrm{i}\theta_N}|^{2k+2}\prod_{p=1}^{N}\mathrm{d}\theta_p. \tag{2.8}$$

Integrating first over $\theta_1\cdots\theta_{N-1}$ and then over θ_N,

$$\int_{U(N)}\frac{1}{N}\sum_{n=1}^{N}|Z'(\theta_n)|^{2k}\,\mathrm{d}\mu_N=\int_{-\pi}^{\pi}\left\{\frac{1}{2\pi N}\int_{U(N-1)}|Z(\theta_N)|^{2k+2}\,\mathrm{d}\mu_{N-1}\right\}\mathrm{d}\theta_N \tag{2.9}$$

$$=\frac{1}{N}\prod_{j=1}^{N-1}\frac{\Gamma(j)\Gamma(j+2k+2)}{(\Gamma(j+k+1))^2}, \tag{2.10}$$

which is valid for $\mathrm{Re}(k)>-3/2$. (The evaluation of $\int_{U(N-1)}|Z(\theta_N)|^{2k+2}\,\mathrm{d}\mu_{N-1}$ is essentially (1.12), since it is independent of θ_N.)

From the recurrence relation for the G-function (see the appendix), (2.10) equals

$$\frac{G^2(k+2)}{G(2k+3)}\frac{1}{N}\frac{G(N+2k+2)G(N)}{G^2(N+k+1)}. \tag{2.11}$$

Assuming k to be bounded, then as $N\to\infty$, the asymptotics for G (A 2) imply

$$\frac{1}{N}\frac{G(N+2k+2)G(N)}{G^2(N+k+1)}=N^{k(k+2)}\left(1+O\left(\frac{1}{N}\right)\right). \tag{2.12}$$

This proves theorem 1.2. ■

Remark 2.1. If k is a non-negative integer, then the recurrence relation for G implies

$$\frac{G^2(k+2)}{G(2k+3)}=\prod_{j=0}^{k}\frac{j!}{(k+1+j)!}. \tag{2.13}$$

Remark 2.2. By comparing the Taylor expansions of both sides, one can show that

$$\frac{G^2(k+2)}{G(2k+3)}=\frac{\exp(3\zeta'(-1)+\ln\pi-\frac{11}{12}\ln 2+k\ln\pi-3k\ln 2-2k^2\ln 2)}{\Gamma(k+\frac{3}{2})G^2(k+\frac{3}{2})}, \tag{2.14}$$

which has the advantage of making the poles at $k=-\frac{1}{2}(2n+1)$, $n=1,2,3,\ldots$, explicit. (The poles are of order $2n-1$.)

The existence of a pole at $k=-3/2$ in (2.10) means that the random matrix average diverges (for any $N\geqslant 2$) for $\mathrm{Re}(k)\leqslant -3/2$. Its analytic continuation into this region is given by (2.10).

(*b*) *A heuristic analysis of* $J_k(T)$

Define, for $x > 0$, with $x \gg \ln\ln T$

$$P(T,x) = \frac{1}{T}|\{t : 0 \leqslant t \leqslant T,\ \ln|\zeta(\tfrac{1}{2}+\mathrm{i}t)| \leqslant -x\}|, \tag{2.15}$$

so $P(T,x)$ is the proportion of space $0 \leqslant t \leqslant T$, where $|\zeta(\frac{1}{2}+\mathrm{i}t)| \leqslant \mathrm{e}^{-x}$.

In the limit as $x \to \infty$, the regions in $0 \leqslant t \leqslant T$ where $|\zeta(\frac{1}{2}+\mathrm{i}t)| \leqslant \mathrm{e}^{-x}$ each contain exactly one zero, provided all the zeros are simple. At such a zero, we wish to solve $|\zeta(\frac{1}{2}+\mathrm{i}(\gamma_n+\epsilon))| = \mathrm{e}^{-x}$ for ϵ. To do this, we Taylor expand the zeta function, then take the modulus, obtaining

$$|\zeta(\tfrac{1}{2}+\mathrm{i}\gamma_n+\mathrm{i}\epsilon)| = |\epsilon|\,|\zeta'(\tfrac{1}{2}+\mathrm{i}\gamma_n)| + O_T(\epsilon^2), \tag{2.16}$$

which equals e^{-x} when

$$|\epsilon| = \frac{\mathrm{e}^{-x}}{|\zeta'(\frac{1}{2}+\mathrm{i}\gamma_n)|} + O_T(\mathrm{e}^{-2x}), \tag{2.17}$$

and so the length of each region is $2|\epsilon| + O_T(\epsilon^2)$.

Thus,

$$\lim_{x\to\infty} \mathrm{e}^{x} P(T,x) = \frac{2}{T}\sum_{0<\gamma_n\leqslant T} |\zeta'(\tfrac{1}{2}+\mathrm{i}\gamma_n)|^{-1}. \tag{2.18}$$

A different evaluation of $P(T,x)$ comes from conjecture 1.1, which suggests that, for large T,

$$P(T,x) \sim \int_{-\infty}^{-x} \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}yz}\frac{G^2(1+\frac{1}{2}\mathrm{i}y)}{G(1+\mathrm{i}y)}\left(\ln\frac{T}{2\pi}\right)^{-y^2/4} a(\tfrac{1}{2}\mathrm{i}y)\,\mathrm{d}y\mathrm{d}z, \tag{2.19}$$

for $N = \ln(T/2\pi)$.

For $x > 0$, calculating the Fourier integral by the residue theorem (cf. § 3 *c*), we find that for x sufficiently large ($x \gg \ln\ln T$), $P(T,x)$ is dominated by the (simple) pole at $y = \mathrm{i}$,

$$P(T,x) = \mathrm{e}^{-x}G^2(\tfrac{1}{2})\left(\ln\frac{T}{2\pi}\right)^{1/4} a(-\tfrac{1}{2}) \operatorname*{Res}_{s=0}\{(G(s))^{-1}\} + O_T(\mathrm{e}^{-3x+\epsilon}), \tag{2.20}$$

and so, as $T \to \infty$,

$$\lim_{x\to\infty} \mathrm{e}^{x} P(T,x) \sim \left(\ln\frac{T}{2\pi}\right)^{1/4} \exp(3\zeta'(-1) + \tfrac{1}{12}\ln 2 - \tfrac{1}{2}\ln\pi)a(-\tfrac{1}{2}). \tag{2.21}$$

Combining (2.18) and (2.21), we obtain the following conjecture.

Conjecture 2.3.

$$J_{-1/2}(T) \sim \exp(3\zeta'(-1) + \tfrac{1}{12}\ln 2 + \tfrac{1}{2}\ln\pi)a(-\tfrac{1}{2})\left(\ln\frac{T}{2\pi}\right)^{-3/4}. \tag{2.22}$$

Note that for $k=-1/2$, the random matrix moment (1.15) is asymptotic to

$$\exp(3\zeta'(-1)+\tfrac{1}{12}\ln 2+\tfrac{1}{2}\ln\pi)N^{-3/4}, \tag{2.23}$$

as $N\to\infty$. Since $a(k)$ is exactly the zeta-function-specific term in conjecture 1.1, and $N=\ln(T/2\pi)$, this in turn leads us to conjecture 1.3, that as $T\to\infty$,

$$J_k(T)\sim\frac{G^2(k+2)}{G(2k+3)}a(k)\left(\ln\frac{T}{2\pi}\right)^{k(k+2)}, \tag{2.24}$$

for $k>-3/2$ fixed.

(i) *Comparison with known results*

If the tails of the distribution (1.18) are sufficiently small, one might expect (Gonek 1989; Hejhal 1989)

$$J_k(T)\asymp(\ln T)^{k(k+2)}. \tag{2.25}$$

We show in § 3 *c* that the singularity at $k=-3/2$ in (2.24) comes from a large left tail of the distribution of $\ln|Z'(\theta_1)|$.

Under RH, Gonek (1984) has proved that $J_1(T)\sim\frac{1}{12}(\ln T)^3$. Under the additional assumption that all the zeros are simple, he has conjectured that $J_{-1}(T)\sim(6/\pi^2)(\ln T)^{-1}$ (Gonek 1989).

We observe that our conjecture agrees with all these results.

(*c*) *Discussion on the 'pole' at* $k=-3/2$

Due to the divergence of the random matrix average, conjecture 1.3 is restricted to $2k>-3$. In this section we argue that this restriction is necessary.

For k negative and $|k|$ large, the sum over zeros of the zeta function may be dominated by the few points where $|\zeta'(1/2+\mathrm{i}\gamma_n)|$ is close to zero. These points are expected to be where two zeros lie very close together (an occurrence of Lehmer's phenomena).

Gonek (1999) defined

$$\Theta=\inf\{\theta:|\zeta'(\tfrac{1}{2}+\mathrm{i}\gamma_n)|^{-1}=O(|\gamma_n|^{\theta}),\ \forall n\}. \tag{2.26}$$

He observed that RH implies $\Theta\geqslant 0$, provided all the zeros are simple, and that $\Theta\leqslant 1$ if the averaged Mertens hypothesis holds, that is if

$$\int_1^X\frac{1}{x^2}\left(\sum_{n\leqslant x}\mu(n)\right)^2\mathrm{d}x=O(\ln X), \tag{2.27}$$

where $\mu(n)$ is the Möbius function.

If Θ is finite, then there exists an infinite subsequence of the $\{\gamma_n\}$, such that, for all $\epsilon>0$,

$$|\zeta'(\tfrac{1}{2}+\mathrm{i}\gamma_n)|^{-1}>|\gamma_n|^{\Theta-\epsilon}. \tag{2.28}$$

Choosing a γ from this subsequence and setting $T=\gamma$, we have, for $k<0$,

$$J_k(T) > \frac{1}{N(T)}|\zeta'(\tfrac{1}{2}+\mathrm{i}\gamma)|^{2k} \tag{2.29}$$

$$> \frac{2\pi}{T\ln T}T^{-2k(\Theta-\epsilon)}. \tag{2.30}$$

If $\Theta>0$, then

$$\frac{2\pi}{T\ln T}T^{-2k(\Theta-\epsilon)} \gg (\ln T)^{k(k+2)} \tag{2.31}$$

when

$$2k < -1/\Theta, \tag{2.32}$$

implying that the conjectured scaling (2.25) is too small for $2k<-1/\Theta$. It follows from theorem 1.2 that the analogue of (2.25) for $Z'(\theta_1)$ fails for $2k\leqslant -3$, which implies, via conjecture 1.3, that $\Theta=1/3$. This is precisely the value conjectured by Gonek (1999), and is in line with the fact that the pair correlation conjecture of Montgomery (1973) suggests that $\Theta\geqslant 1/3$.

In the region $2k<-1/\Theta$, all we can say is that for any $\epsilon>0$

$$J_k(T) = \Omega(T^{2|k|\Theta-1-\epsilon}). \tag{2.33}$$

For $k<0$ we have the trivial upper bound of

$$J_k(T) = O(T^{2|k|\Theta+\epsilon}), \tag{2.34}$$

which comes from noting that $|\zeta'(1/2+\mathrm{i}\gamma_n)|^{-1}=O(|\gamma_n|^{\Theta+\epsilon})$ for all n.

Remark 2.4. If all the zeros are simple, then for $k\leqslant -3/2$, $J_k(T)$ is still defined, but our results do not predict its asymptotic behaviour. However, if one redefines $J_k(T)$ to exclude these rare points, where $|\zeta'(1/2+\mathrm{i}\gamma_n)|$ is very close to zero, then RMT should still predict the universal behaviour.

3. The distribution of $\ln|Z'(\theta_1)|$

(a) Central limit theorem

Proof of theorem 1.5. From theorem 1.2 we have

$$\int_{U(N)} \frac{1}{N}\sum_{n=1}^{N}|Z'(\theta_n)|^{\lambda}\,\mathrm{d}\mu_N = \int_{U(N)}|Z'(\theta_1)|^{\lambda}\,\mathrm{d}\mu_N \tag{3.1}$$

$$= \frac{G^2(2+\frac{1}{2}\lambda)}{G(3+\lambda)}\frac{G(N+2+\lambda)G(N)}{G^2(N+1+\frac{1}{2}\lambda)N} \tag{3.2}$$

$$= F(\lambda,N), \tag{3.3}$$

which we can think of as a moment-generating function for $\ln|Z'(\theta_1)|$.

By definition, the cumulants of this moment-generating function are

$$C_n = \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}\{\ln F(\lambda,N)\}\bigg|_{\lambda=0}. \tag{3.4}$$

Evaluating these, and asymptotically expanding for large N, we see that

$$C_1 = \Phi(2) - \Phi(3) + \Phi(N+2) - \Phi(N+1) \tag{3.5}$$
$$\sim \ln N + \gamma - 1, \tag{3.6}$$

$$C_2 = \tfrac{1}{2}\Phi^{(1)}(2) - \Phi^{(1)}(3) + \Phi^{(1)}(N+2) - \tfrac{1}{2}\Phi^{(1)}(N+1) \tag{3.7}$$
$$\sim \tfrac{1}{2}(\ln N + \gamma + 3 - \pi^2/2), \tag{3.8}$$

$$C_n = O(1), \quad \text{for } n \geqslant 3, \tag{3.9}$$

where

$$\Phi^{(n)}(x) = \frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}} \ln G(x). \tag{3.10}$$

This implies that the mean of the distribution is C_1 and the variance is C_2, with the other cumulants subdominant to the variance, which is sufficient to show that $(\ln|Z'(\theta_1)| - C_1)/\sqrt{C_2}$ converges in distribution (as $N \to \infty$) to a standard normal random variable (see, for example, Billingsley 1979, § 30).

Writing the result out explicitly, for $a < b$,

$$\lim_{N\to\infty} \mathrm{P}\left\{ \frac{\ln\left|\frac{Z'(\theta_1)}{N\exp(\gamma-1)}\right|}{\sqrt{\frac{1}{2}(\ln N + 3 + \gamma - \pi^2/2)}} \in (a,b) \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-x^2/2}\,\mathrm{d}x, \tag{3.11}$$

where $\mathrm{P}\{f(U) \in A\}$ is the probability that $f(U)$ lies in a set A and is defined to equal

$$\int_{U(N)} 1_{\{f(U)\in A\}}\,\mathrm{d}\mu_N, \tag{3.12}$$

where $1_{\{\cdot\}}$ is the indicator function. ■

Recalling (1.6), that $N = \ln(T/2\pi)$, this is in line with theorem 1.4. (Note that the $O(1)$ differences in the mean and variance are subdominant in the large-N, large-T limit.)

Odlyzko (1992) found numerically that, around the 10^{20}th zero, $\ln|\zeta'|$ had mean 3.35 and variance 1.14. Compare this with the leading-order asymptotic prediction in (1.18) of 1.91 and 1.89, and the above RMT prediction of 3.33 and 1.21, respectively.

(*b*) *Large deviations*

In this section we study the tails of the distribution of $\ln|Z'(\theta_1)|$, beyond the scope of the above central limit theorem. (In fact, we consider the random variable $\ln|Z'(\theta_1)/\exp(C_1)|$, since this has zero mean.)

Define a new family of random variables, R_N^A, by

$$R_N^A = \frac{\ln|Z'(\theta_1)/\exp(C_1)|}{A(N)}, \tag{3.13}$$

where $A(N)$ is a given function, much greater than $\sqrt{C_2}$.

Denote the logarithmic moment-generating function of R_N^A by

$$\Lambda_N(\lambda) = \ln \int_{U(N)} \mathrm{e}^{\lambda R_N^A} \, \mathrm{d}\mu_N \tag{3.14}$$

$$= \begin{cases} \ln F\left(\dfrac{\lambda}{A(N)}, N\right) - \dfrac{\lambda}{A(N)} C_1, & \text{for } \dfrac{\lambda}{A(N)} > -3, \\ \infty, & \text{for } \dfrac{\lambda}{A(N)} \leqslant -3. \end{cases} \tag{3.15}$$

A standard theorem in large deviation theory (see, for example, Bucklew 1990, § II.B) allows one to establish the log-asymptotics of the probability distribution of R_N^A. In order to apply this theorem, we need the following assumption.

Assumption 3.1. *There exists a function $B(N)$ (which tends to infinity as $N \to \infty$), such that*

$$\Lambda(\lambda) = \lim_{N\to\infty} \frac{1}{B(N)} \Lambda_N(B(N)\lambda) \tag{3.16}$$

exists as an extended real number, for each λ (i.e. the pointwise limit exists in the extended reals).

Definition 3.1. The effective domain of $\Lambda(\cdot)$ is

$$\mathcal{D} = \{\lambda \in \mathbb{R} : \Lambda(\lambda) < \infty\}, \tag{3.17}$$

and its interior is denoted by $\mathcal{D}^\circ$.

Definition 3.2. The Fenchel–Legendre transform of $\Lambda(\cdot)$ is

$$\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}} \{\lambda x - \Lambda(\lambda)\}. \tag{3.18}$$

Theorem 3.3. *If assumption 3.1 holds, then for $-\infty \leqslant a < b \leqslant \infty$,*

$$\limsup_{N\to\infty} \frac{1}{B(N)} \ln \mathrm{P}\{R_N^A \in [a,b]\} \leqslant - \inf_{x\in[a,b]} \Lambda^*(x). \tag{3.19}$$

If, in addition, $\Lambda(\cdot)$ is differentiable in $\mathcal{D}^\circ$, and $(a,b) \subseteq \{\Lambda'(\lambda) : \lambda \in \mathcal{D}^\circ\}$, then

$$\lim_{N\to\infty} \frac{1}{B(N)} \ln \mathrm{P}\{R_N^A \in (a,b)\} = - \inf_{x\in(a,b)} \Lambda^*(x). \tag{3.20}$$

We say that R_N^A satisfies the *large deviation principle* (LDP) at speed $B(N)$ with rate function $\Lambda^*(\cdot)$ if (3.20) holds.

Thus, for example, if $a > 0$, then

$$\mathrm{P}\left\{\ln\left|\frac{Z'(\theta_1)}{N\exp(\gamma-1)}\right| > aA(N)\right\} \approx \mathrm{e}^{-B(N)\Lambda^*(a)}, \tag{3.21}$$

where the meaning of the symbol $\approx$ is made precise in the equation above.

In order to apply theorem 3.3, we need to know the leading-order asymptotics of $\Lambda_N(B\lambda)$. Writing $\eta(N)=\lambda B(N)/A(N)$ for simplicity, then (3.15), (3.3), (3.6) and (A 2) imply that as $N\to\infty$, for $\eta>-3$,

$$\begin{aligned}\Lambda_N(B\lambda) = {}& \tfrac{1}{2}(N+\eta+1)^2\ln(N+\eta)+\tfrac{1}{2}(N-1)^2\ln N-(N+\tfrac{1}{2}\eta)^2\ln(N+\tfrac{1}{2}\eta)\\ &-\tfrac{1}{12}\ln(N+\eta)-\tfrac{13}{12}\ln N+\tfrac{1}{6}\ln(N+\tfrac{1}{2}\eta)-\eta\ln N\\ &+2\ln G(2+\tfrac{1}{2}\eta)-\ln G(3+\eta)-\gamma\eta-\tfrac{3}{8}\eta^2+O\left(\frac{1}{N}\right). \qquad (3.22)\end{aligned}$$

This can be simplified if we restrict $B(N)/A(N)$ to various regimes, and hence we are able to find $B(N)$ and $\Lambda^*(x)$ as follows.

Case 1. $1/\ln N \ll B/A \ll 1$:

$$\frac{1}{B}\Lambda_N(B\lambda)=\frac{B}{4A^2}\lambda^2\ln N+O\left(\frac{1}{BN}\right)+O\left(\frac{1}{A}\right), \qquad (3.23)$$

so we take $B=A^2/\ln N$, and obtain $\Lambda^*(x)=x^2$.
This is valid for $\sqrt{\ln N}\ll A\ll\ln N$.

Case 2. $B/A=1$:

$$\frac{1}{B}\Lambda_N(B\lambda)=\begin{cases}\dfrac{B}{4A^2}\lambda^2\ln N+O\left(\dfrac{1}{B}\right), & \text{if } \lambda>-3,\\ \infty, & \text{if } \lambda\leqslant-3.\end{cases} \qquad (3.24)$$

If $A=\ln N$, then the supremum of $\{\lambda x-\Lambda(\lambda)\}$ occurs at

$$\lambda=\begin{cases}2x, & \text{if } x\geqslant-\tfrac{3}{2},\\ -3, & \text{if } x\leqslant-\tfrac{3}{2},\end{cases} \qquad (3.25)$$

and so,

$$\Lambda^*(x)=\begin{cases}x^2, & \text{if } x\geqslant-\tfrac{3}{2},\\ -3x-\tfrac{9}{4}, & \text{if } x\leqslant-\tfrac{3}{2}.\end{cases} \qquad (3.26)$$

Remark 3.4. Note that in this case, theorem 3.3 only gives the upper bound (3.19) on the probabilities for $x<-\tfrac{3}{2}$.

If we keep the condition $B/A=1$, but have $A\gg\ln N$, then

$$\frac{1}{B}\Lambda_N(B\lambda)\to\begin{cases}0, & \text{for } \lambda>-3,\\ \infty, & \text{for } \lambda\leqslant-3,\end{cases} \quad \text{as } N\to\infty, \qquad (3.27)$$

and thus,

$$\Lambda^*(x)=\begin{cases}\infty, & \text{for } x>0,\\ -3x, & \text{for } x<0.\end{cases} \qquad (3.28)$$

Remark 3.5. Again, theorem 3.3 only gives the upper bound on the probabilities for $x < 0$. However, in § 3 *c* the probability density is evaluated in such a way as to prove the full LDP (3.20). (We obtain, in fact, a much stronger result: the asymptotics of the probability density function, not just the log-asymptotics.)

Remark 3.6. The fact that the rate function is infinite for $x > 0$ means that for $A \gg \ln N$ the deviations to the right ($x > 0$) tend to zero much faster than the deviations to the left ($x < 0$). We will now study these far-right deviations.

Case 3. $\lambda > 0$ with $1 \ll B/A$ and $\ln(B/A) \sim \epsilon \ln N$ with $0 \leqslant \epsilon < 1$ fixed:

$$\frac{1}{B} \Lambda_N(B\lambda) = \frac{B}{4A^2} \lambda^2 (1-\epsilon) \ln N + O\left(\frac{B}{A^2}\right). \tag{3.29}$$

Hence, we require $B = A^2/\ln N$.
The rate function is, therefore:

$$\Lambda^*(x) = \begin{cases} \dfrac{x^2}{1-\epsilon}, & \text{if } x \geqslant 0, \\ 0, & \text{if } x \leqslant 0. \end{cases} \tag{3.30}$$

This is valid for $A \gg \ln N$ but $\limsup_{N\to\infty}(\ln A/\ln N) < 1$.

Case 4. $\lambda > 0$ with $B/A = N$:

$$\frac{1}{B}\Lambda_N(B\lambda) = \frac{N^2}{B}\{\tfrac{1}{2}(1+\lambda)^2 \ln(1+\lambda) - (1+\tfrac{1}{2}\lambda)^2 \ln(1+\tfrac{1}{2}\lambda) - \tfrac{1}{4}\lambda^2 \ln 2\lambda\} + O\left(\frac{N \ln N}{B}\right). \tag{3.31}$$

Hence, we require $B = N^2$, which means $A = N$, and so the rate function is

$$I_c(x) = \sup_{\lambda>0}\{\lambda x - \tfrac{1}{2}(1+\lambda)^2 \ln(1+\lambda) + (1+\tfrac{1}{2}\lambda)^2 \ln(1+\tfrac{1}{2}\lambda) + \tfrac{1}{4}\lambda^2 \ln 2\lambda\}. \tag{3.32}$$

Assuming $x \geqslant 0$, the supremum occurs when

$$x = \tfrac{1}{2}\lambda \ln\left(\frac{(\lambda+1)^2}{\lambda(\lambda+2)}\right) + \ln\left(\frac{\lambda+1}{\lambda+2}\right) + \ln 2. \tag{3.33}$$

Note that the right-hand side is an increasing function of λ (for $\lambda > 0$) bounded between 0 and $\ln 2$. This means, for $x \geqslant \ln 2$ the supremum is ∞, implying that any scaling $A(N)$ greater than N has rate function ∞, independent of x.
Hence,

$$\Lambda^*(x) = \begin{cases} 0, & \text{for } x \leqslant 0, \\ I_c(x), & \text{for } 0 \leqslant x \leqslant \ln 2, \\ \infty, & \text{for } x \geqslant \ln 2. \end{cases} \tag{3.34}$$

This only leaves the regime $x > 0$ with $\lim_{N\to\infty}(\ln A/\ln N) = 1$ but $A \ll N$ unconsidered. This can be calculated in a similar way to the above.

Conclusion for deviations to the right ($x > 0$)

scaling $A(N)$	speed $B(N)$	rate function $\Lambda^*(x)$
$A \gg \sqrt{\ln N}$ but $\ln A \ll \ln N$	$A^2/\ln N$	x^2
$\ln A \sim \epsilon \ln N$, $\epsilon < 1$	$A^2/\ln N$	$x^2/(1-\epsilon)$
$A = N$	N^2	$\begin{cases} I_c(x), & \text{if } 0 \leqslant x \leqslant \ln 2, \\ \infty, & \text{if } x \geqslant \ln 2 \end{cases}$

Conclusion for deviations to the left ($x < 0$)

scaling $A(N)$	speed $B(N)$	rate function $\Lambda^*(x)$
$\sqrt{\ln N} \ll A \ll \ln N$	$A^2/\ln N$	x^2
$A = \ln N$	$\ln N$	$\begin{cases} x^2, & \text{if } -\frac{3}{2} \leqslant x \leqslant 0, \\ 3\lvert x\rvert - \frac{9}{4}, & \text{if } x \leqslant -\frac{3}{2} \end{cases}$
$A \gg \ln N$	A	$3\lvert x\rvert$

Remark 3.7. Note that the LDP for the deviations to the right is identical to that found for $\ln|Z(\theta)|$ in Hughes *et al.* (2000). The LDP for deviations to the left is very similar, but the rate function there is linear for $x < -1/2$ rather than $x < -3/2$.

Remark 3.8. This later arrival of the linear rate function is consistent with the observation that the value distribution $\ln|\zeta'|$ is closer to the standard normal curve than $\ln|\zeta|$ in Odlyzko's numerical data (see Odlyzko 1992, p. 55).

(*c*) *Refined asymptotics for deviations to the left*

Due to the singularity in $\Lambda(\lambda)$ for $\lambda < -3$ and $A \geqslant \ln N$, theorem 3.3 only gives the upper bound on the probabilities (3.19). In order to complete the proof of the LDP in this region (that is, to prove (3.20)), we will actually prove a much stronger result, namely the asymptotics for the probability density.

Using the Fourier inversion theorem, the probability density function, $p(t)$, of $\ln|Z'(\theta_1)/\exp(C_1)|$ exists and is given by

$$p(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}y(t+C_1)} F(\mathrm{i}y, N)\,\mathrm{d}y \tag{3.35}$$

(that is, for any measurable set A, $\mathrm{P}\{\ln|Z'(\theta_1)| - C_1 \in A\} = \int_A p(t)\,\mathrm{d}t$).

Integrating over the rectangle with vertices $-M$, M, $M+(3+\epsilon)\mathrm{i}$, $-M+(3+\epsilon)\mathrm{i}$ (where ϵ is a fixed number satisfying $0 < \epsilon < 1$) and letting $M \to \infty$, we see that

$$p(t) = \operatorname*{Res}_{y=3\mathrm{i}}\{\mathrm{e}^{-\mathrm{i}y(t+C_1)} F(\mathrm{i}y, N)\} + E, \tag{3.36}$$

where

$$E = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{(-\mathrm{i}y+3+\epsilon)(t+C_1)} F(\mathrm{i}y - 3 - \epsilon, N)\,\mathrm{d}y. \tag{3.37}$$

Asymptotic analysis shows that

$$|E| \leqslant \frac{1}{2\pi} \mathrm{e}^{(3+\epsilon)(t+C_1)} \int_{-\infty}^{\infty} |F(\mathrm{i}y - 3 - \epsilon, N)| \, \mathrm{d}y \tag{3.38}$$

$$\sim \frac{1}{\sqrt{\pi}} \left| \frac{G^2(\frac{1}{2} - \frac{1}{2}\epsilon)}{G(-\epsilon)} \right| \mathrm{e}^{(3+\epsilon)(\gamma-1)} \mathrm{e}^{(3+\epsilon)t} N^{9/4+3\epsilon/2+\epsilon^2/4} (\ln N)^{-1/2}, \tag{3.39}$$

and that

$$\operatorname*{Res}_{y=3\mathrm{i}} \{\mathrm{e}^{-\mathrm{i}y(t+C_1)} F(\mathrm{i}y, N)\} \sim \mathrm{e}^{3t} N^{9/4} \mathrm{e}^{3\gamma-3} G^2(1/2). \tag{3.40}$$

If $\lim_{N\to\infty}(t/\ln N) < -3/2$, then choosing

$$0 < \epsilon < \min\left\{-6 - 4 \limsup_{N\to\infty} \frac{t}{\ln N}, 1\right\} \tag{3.41}$$

shows that the residue gives the dominant contribution to $p(t)$ in this region, that is

$$p(t) \sim \mathrm{e}^{3t} N^{9/4} \mathrm{e}^{3\gamma-3} G^2(1/2) \tag{3.42}$$

if $\limsup_{N\to\infty}(t/\ln N) < -3/2$.

Remark 3.9. The asymptotics for $\ln p(A(N)x)$ completes the proof of the LDP for scaling $A(N) = \ln N$ with $x < -3/2$ (3.26), and for $A \gg \ln N$ with $x < 0$, (3.28).

Remark 3.10. Due to the e^{3t} term in (3.42),

$$\int_{U(N)} |Z'(\theta_1)|^{2k} \, \mathrm{d}\mu_N \tag{3.43}$$

diverges for $k \leqslant -3/2$.

4. Other unitary ensembles

The other unitary ensembles—the circular orthogonal (COE: $\beta = 1$) and circular symplectic (CSE: $\beta = 4$) ensembles—can be dealt with in the same manner as the CUE ($\beta = 2$), the ensemble considered in all of the above.

The normalized measures on these spaces, $U_\beta(N)$, are (Mehta 1991)

$$\mathrm{d}\mu_N^\beta = \frac{((\beta/2)!)^N}{(N\beta/2)!(2\pi)^N} \prod_{1 \leqslant j<k \leqslant N} |\mathrm{e}^{\mathrm{i}\theta_j} - \mathrm{e}^{\mathrm{i}\theta_k}|^\beta \prod_{n=1}^{N} \mathrm{d}\theta_n, \tag{4.1}$$

and Keating & Snaith (2000) found that

$$\int_{U_\beta(N)} |Z(\theta)|^s \, \mathrm{d}\mu_N^\beta = \prod_{j=0}^{N-1} \frac{\Gamma(1 + j\beta/2)\Gamma(1 + s + j\beta/2)}{(\Gamma(1 + s/2 + j\beta/2))^2} \tag{4.2}$$

$$= M_N(\beta, s). \tag{4.3}$$

As in § 2 *a*, we find that

$$\int_{U_\beta(N)} \frac{1}{N} \sum_{n=1}^{N} |Z'(\theta_n)|^s \, \mathrm{d}\mu_N^\beta = \int_{U_\beta(N)} |Z'(\theta_1)|^s \, \mathrm{d}\mu_N^\beta \tag{4.4}$$

$$= (\beta/2)! \frac{((N-1)\beta/2)!}{(N\beta/2)!} M_{N-1}(\beta, s+\beta). \tag{4.5}$$

Calculating the cumulants,

$$C_1^\beta = \sum_{j=0}^{N-2} \Psi(1+\beta+j\beta/2) - \Psi(1+\beta/2+j\beta/2) \tag{4.6}$$

$$= \ln N + O(1), \tag{4.7}$$

$$C_2^\beta = \sum_{j=0}^{N-2} \Psi^{(1)}(1+\beta+j\beta/2) - \tfrac{1}{2}\Psi^{(1)}(1+\beta/2+j\beta/2) \tag{4.8}$$

$$= \frac{1}{\beta} \ln N + O(1), \tag{4.9}$$

$$C_n^\beta = \sum_{j=0}^{N-2} \Psi^{(n-1)}(1+\beta+j\beta/2) - 2^{-(n-1)}\Psi^{(n-1)}(1+\beta/2+j\beta/2) \tag{4.10}$$

$$= O(1), \quad \text{for } n \geqslant 3, \tag{4.11}$$

which shows that

$$\lim_{N\to\infty} \mathrm{P}_\beta \left\{ \frac{\ln |Z'(\theta_1)| - C_1^\beta}{\sqrt{C_2^\beta}} \in (a, b) \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b \mathrm{e}^{-x^2/2} \, \mathrm{d}x. \tag{4.12}$$

We are grateful to Brian Conrey and Steve Gonek for discussions which stimulated part of the research reported here, and to Steve Gonek for making a number of helpful comments on an early draft of the manuscript. C.P.H. and J.P.K. also acknowledge the Mathematical Sciences Research Institute, Berkeley, for support to attend the conference there in June 1999 on the Riemann zeta function and random matrices, where conjecture 2.3 was first announced.

Appendix A. The Barnes *G*-function

Barnes (1900) defined the G-function for all z by

$$G(z+1) = (2\pi)^{z/2} \exp(-\tfrac{1}{2}(z^2 + \gamma z^2 + z)) \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right)^n \mathrm{e}^{-z+z^2/2n}. \tag{A 1}$$

It is an entire function of order two, such that $G(z+1)$ has zeros at $z = -n$ of multiplicity n, where $n = 1, 2, \ldots$.

It has the following properties (Barnes 1900; Voros 1987):

Recurrence relation $G(z+1) = \Gamma(z)G(z)$.

Asymptotic formula for $|z| \to \infty$ with $|\arg(z)| < \pi$:

$$\ln G(z+1) = z^2(\tfrac{1}{2}\ln z - \tfrac{3}{4}) + \tfrac{1}{2}z\ln 2\pi - \tfrac{1}{12}\ln z + \zeta'(-1) + O(1/z). \qquad \text{(A 2)}$$

Taylor expansion for $|z| < 1$:

$$\ln G(z+1) = \tfrac{1}{2}(\ln 2\pi - 1)z - \tfrac{1}{2}(1+\gamma)z^2 + \sum_{n=3}^{\infty}(-1)^{n-1}\zeta(n-1)\frac{z^n}{n}. \qquad \text{(A 3)}$$

Special values $G(1) = 1$ and $G(1/2) = \mathrm{e}^{3\zeta'(-1)/2}\pi^{-1/4}2^{1/24}$.

Logarithmic differentiation

$$\frac{\mathrm{d}^{n+1}}{\mathrm{d}z^{n+1}}\ln G(z) = \Phi^{(n)}(z), \qquad \text{(A 4)}$$

which can be written in terms of the polygamma functions, $\Psi^{(n)}(z)$, with

$$\Phi^{(0)}(z) = \tfrac{1}{2}\ln 2\pi - z + \tfrac{1}{2} + (z-1)\Psi^{(0)}(z), \qquad \text{(A 5)}$$

for example.

References

Barnes, E. W. 1900 The theory of the G-function. *Q. J. Math.* **31**, 264–314.

Berry, M. V. 1988 Semiclassical formula for the number variance of the Riemann zeros. *Nonlinearity* **1**, 399–407.

Berry, M. V. & Keating, J. P. 1999 The Riemann zeros and eigenvalue asymptotics. *SIAM Rev.* **41**, 236–266.

Billingsley, P. 1979 *Probability and measure*. Wiley.

Bogomolny, E. B. & Keating, J. P. 1995 Random matrix theory and the Riemann zeros I: three- and four-point correlations. *Nonlinearity* **8**, 1115–1131.

Bogomolny, E. B. & Keating, J. P. 1996 Random matrix theory and the Riemann zeros II: n-point correlations. *Nonlinearity* **9**, 911–935.

Bucklew, J. A. 1990 *Large deviation techniques in decision, simulation, and estimation*. Wiley.

Coram, M. & Diaconis, P. 1999 New tests of the correspondence between unitary eigenvalues and the zeros of Riemann's zeta function. Preprint.

Gonek, S. M. 1984 Mean values of the Riemann zeta function and its derivatives. *Inventiones Math.* **75**, 123–141.

Gonek, S. M. 1989 On negative moments of the Riemann zeta-function. *Mathematika* **36**, 71–88.

Gonek, S. M. 1999 The second moment of the reciprocal of the Riemann zeta function and its derivative. Talk at Mathematical Sciences Research Institute, Berkeley, June 1999. Available from http://www.msri.org/publications/.

Hejhal, D. A. 1989 On the distribution of $\log|\zeta'(1/2+\mathrm{i}t)|$. In *Number theory, trace formulas, and discrete groups* (ed. K. E. Aubert, E. Bombieri & D. M. Goldfeld). *Proc. 1987 Selberg Symp.*, pp. 343–370.

Hughes, C. P., Keating, J. P. & O'Connell, N. 2000 On the characteristic polynomial of a random unitary matrix. Preprint. BRIMS Technical Report HPL-BRIMS-2000-17.

Keating, J. P. & Snaith, N. C. 2000 Random matrix theory and $\zeta(1/2+\mathrm{i}t)$. *Commun. Math Phys.* **214**, 57–89.

Mehta, M. L. 1991 *Random matrices*, 2nd edn. Academic.

Montgomery, H. L. 1973 The pair correlation of zeros of the zeta function. In *Analytic number theory. Proc. Symposia in Pure Mathematics*, no. 24, pp. 181–193.

Odlyzko, A. M. 1992 The 10^{20}th zero of the Riemann zeta function and 175 million of its neighbors. Preprint. Available from http://www.research.att.com/amo/unpublished/index.html.

Rudnick, Z. & Sarnak, P. 1996 Zeros of principal L-functions and random matrix theory. *Duke Math. J.* **81**, 269–322.

Titchmarsh, E. C. 1986 *The theory of the Riemann zeta-function*, 2nd edn (revised by D. R. Heath-Brown). Oxford: Oxford Science.

Voros, A. 1987 Spectral functions, special functions and the Selberg zeta function. *Commun. Math. Phys.* **110**, 439–465.

Weyl, H. 1946 *The classical groups*. Princeton, NJ: Princeton University Press.