ON THE CHARACTERISTIC POLYNOMIAL OF A RANDOM UNITARY MATRIX AND THE RIEMANN ZETA FUNCTION

Christopher Paul Hughes
School of Mathematics

September 2001

A dissertation submitted to the University of Bristol
in accordance with the requirements of the degree
of Doctor of Philosophy in the Faculty of Science

Abstract

The analogy between the characteristic polynomial of a random unitary matrix and Riemann's zeta function was first studied by Keating and Snaith in [69]. For example, they were able to conjecture the asymptotic form of the moments of the zeta function high up the critical line from a random matrix calculation. The purpose of this thesis is to develop the analogy further. In particular, we present a range of fluctuation and large deviation results for the logarithm of the characteristic polynomial, as the matrix size tends to infinity. These are then compared with the zeta function, and a conjecture for the left large deviations of $\log \zeta|1/2 + \mathrm{i}t|$ is proposed.

This naturally leads to the study discrete moments of the derivative of zeta, evaluated at the non-trivial zeros, whose asymptotics are then conjectured from a random matrix calculation. This conjecture is consistent with the earlier large deviation result, as well as some unconditional theorems.

The Keating-Snaith conjecture and the derivatives conjecture are unified as being special cases of one particular calculation: the discrete moments of the modulus of zeta, evaluated close to its non-trivial zeros. Large gaps between zeros of the zeta function can also be studied using these results.

Finally we use techniques developed in the above work to study the joint moments of the modulus of zeta with its logarithmic derivative.

Acknowledgements

First of all I would like to thank my supervisors, Jon Keating and Neil O'Connell, for their advice and help throughout my time in Bristol. If either of them had not been involved in this research, then my thesis would be very different, and much poorer, as a result.

I would also like to thank Nina Snaith not only for her friendship but also for putting up with all my questions about her work, and for taking the time to answer them. I had many interesting conversations with Kelly Wieand during her stay in Bristol and they led to several improvements in this work, for which I am most grateful. Finally, I would like to thank the various people with whom I have had helpful discussions, including Brian Conrey, Persi Diaconis, Steve Evans, Steve Gonek, Richard Hall and Zeév Rudnick.

I have been kept going these past three years by the encouragement and friendship of many people: most significantly my family, but also members of the maths department, of BRIMS, and of my church. My thanks to them all.

My PhD was financially supported by EPSRC and BRIMS, Hewlett-Packard Laboratories. About half my time was spent at BRIMS, the other half in the maths department.

Author's Declaration

I declare that the work in this thesis was carried out in accordance with the Regulations of the University of Bristol. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. Any views expressed in the dissertation are those of the author and do not necessarily represent those of the University of Bristol. The thesis has not been presented to any other university for examination either in the United Kingdom or overseas.

Christopher Paul Hughes

Date: 11th September 2001

Contents

List of Tables

List of Figures

Chapter 1

Introduction and Summary of Related Work

1.1 The Riemann zeta function

1.1.1 History

The Riemann zeta function is defined for $\mathfrak{Re}(s) > 1$ to be

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s} \tag{1.1}$$

$$= \prod_{\substack{p \\ \text{prime}}} \left(1 - \frac{1}{p^s}\right)^{-1}. \tag{1.2}$$

The second line is called the Euler product of the Riemann[1] zeta function. Immediately we see why the zeta function is important in analytic number theory, for the Euler product explicitly gives a connection between natural numbers and primes. This connection was developed further by Riemann in 1859 when he published an eight page manuscript, [81], on the connection between the asymptotic value of the number of primes less than x and properties of zeros of the zeta function. One of the first things Riemann showed was that $\zeta(s)$ can be meromorphically continued from $\mathfrak{Re}(s) > 1$ into the whole complex plane $\mathbb{C}$. This fact follows from

$$\zeta(s) = \frac{e^{-\mathrm{i}\pi s}\Gamma(1-s)}{2\pi\mathrm{i}} \int_C \frac{(-z)^{s-1}}{e^z - 1}\, \mathrm{d}z \tag{1.3}$$

[1] Though it was discovered by Euler [39] in 1737, 89 years before Riemann was born!

where C is the contour starting at $+\infty$, going parallel and above the positive real axis, encircling the origin once in the positive direction on a circle of radius less than 2π, and going back, below the positive real axis, to $+\infty$. In fact by deforming the contour C Riemann was able to find a functional equation:

$$\zeta(s) = \chi(s)\zeta(1-s) \tag{1.4}$$

where[2]

$$\chi(s) = 2^s \pi^{s-1} \sin(\tfrac{1}{2}\pi s)\Gamma(1-s) \tag{1.5}$$

$$= \pi^{s-1/2} \frac{\Gamma(\frac{1}{2} - \frac{1}{2}s)}{\Gamma(\frac{1}{2}s)} \tag{1.6}$$

From this it follows that $\zeta(s)$ is analytic everywhere, other than for a simple pole at $s = 1$, and that it has simple zeros at the negative even integers. Combining the functional equation with (1.1) means that we effectively know everything about $\zeta(s)$ for $\mathfrak{Re}(s) > 1$ and $\mathfrak{Re}(s) < 0$, and thus it is just the so-called critical strip, $0 \leq \mathfrak{Re}(s) \leq 1$, that remains mysterious. Finally, in his paper, Riemann showed that $\zeta(s)$ vanishes infinitely often in the critical strip (such points being called the non-trivial zeros of the Riemann zeta function) and sketched a proof to show that if no such zero lies on the very edge of the critical strip, (that is, no zero lies on the line $\mathfrak{Re}(s) = 1$), then the prime number theorem is true, which is to say that

$$\pi(x) := \sum_{\substack{p \leq x \\ p \text{ prime}}} 1 \sim \frac{x}{\log x}. \tag{1.7}$$

(Actually Riemann did more than this: he gave an exact formula for $\pi(x)$ in terms of logarithmic integrals, and found that $\mathrm{Li}(x)$ (which is asymptotic to $x/\log x$) is the dominant asymptotic term in the expansion if all the non-trivial zeros lie strictly inside the critical strip). The proof of the prime number theorem was finally completed in 1896 by de la Vallée Poussin [93] and (independently) Hadamard [49]. There is a much fuller account of the history of the Riemann zeta function given in Edwards' book, [38]. A complete account of all the standard results in the theory of the Riemann zeta function can be found in [90].

[2] I have translated Riemann's equations into standard modern mathematical notation. In particular, Riemann had a different notation for the Euler Gamma function. Riemann's notation is preserved in [38].

From the functional equation, (1.4), it is clear that

$$\mathcal{Z}(t) := \sqrt{\chi(\tfrac{1}{2} - \mathrm{i}t)}\zeta(\tfrac{1}{2} + \mathrm{i}t) \tag{1.8}$$

is real for real t with the useful property that $|\zeta(1/2 + \mathrm{i}t)| = |\mathcal{Z}(t)|$. Note that from (1.6)

$$\sqrt{\chi(\tfrac{1}{2} - \mathrm{i}t)} = e^{\mathrm{i}\vartheta(t)} \tag{1.9}$$

where

$$\vartheta(t) = \mathfrak{Im} \log \left\{ \pi^{-\mathrm{i}t/2} \Gamma(\tfrac{1}{4} + \tfrac{1}{2}\mathrm{i}t) \right\} \tag{1.10}$$

is taken to be a real continuous function with the initial value $\vartheta(0) = 0$.

1.1.2 Zeros of zeta

It is known that $\zeta(s)$ has a countably infinite number of zeros (called the Riemann zeros or non-trivial zeros) in the critical strip, and that none of them are real (so the only real zeros of the zeta function are at $-2n$, n a positive integer). If we denote the non-trivial zeros by $\rho_n = \beta_n + \mathrm{i}\gamma_n$ with $\gamma_{-1} < 0 < \gamma_1 \leq \gamma_2 \leq \ldots$, then what Hadamard and de la Vallée Poussin did in order to complete the proof of the prime number theorem was to show that $0 < \beta_n < 1$ for all n. Let $N(T)$ be the number of zeros up to height T, counted according to multiplicity, that is if there is a double zero at z, say, then $\rho_n = \rho_{n+1} = z$ for some n. It is known [73] that $N(T) = \overline{N}(T) + S(T)$ with

$$\overline{N}(T) := 1 + \frac{1}{\pi}\vartheta(T) \tag{1.11}$$

$$= \frac{T}{2\pi} \log \frac{T}{2\pi e} + \frac{7}{8} + \mathcal{O}\left(\frac{1}{T}\right) \tag{1.12}$$

being the smoothed mean counting function ($\vartheta(T)$ being defined in (1.10)), and where

$$S(T) := \frac{1}{\pi}\mathfrak{Im} \log \zeta\left(\tfrac{1}{2} + \mathrm{i}T\right) \tag{1.13}$$

$$= \mathcal{O}(\log T) \tag{1.14}$$

describes the fluctuations around the mean. The imaginary part of the logarithm is defined so that $\mathfrak{Im} \log \zeta(s)$ varies continuously along the straight lines joining 2 to $2 + \mathrm{i}T$, and $2 + \mathrm{i}T$ to $1/2 + \mathrm{i}T$, with the initial value of 0 at $s = 2$. When $1/2 + \mathrm{i}T$ is

a Riemann zero $S(T)$ has a jump discontinuity. This asymptotic formula for $N(T)$ can again be found in Riemann's paper, [81], but it is not rigorously proved there. Such a proof was first given by von Mangoldt, [73].

Due to the functional equation, if $\beta + \mathrm{i}\gamma$ is a zero, then so is $1 - \beta - \mathrm{i}\gamma$. And since $\zeta(s)^* = \zeta(s^*)$, then $\beta - \mathrm{i}\gamma$ and $1 - \beta + \mathrm{i}\gamma$ must also be zeros of $\zeta(s)$. That is, the zeros are located so as to be symmetric about the real axis, and the line $\mathfrak{Re}(s) = \frac{1}{2}$ (the critical line). In [81], Riemann says that it is very probable ("sehr wahrscheinlich") that all the zeros lie on the symmetry line, that $\beta_n = \frac{1}{2}$ for all n. This is the Riemann Hypothesis (RH). He goes on to say

> "Hiervon wäre allerdings ein strenger Beweis zu wünschen; ich habe indess die Aufsuchung desselben nach einigen flüchtigen vergeblichen Versuchen vorläufig bei Seite gelassen, da er für den nächsten Zweck meiner Untersuchung entbehrlich schien."[3]

As we have seen, the first proofs of the prime number theorem depended on the fact the no zero lay on the line $\mathfrak{Re}(s) = 1$. It is pleasing to note the connection is deeper, for the horizontal distribution of zeros and the error term in the prime number theorem are closely related: The statement

$$\pi(x) = \mathrm{Li}(x) + \mathcal{O}(\sqrt{x}\log x) \tag{1.15}$$

is equivalent to RH. In fact, it is true that

$$\pi(x) = \mathrm{Li}(x) + \mathcal{O}(x^{\Theta}\log x) \tag{1.16}$$

where

$$\Theta = \sup_{\zeta(\rho)=0} \mathfrak{Re}(\rho) = \sup_n \beta_n \tag{1.17}$$

(so that, unconditionally, $\frac{1}{2} \leq \Theta \leq 1$, and that $\Theta = \frac{1}{2}$ is the Riemann hypothesis).

Although no valid proof of the Riemann hypothesis has been found, there is extensive evidence for its truth—for example, the first $1.5 \times 10^9 + 1$ zeros lie on the critical line, and are simple [72], as do 175 million zeros around zero number 10^{20} [78]. Also, at least 40% of all Riemann zeros lie on the line and are simple [23].

[3] In English: "Though one would wish for a strict proof here; I have meanwhile temporarily put aside the search for this after some fleeting futile attempts, as it appears unnecessary for the next objective of my investigation." [Which was the proof of the prime number theorem].

Many of the theorems in analytic number theory are conditional on the truth of RH, and its manifold ramifications make it perhaps the most important unresolved question in modern mathematics[4].

It is possible to ask much deeper questions about the distribution of zeros, beyond their horizontal distribution (which is obviously completely solved if one assumes RH!). Montgomery [76] in 1973 was one of the first people to consider vertical spacings for the zeros of the zeta function. He considered

$$F(\alpha, T) = \frac{1}{N(T)} \sum_{0<\gamma,\gamma'\leq T} T^{\mathrm{i}\alpha(\gamma-\gamma')} w(\gamma - \gamma') \tag{1.18}$$

where $w(u) = 4/(4+u^2)$ is a weighting function, α is real. In order for this to make sense, $N(T)$ must be nonzero[5], which requires $T > \gamma_1 \approx 14.13$. Montgomery proved:

Theorem 1.1. *Assume RH. Then for* $0 \leq \alpha < 1$,

$$F(\alpha, T) = \alpha + (1 + o(1)) \frac{\log T}{T^{2\alpha}} + o(1) \tag{1.19}$$

as $T \to \infty$, *uniformly for* $0 \leq \alpha < 1$.

Furthermore he conjectured that if the primes are distributed sufficiently uniformly in arithmetic progressions, then

Conjecture 1.2. *For* $\alpha \geq 1$

$$F(\alpha, T) = 1 + o(1) \tag{1.20}$$

as $T \to \infty$, *uniformly in* α. *Using the fact that* $F(\alpha, T)$ *is an even function of* α, *Montgomery's conjecture and theorem together state*

$$\lim_{T\to\infty} F(\alpha, T) = \begin{cases} |\alpha| + \delta(\alpha) & \text{for } |\alpha| \leq 1 \\ 1 & \text{for } |\alpha| \geq 1 \end{cases} \tag{1.21}$$

where $\delta(\alpha)$ *is the Dirac delta function.*

[4] so much so that the Clay Mathematical Institute has made the proof of RH and its generalization to other L–functions one of the problems of the millennium, with a \$1,000,000 prize for a valid proof of it.

[5] In [76], $N(T)$ is replaced by $\frac{T}{2\pi} \log T$, but this makes no difference to the large T asymptotics.

Assuming this conjecture, then Montgomery showed, by taking the Fourier transform of $F(\alpha, T)$, that for $\alpha < \beta$,

$$\lim_{T\to\infty} \frac{1}{N(T)} \# \left\{ 0 < \gamma, \gamma' \leq T \,:\, \alpha \leq (\gamma - \gamma') \frac{1}{2\pi} \log \frac{T}{2\pi} \leq \beta \right\} = \int_\alpha^\beta 1 - \left(\frac{\sin(\pi u)}{\pi u} \right)^2 \, \mathrm{d}u + \delta(\alpha, \beta) \quad (1.22)$$

where $\delta(\alpha, \beta) = 1$ if $0 \in [\alpha, \beta]$ and zero otherwise. This term reflects the fact that if $\alpha \leq 0 \leq \beta$, then the sum includes the terms $\gamma = \gamma'$.

From (1.12) it is clear that high up the critical line the zeros around height T "bunch up" with a density asymptotic to $\frac{1}{2\pi} \log \frac{T}{2\pi}$. And so scaling the zeros by this density, that is mapping γ_n to $\frac{\gamma_n}{2\pi} \log \frac{\gamma_n}{2\pi}$, makes the average distance between them tend to one. Thus Montgomery's conjecture essentially says that when the high non-trivial zeros are scaled so that they have mean spacing equal to one, the expected[6] number of scaled zeros less than x away from another scaled zero is

$$\int_{-x}^{x} 1 - \left(\frac{\sin(\pi u)}{\pi u} \right)^2 \mathrm{d}u + 1 \quad (1.23)$$

We will return to Montgomery's conjecture later in the chapter.

1.1.3 Moments of $|\zeta(1/2 + \mathrm{i}t)|$

The Lindelöf hypothesis is the conjecture that

$$\zeta(\tfrac{1}{2} + \mathrm{i}t) = \mathcal{O}(t^\epsilon) \quad (1.24)$$

for every positive ϵ (which, incidently, is equivalent to $\zeta(\sigma + \mathrm{i}t) = \mathcal{O}(t^\epsilon)$ for all $\sigma \geq \frac{1}{2}$).

The $2k^{\text{th}}$ moment (or mean-value) of the modulus of the Riemann zeta function,

$$I_k = \frac{1}{T} \int_0^T \left| \zeta\left(\tfrac{1}{2} + \mathrm{i}t\right) \right|^{2k} \, \mathrm{d}t, \quad (1.25)$$

were originally studied in an attempt to prove the Lindelöf hypothesis, which is equivalent to $I_k = \mathcal{O}(T^\epsilon)$ for all k (see theorem 13.2 of [90], for example). Later on, deeper results on I_k were used as an aid to understanding large values of the zeta function on the critical line (this problem is discussed in section 3.6.1).

[6] By expectation we mean sum over many scaled zeros around some large height T, and divide by the number zeros summed over. Due to the scaling, the answer should be insensitive to the height T.

Some of the known results concerning I_k are as follows: Hardy and Littlewood [52] proved that

$$I_1(T) \sim \log T \text{ as } T \to \infty, \tag{1.26}$$

and Ingham [61] proved

$$I_2(T) \sim \frac{1}{2\pi^2}(\log T)^4. \tag{1.27}$$

Conrey and Ghosh [28] have conjectured that

$$I_3(T) \sim \frac{42}{9!}a(3)(\log T)^9, \tag{1.28}$$

(where $a(k)$ is given below) and Conrey and Gonek [30][7] that

$$I_4(T) \sim \frac{24024}{16!}a(4)(\log T)^{16}. \tag{1.29}$$

The Keating-Snaith conjecture comes from a random matrix calculation, and covers all fixed k with $\mathfrak{Re}(k) > -1/2$. They conjecture

$$I_k(T) \sim \frac{G^2(k+1)}{G(2k+1)}a(k)(\log T)^{k^2} \tag{1.30}$$

where

$$a(k) = \prod_{\substack{p \\ \text{prime}}} \left(1 - \frac{1}{p}\right)^{k^2} \sum_{m=0}^{\infty} \left(\frac{\Gamma(m+k)}{m!\,\Gamma(k)}\right)^2 p^{-m}. \tag{1.31}$$

The evidence for this conjecture (including the random matrix calculations which lead to it) will be dealt with in more depth in section 1.3.1.

In the cases $k = 1, 2$ more is known, beyond the leading order term of the asymptotic expansion.

$$I_1(T) = \log \frac{T}{2\pi} + 2\gamma - 1 + \mathcal{O}(T^{-15/22}(\log T)^{111/22}) \tag{1.32}$$

was proved by Ingham [61] with the error term due to Heath-Brown and Huxley [55]. Heath-Brown [53] has found

$$I_2(T) = P_4\left(\log \frac{T}{2\pi}\right) + \mathcal{O}(T^{-1/8+\epsilon}) \tag{1.33}$$

where

$$P_4(x) = \sum_{n=0}^{4} c_n x^n \tag{1.34}$$

[7] Although [30] was published in 2001, this conjecture was first announced in 1998. In fact, at the same conference the Keating-Snaith conjecture (conjecture 1.4 in this thesis) was also announced.

with $c_4 = \frac{1}{2\pi^2}$ and $c_3 = \frac{2}{\pi^2}\left(4\gamma - 1 - \frac{12}{\pi^2}\zeta'(2)\right)$. Conrey [24] improved this to

$$P_4(x) = g_0(x) + g_1(x) \tag{1.35}$$

with

$$g_0(x) = \operatorname*{Res}_{s=0} \frac{2e^{xs}\zeta(s+1)^4}{s(s+1)\zeta(2s+2)} \tag{1.36}$$

and

$$g_1(x) = \frac{\mathrm{d}^2}{\mathrm{d}s^2} \left. \frac{e^{xs}e^{2s\gamma}\left(s\zeta(s+1)^2 - 2\zeta(2s+1) - 2s\zeta(2s+2)\right)}{2s(s+1)\zeta(s+2)} \right|_{s=0} \tag{1.37}$$

where $\gamma = 0.5772\ldots$ is Euler's constant.

Apart from $k = 0, 1, 2$, none of the asymptotic formulas for $I_k(T)$ have been proven. However, some bounds on $I_k(T)$ have been. For example, Heath-Brown [54] has shown that, for certain k, $T \geq 2$

$$A(k)(\log T)^{k^2} \leq I_k(T) \leq B(k)(\log T)^{k^2} \tag{1.38}$$

where the left hand side holds for all rational $k \geq 0$, or (under RH) for all real $k \geq 0$ and the right hand side holds for all $k = 1/n$, n an integer, or (under RH) for $0 \leq k \leq 2$. (So we have $I_k(T) \asymp (\log T)^{k^2}$ for $k = 1/n$, n an integer, or (under RH) for $0 \leq k \leq 2$). Jutila [65] has shown that the constants $A(k)$ and $B(k)$ can be made independent of k for $k = 1/n$.

Under the Riemann Hypothesis, Conrey and Ghosh [26] have shown that, for any fixed $k \geq 0$,

$$I_k(T) \geq \left(\frac{a(k)}{\Gamma(1+k^2)} + o(1)\right)(\log T)^{k^2}. \tag{1.39}$$

This has been improved for integer k by Balasubramanian and Ramachandra [4], who show (independent of any unproved hypothesis) that for any positive integer k, for $\log\log T << H \leq T$,

$$\frac{1}{H}\int_T^{T+H} \left|\zeta(\tfrac{1}{2} + \mathrm{i}t)\right|^{2k} \mathrm{d}t \geq \left(\frac{a(k)}{\Gamma(1+k^2)} + o(1)\right)(\log H)^{k^2}. \tag{1.40}$$

Putting $H = T$, Conrey and Ghosh's result is recovered, without requiring RH. But note that for $H \ll T$ the order of this lower bound, $(\log H)^{k^2}$, is much less than the anticipated true order, $(\log T)^{k^2}$.

1.2 Random matrix theory

Random matrix theory (RMT) is essentially the probabilistic study of various ensembles of matrices. (An ensemble is a set with an attached probability measure). See [74, 91] for a review of RMT.

Mathematics	Physics
multivariate analysis (1940s)	nuclear physics (1950s)
spectral theory (1960s)	quantum chaology (1970s)
orthogonal polynomials (1960s)	large-colour QCD (1970s)
zeros of zeta and L–functions (1970s)	statistical mechanics of random surfaces (1980s)
combinatorics (1980s)	mesoscopics (1980s)
C^*-algebras and non-commutative probability (1980s)	infinite dimensional integrable systems (1980s)
geometry of Banach spaces (1990s)	2D quantum gravity (1990s)
large-n representation theory (1990s)	semiclassical methods (1990s)

Table 1.1: Random matrix theory in mathematics and physics

Table 1.1[8] shows some of the various areas of applicability of (different ensembles of) random matrices. See also [79], and its 206 references, for a fuller description of random matrices and their applications. In this thesis we will only be concerned with one particular ensemble (the CUE), and one particular application (the Riemann zeta function). But for completeness we will here define the various circular and Gaussian ensembles, as well as the ensembles for the classical compact groups. We will define them in terms of the invariances they possess and state the eigenangle joint probability density such symmetries induce. Further results for the CUE are described in the next subsection.

[8] copied from a talk given by L. Pastur

Each of the classical compact groups we will deal with is a compact Lie group. Therefore it has a unique left- (and right-)invariant probability measure, called Haar measure. The unitary, orthogonal and symplectic group ensemble is the group attached with its respective Haar measure. The calculation of the joint probability density of the eigenangles under Haar measure is due to Weyl [95].

Unitary. The unitary group, $\mathcal{U}(N)$, is the group of all $N \times N$ complex matrices U which satisfy the condition $UU^\dagger = I_N$, where $U^\dagger$ denotes the complex transpose of U, and I_N is the $N \times N$ identity matrix. The joint probability density of eigenangles is

$$\frac{1}{N!(2\pi)^N} \prod_{1 \le j < k \le N} \left| e^{\mathrm{i}\theta_j} - e^{\mathrm{i}\theta_k} \right|^2 \prod_{n=1}^{N} \mathrm{d}\theta_n \tag{1.41}$$

Orthogonal. The orthogonal group, $O(N)$, is the group of all $N \times N$ real matrices O which satisfy the condition $OO^t = I_N$, where O^t denotes transpose of O. The special orthogonal group, $SO(N)$, is the group of all orthogonal matrices with determinant $+1$. Specializing to $SO(2N)$, then the eigenangles are necessarily paired, in the sense that if $e^{\mathrm{i}\theta_1}$ is an eigenvalue, then so is $e^{-\mathrm{i}\theta_1}$. Thus the joint probability density of eigenangles of $SO(2N)$ under Haar measure is given in terms of $\theta_1, \ldots, \theta_N$, which are restricted to lie between 0 and π, and it equals

$$\frac{2^{N^2-2N+1}}{\pi^N N!} \prod_{1 \le j < k \le N} (\cos\theta_j - \cos\theta_k)^2 \prod_{n=1}^{N} \mathrm{d}\theta_n \tag{1.42}$$

Symplectic. The symplectic group, $Sp(N)$, is the group of all $2N \times 2N$ unitary matrices S with satisfy the condition $SJS^t = J$ where

$$J = \begin{pmatrix} 0 & I_N \\ -I_N & 0 \end{pmatrix} \tag{1.43}$$

Again, the eigenangles are paired, and the joint probability density is (in terms of $\theta_1, \ldots, \theta_N$, which are restricted to lie between 0 and π)

$$\frac{2^{N^2}}{\pi^N N!} \prod_{1 \le j < k \le N} (\cos(\theta_j) - \cos(\theta_k))^2 \prod_{n=1}^{N} \sin^2\theta_n \, \mathrm{d}\theta_n \tag{1.44}$$

Other measures can be attached to $\mathcal{U}(N)$. For example, we could consider the set of unitary matrices with probability measure invariant under orthogonal transformations (that is $\mathcal{U}(N)/O(N)$) (this will be called the circular orthogonal ensemble,

or COE) or the set of all unitary matrices with probability measure invariant under symplectic transformations (that is $\mathcal{U}(2N)/Sp(N)$) (this will be called the circular symplectic ensemble or CSE). Each eigenvalue of a CSE matrix has even multiplicity.

The joint probability density of eigenangles of the circular ensembles is

$$\frac{\left((\frac{1}{2}\beta)!\right)^N}{(\frac{1}{2}N\beta)!(2\pi)^N} \prod_{1\leq j<k\leq N} \left|e^{\mathrm{i}\theta_j} - e^{\mathrm{i}\theta_k}\right|^\beta \prod_{n=1}^N \mathrm{d}\theta_n \tag{1.45}$$

where $\beta = 1$ for COE and $\beta = 4$ for CSE. The case $\beta = 2$ is the circular unitary ensemble (CUE), which is also the unitary group ensemble. The circular ensembles were first studied by Dyson in a series of papers, [37].

Another collection of ensembles is the Gaussian ensembles. These were introduced by Wigner[9] to study the energy-level statistics of large nuclei (see [74], for example).

Gaussian orthogonal ensemble (GOE). The set of all $N \times N$ real symmetric matrices, with attached probability measure that is invariant under $H \mapsto O^{-1}HO$, with the matrix elements on and above the diagonal being iid random variables.

Gaussian unitary ensemble (GUE). The set of all $N \times N$ hermitian matrices with attached probability measure invariant under $H \mapsto U^{-1}HU$, with the real and imaginary matrix elements on and above the diagonal being iid random variables.

Gaussian symplectic ensemble (GSE). The set of all $2N \times 2N$ self-dual hermitian[10] matrices with the attached probability measure invariant under $H \mapsto S^{-1}HS$, with the real and imaginary matrix elements on and above the diagonal being iid random variables.

The joint probability density of the eigenvalues (which are real numbers, due to the matrices being hermitian) is proportional to

$$e^{-\frac{1}{2}\beta\sum_{n=1}^N x_n^2} \prod_{j<k} |x_j - x_k|^\beta \tag{1.46}$$

where $\beta = 1$ is GOE, $\beta = 2$ is GUE and $\beta = 4$ is GSE. They are called Gaussian ensembles due to the $\exp(-\frac{1}{2}\beta\sum_{n=1}^N x_n^2)$ factor in front of the Vandermonde determinant.

[9] It is Wigner who is credited with starting the study of random matrix theory as it is known today.

[10] A $2N \times 2N$ hermitian matrix is self-dual if the equivalent $N \times N$ quaternion matrix has elements which satisfy $\overline{M_{ji}} = M_{ij}$, where the bar denotes quaternion conjugation. Each eigenvalue of a self-dual hermitian matrix appears with even multiplicity.

1.2.1 Circular unitary ensemble

The ensemble we use throughout almost all of this report is the circular unitary ensemble (CUE). We discuss here some of its features in more depth. Much of this summary can be found in [91] or [74].

Denote the eigenvalues of $U \in \mathcal{U}(N)$ by $\exp(\mathrm{i}\theta_1), \ldots, \exp(\mathrm{i}\theta_N)$. As has already been mentioned, the joint probability density of eigenangles is

$$P_N(\theta_1, \ldots, \theta_N) = \frac{1}{N!(2\pi)^N} \prod_{1 \le j < k \le N} \left| e^{\mathrm{i}\theta_j} - e^{\mathrm{i}\theta_k} \right|^2 \tag{1.47}$$

$$= \frac{1}{N!} \det \{K_N(\theta_j - \theta_k)\}_{1 \le j,k \le N} \tag{1.48}$$

where

$$K_N(\theta) = \frac{1}{2\pi} \frac{\sin(N\theta/2)}{\sin(\theta/2)} \tag{1.49}$$

(this follows from rearranging the Vandermonde matrices). Almost all the random matrix calculations in this thesis will be integrating functions of eigenangles against this measure.

Define

$$R_n^{(N)}(\theta_1, \ldots, \theta_n) = \frac{N!}{(N-n)!} \int \cdots \int_{-\pi}^{\pi} P_N(\theta_1, \ldots, \theta_N) \, \mathrm{d}\theta_{n+1} \ldots \mathrm{d}\theta_N \tag{1.50}$$

which is like the probability density of finding eigenangles (regardless of labelling) at each of the angles $\theta_1, \theta_2, \ldots, \theta_n$, ignoring the position of the remaining $N - n$ eigenangles. (It is not actually a probability density, due to the lack of normalization. Integrating $R_n^{(N)}$ over all arguments gives the number of n-tuples of eigenvalues regardless of labelling, that is, $N!/(N-n)!$.) It is shown in [74] that

$$R_n^{(N)}(\theta_1, \ldots, \theta_n) = \det \{K_N(\theta_j - \theta_k)\}_{1 \le j,k \le n} \tag{1.51}$$

$R_1^{(N)}(\theta_1) = \frac{N}{2\pi}$ is just the density of eigenvalues on the unit circle. Note it is independent of position, θ_1. Therefore to get a non-trivial limit as $N \to \infty$ the eigenangles must be scaled by their density. Writing $x_j = \frac{N}{2\pi}\theta_j$, then

$$R_n(x_1, \ldots, x_n) = \lim_{N \to \infty} R_n^{(N)} \left(\frac{2\pi}{N} x_1, \ldots, \frac{2\pi}{N} x_n \right) \tag{1.52}$$

$$= \det \{K(x_j - x_k)\}_{1 \le j,k \le n} \tag{1.53}$$

where

$$K(x) = \frac{\sin(\pi x)}{\pi x} \tag{1.54}$$

In particular, for $\alpha < \beta$

$$\lim_{N\to\infty} \mathbb{E}_N \frac{1}{N} \# \left\{ \theta_m, \theta_n \ : \ \alpha \leq (\theta_m - \theta_n)\frac{N}{2\pi} \leq \beta \right\} = \int_\alpha^\beta R_2(u,0)\, \mathrm{d}u$$
$$= \int_\alpha^\beta 1 - \left(\frac{\sin(\pi u)}{\pi u}\right)^2 \mathrm{d}u + \delta(\alpha,\beta) \tag{1.55}$$

where $\mathbb{E}_N$ denotes expectation with respect to the Haar measure on $\mathcal{U}(N)$, and where $\delta(\alpha,\beta) = 1$ if $\alpha \leq 0 \leq \beta$ and equals 0 otherwise. $R_2(u,0)$ is called the two-point correlation function. (In fact, for almost all matrices, (1.55) holds without requiring the average over $\mathcal{U}(N)$ to be taken—an example of ergodicity).

The final piece of notation to be introduced here is $E(n;s)$, the limiting probability of finding exactly n scaled eigenangles in the interval $[0,s]$. (It is simple to generalize this to subsets $I \subset \mathbb{R}$, but that is unnecessary here). In particular, $p(s) := \frac{\mathrm{d}^2}{\mathrm{d}s^2} E(0;s)$ is the limiting probability density of the gap between consecutive scaled eigenangles. It is known [62] (see [91] for a simpler proof) that

$$E(0;s) = \exp\left(\int_0^{\pi s} \frac{\sigma(u)}{u} \mathrm{d}u\right) \tag{1.56}$$

where $\sigma(u)$ is of Painlevé V type, which means it satisfies

$$\left(u\sigma''\right)^2 + 4\left(u\sigma' - \sigma\right)\left(u\sigma' - \sigma + (\sigma')^2\right) = 0 \tag{1.57}$$

$$\sigma(u) \sim -\tfrac{1}{\pi}u - \tfrac{1}{\pi^2}u^2 \quad \text{as } u \to 0 \tag{1.58}$$

and [40]

$$p(s) = -\frac{\tilde{\sigma}(\pi s)}{s} \exp\left(\int_0^{\pi s} \frac{\tilde{\sigma}(u)}{u} \mathrm{d}u\right) \tag{1.59}$$

where

$$\left(u\tilde{\sigma}''\right)^2 + 4\left(u\tilde{\sigma}' - \tilde{\sigma}\right)\left(u\tilde{\sigma}' - \tilde{\sigma} + (\tilde{\sigma}')^2\right) - 4\left(\tilde{\sigma}'\right)^2 = 0 \tag{1.60}$$

$$\tilde{\sigma}(u) \sim -\tfrac{1}{3\pi}u^3 \quad \text{as } u \to 0 \tag{1.61}$$

As $N \to \infty$ the spacing distribution and k-point correlation functions in the GUE are found to be identical to those in the CUE. This is an example of universality.

1.3 RMT and $\zeta(1/2 + \mathrm{i}t)$

Comparing (1.55) with (1.22) we see that (conjecturally) the two-point correlation function for the non-trivial zeros is asymptotically the same as the two-point correlation function for eigenangles of Haar-distributed unitary matrices. This conjecture can be derived using heuristic methods, see [67, 76], for example. Furthermore, 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 has been proven in restricted ranges [57, 82] (the restriction being essentially the same as the restriction $|\alpha| \leq 1$ in theorem 1.1), and heuristically calculated, [13, 14]. Thus it appears that at the local level, the correlations between the zeros of the zeta function equal the correlations between eigenangles of unitary matrices. This theoretical work has been verified by extensive numerical evidence [78].

1.3.1 The characteristic polynomial

In order to develop a heuristic understanding of the value distribution and moments of the Riemann zeta function on the critical line, Keating and Snaith [69, 86] considered the characteristic polynomial of an $N \times N$ unitary matrix,

$$Z_U(\theta) = \det(I - Ue^{-\mathrm{i}\theta}) \tag{1.62}$$

$$= \prod_{n=1}^{N} \left(1 - e^{\mathrm{i}(\theta_n - \theta)}\right). \tag{1.63}$$

Central limit theorems

Note that the law of $Z_U(\theta)$ over the CUE is independent of $\theta \in \mathbb{T}$ (the unit circle). In [69] it is shown that:

Theorem 1.3.

$$\frac{\log Z_U(0)}{\sigma} \Longrightarrow X + \mathrm{i}Y, \tag{1.64}$$

where X, Y are iid standard normal random variables, and

$$\sigma = \sqrt{\tfrac{1}{2}\log N}. \tag{1.65}$$

(Baker and Forrester [2] had previously shown that the real and imaginary parts of $\log Z_U(\theta)/\sigma$ separately converge in distribution to a standard normal random

variable using a method similar to that employed by Keating and Snaith to show convergence of the joint distribution. The central limit theorem for the imaginary part of the log can be deduced from results on the counting function, due to Wieand [96, 97] and, for the GUE, to Costin and Lebowitz [33].) In fact, in [69] the variance of the real and imaginary parts of $\log Z_U(0)$ is calculated exactly at finite N:

$$\mathbb{E}\left\{\mathfrak{Re}\log Z_U(0)^2\right\} = \mathbb{E}\left\{\mathfrak{Im}\log Z_U(0)^2\right\} \tag{1.66}$$

$$= \tfrac{1}{2}(\Psi(N)+\gamma+N\Psi^{(1)}(N)) \tag{1.67}$$

$$\sim \tfrac{1}{2}\log N = \sigma^2. \tag{1.68}$$

In order to make the imaginary part of the logarithm well-defined, the branch is chosen so that

$$\log Z_U(\theta) = \sum_{n=1}^{N} \log\left(1-e^{\mathrm{i}(\theta_n-\theta)}\right) \tag{1.69}$$

and

$$-\tfrac{1}{2}\pi < \mathfrak{Im}\log\left(1-e^{\mathrm{i}(\theta_n-\theta)}\right) \leq \tfrac{1}{2}\pi. \tag{1.70}$$

Compare the above central limit theorem with a central limit theorem, due to Selberg, for the value distribution of the log of the Riemann zeta function along the critical line. Selberg proved (see, for example, §4 of [71]) that, for rectangles $B \subseteq \mathbb{C}$,

$$\lim_{T\to\infty}\frac{1}{T}\operatorname{meas}\left\{T\leq t\leq 2T \;:\; \frac{\log\zeta(\frac{1}{2}+\mathrm{i}t)}{\sqrt{\frac{1}{2}\log\log T}}\in B\right\} = \frac{1}{2\pi}\iint_B e^{-(x^2+y^2)/2}\,\mathrm{d}x\,\mathrm{d}y. \tag{1.71}$$

In order to compare these two central limit theorems, a connection between N and T must be established. Note that to get a $T\to\infty$ limit in (1.22), the zeros are scaled by their mean density at height T, which is $\frac{1}{2\pi}\log\frac{T}{2\pi}$. Similarly, in (1.55), the scaling was the mean density of eigenangles of an $N\times N$ unitary matrix, which is $\frac{N}{2\pi}$. These are the only scaling parameters, so it makes sense to set them equal to each other, which gives

$$N = \log\frac{T}{2\pi}. \tag{1.72}$$

With this connection between N and T, the two central limit theorems are consistent.

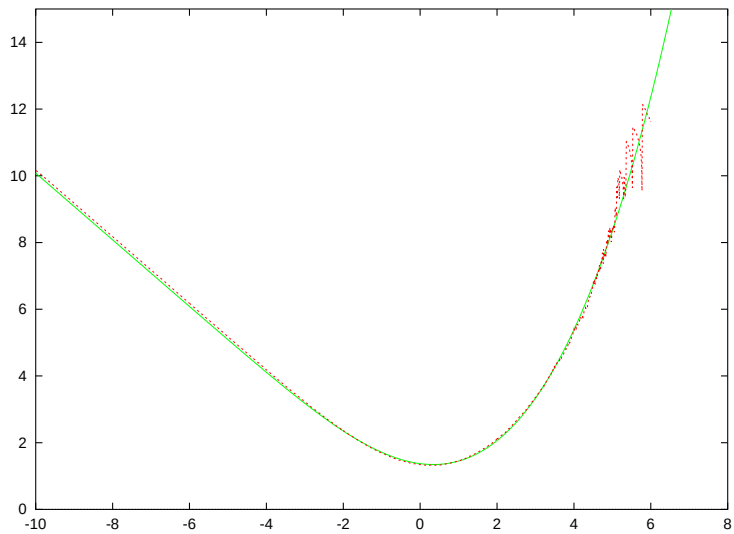

Figure 1.1: Graph of the negative log of the value distribution for the Riemann Zeta function (from A. Odlyzko) around the 10^{20}th zero (red), against the negative log of the probability density of X_N (from N. Snaith) with $N = 42$ (green).

Coram and Diaconis [31] have subsequently shown that making the identification (1.72) leads to close agreement with empirical statistics other than the characteristic polynomial.

The Keating-Snaith conjecture

In order the prove the central limit theorem, Keating and Snaith evaluated

$$\mathbb{E}\exp\left(s\mathfrak{Re}\log Z_U(\theta) + t\mathfrak{Im}\log Z_U(\theta)\right) = \frac{G(1+\frac{1}{2}s+\frac{1}{2}\mathrm{i}t)G(1+\frac{1}{2}s-\frac{1}{2}\mathrm{i}t)G(1+N)G(1+N+s)}{G(1+N+\frac{1}{2}s+\frac{1}{2}\mathrm{i}t)G(1+N+\frac{1}{2}s-\frac{1}{2}\mathrm{i}t)G(1+s)} \tag{1.73}$$

valid for $\mathfrak{Re}(s \pm \mathrm{i}t) > -1$, where $G(\cdot)$ is the Barnes G–function, described in appendix A, and $\mathbb{E}$ denotes expectation over the CUE. When N is large, the asymptotics for the G–function imply that

$$\mathbb{E}\exp\left(s\mathfrak{Re}\log Z_U(\theta) + t\mathfrak{Im}\log Z_U(\theta)\right) \sim \frac{G(1+\frac{1}{2}s+\frac{1}{2}\mathrm{i}t)G(1+\frac{1}{2}s-\frac{1}{2}\mathrm{i}t)}{G(1+s)} N^{s^2/4+t^2/4} \tag{1.74}$$

uniformly for bounded s, t.

Baker and Forrester [2] have found similar central limit theorems (they consider the real and imaginary parts of $\log Z_U(0)$ separately), by calculating the single

moment generating functions, which we record here as

$$M_N(s) := \mathbb{E}\exp(s\mathfrak{Re}\log Z_U(0)) \tag{1.75}$$

$$= \prod_{j=1}^{N} \frac{\Gamma(j)\Gamma(j+s)}{\Gamma^2(j+\frac{1}{2}s)} \tag{1.76}$$

$$= \frac{G^2(1+\frac{1}{2}s)G(N+1)G(N+1+s)}{G(1+s)G^2(N+1+\frac{1}{2}s)}, \tag{1.77}$$

and

$$L_N(\mathrm{i}t) := \mathbb{E}\exp(t\mathfrak{Im}\log Z_U(0)) \tag{1.78}$$

$$= \prod_{j=1}^{N} \frac{\Gamma^2(j)}{\Gamma(j+\frac{1}{2}\mathrm{i}t)\Gamma(j-\frac{1}{2}\mathrm{i}t)} \tag{1.79}$$

$$= \frac{G(1+\frac{1}{2}\mathrm{i}t)G(1-\frac{1}{2}\mathrm{i}t)G^2(N+1)}{G(N+1+\frac{1}{2}\mathrm{i}t)G(N+1-\frac{1}{2}\mathrm{i}t)}. \tag{1.80}$$

Note that $M_N(2k)$ is the $2k^{\text{th}}$ moment of $|Z_U(0)|$, and that as $N \to \infty$

$$M_N(2k) = \frac{G^2(k+1)}{G(2k+1)}N^{k^2} + \mathcal{O}(N^{k^2-1}). \tag{1.81}$$

Brézin and Hikami [19] have shown that $\mathbb{E}\,|Z_M(0)|^{2k}$ (where M is taken from other $\beta = 2$ ensembles, like the GUE) have the same leading order asymptotics as $M_N(2k)$.

Given the success of $\log Z_U(0)$ in modelling $\log\zeta(1/2+\mathrm{i}t)$, it is natural to ask whether random matrix theory can model the moments of $|\zeta(1/2+\mathrm{i}t)|$. However, the naive approach fails: it is known [61] that

$$\frac{1}{T}\int_0^T \left|\zeta\left(\tfrac{1}{2}+\mathrm{i}t\right)\right|^4 \,\mathrm{d}t \sim \frac{1}{2\pi^2}\left(\log\frac{T}{2\pi}\right)^4, \tag{1.82}$$

where as

$$\mathbb{E}\,|Z_U(0)|^4 = M_N(4) \sim \frac{1}{12}N^4, \tag{1.83}$$

and so the CUE model gets the correct rate of growth (since we identify N with $\log\frac{T}{2\pi}$), but fails to predict the correct asymptotic result. The reason for the failure might be explained as follows: Moment generating functions contain much more information than central limit theorems, in the sense that to get the leading order term for the moment generating function one would need to know the asymptotic expansion of the moments of $\log|\zeta(1/2+\mathrm{i}t)|$ all the way down to the constant order term. Random matrix theory only predicts the leading-order term of such an expansion.

The Keating-Snaith conjecture, which has already been introduced in section 1.1.3, "corrects" the random matrix result, $M_N(2k)$, for the zeta function.

Conjecture 1.4. (Keating & Snaith [69]). *For fixed k with $\mathfrak{Re}(k) > -1/2$,*

$$\frac{1}{T}\int_0^T \left|\zeta\left(\tfrac{1}{2}+\mathrm{i}t\right)\right|^{2k}\,\mathrm{d}t \sim a(k)\frac{G^2(k+1)}{G(2k+1)}\left(\log\frac{T}{2\pi}\right)^{k^2} \tag{1.84}$$

as $T \to \infty$, where

$$a(k) = \prod_{\substack{p\\ \text{prime}}}\left(1-\frac{1}{p}\right)^{k^2}\sum_{m=0}^{\infty}\left(\frac{\Gamma(m+k)}{m!\,\Gamma(k)}\right)^2 p^{-m} \tag{1.85}$$

Note that the conjecture agrees with what is already known for the zeta function when $k = 1, 2, 3, 4$ (see section 1.1.3). There are various other theoretical, numerical and heuristic results suggesting the truth of the Keating-Snaith conjecture. See [69, 86] for a review of such evidence.

Furthermore, this conjecture is in line with previous results for other statistics concerning the Riemann zeta function, where long range deviations from random matrix theory have also been related to the primes [9, 10, 15, 42, 68]. Many of the conjectures developed in this thesis are similar in form to the Keating-Snaith conjecture, in the sense that they involve a random matrix factor (derived from a calculation using $Z_U(\theta)$) times an arithmetic factor (which is almost always $a(k)$), and will in special cases be shown to agree with rigorous theorems about the Riemann zeta function.

1.4 Probability theory

Let Ω be a topological space, and let 2^Ω denote the set of all subsets of Ω.

Definition 1.1. *$\mathcal{A} \subseteq 2^\Omega$ is a σ-algebra of Ω if*

1. *$\Omega \in \mathcal{A}$*

2. *If $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$ where A^c is the compliment.*

3. *$\mathcal{A}$ is closed under countable unions and countable intersections.*

Definition 1.2. *The Borel σ-algebra associated with the topological space Ω is the σ-algebra generated by the open sets (i.e. the smallest σ-algebra containing all the open sets in that topology).*

Definition 1.3. *A probability measure defined on a σ-algebra $\mathcal{A}$ of Ω is a function $\mathbb{P} : \mathcal{A} \mapsto [0, 1]$ that satisfies*

1. $\mathbb{P}\{\Omega\} = 1$

2. *For every countable sequence $\{A_n\}_{n\geq 1}$ of elements of $\mathcal{A}$ which are pairwise disjoint,*

$$\mathbb{P}\left\{\bigcup_{n=1}^{\infty} A_n\right\} = \sum_{n=1}^{\infty} \mathbb{P}\{A_n\} \tag{1.86}$$

Definition 1.4. *Let $\mathcal{F}$ be the Borel σ-algebra of a topological space F. A function $X : \Omega \mapsto F$ is a random variable if $X^{-1}(\Lambda) \in \mathcal{A}$ for all $\Lambda \in \mathcal{F}$.*

We write $\mathbb{P}\{X \in \Lambda\} = \mathbb{P}\{\omega : X(\omega) \in \Lambda\}$ for the probability that a random variable X takes a value lying in some set $\Lambda \in \mathcal{F}$.

If $F = \mathbb{R}^d$, then the notion of probability density is useful:

Definition 1.5. *We say X has a probability density function $p : \Omega \mapsto \mathbb{R}^d$ if for all $\Lambda \in \mathcal{F}$*

$$\mathbb{P}\{X \in \Lambda\} = \int_{\Lambda} p(x)\, \mathrm{d}x. \tag{1.87}$$

Definition 1.6. *$\mathbb{E}\, f(X)$ denotes the expectation of $f(X)$, and if X has a density function,*

$$\mathbb{E}\, f(X) = \int_{\mathbb{R}^d} f(x)p(x)\, \mathrm{d}x. \tag{1.88}$$

Definition 1.7. *The moment generating function of X is $\mathbb{E}\, e^{\langle\lambda, X\rangle}$, and the characteristic function of X is $\mathbb{E}\, e^{\mathrm{i}\langle\lambda, X\rangle}$, where $\langle\lambda, X\rangle = \sum_{j=1}^{d} \lambda_j X_j$.*

If $F = \mathbb{R}$ and $c(\lambda) = \mathbb{E}\, e^{\mathrm{i}\lambda X}$ then by Fourier inversion, if $p(\cdot)$ exists then

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\mathrm{i}\lambda x} c(\lambda)\, \mathrm{d}\lambda. \tag{1.89}$$

1.4.1 Weak convergence and central limit theorems

Definition 1.8. *A sequence of random variables, $\{X_N\}$, taking values in some topological space F is said to converge weakly to X (written $X_N \Longrightarrow X$) as $N \to \infty$ if, for all bounded continuous functions $f : F \longrightarrow \mathbb{R}$,*

$$\lim_{N\to\infty} \mathbb{E}\, f(X_N) = \mathbb{E}\, f(X). \tag{1.90}$$

Note that the topology determines the set of test functions $f(\cdot)$.

Proving weak convergence using the above definition is usually difficult, so, in the case $F = \mathbb{R}^d$, we use:

Theorem 1.5. *If there exists a $B > 0$ such that for for all $\lambda \in \mathbb{R}^d$ with $|\lambda| \leq B$,*

$$\lim_{N\to\infty} \mathbb{E}\, e^{\langle \lambda, X_N\rangle} = \mathbb{E}\, e^{\langle \lambda, X\rangle} < \infty \tag{1.91}$$

then $X_N \Longrightarrow X$ as $N \to \infty$.

For a proof of this fact, see, for example, p.345 of [12].

Remark. If X has a density on $\mathbb{R}^d$ and $X_N \Longrightarrow X$ then for any measurable set $A \subseteq \mathbb{R}^d$,

$$\mathbb{P}\{X_N \in A\} \to \mathbb{P}\{X \in A\} \qquad \text{as } N \to \infty. \tag{1.92}$$

Central limit theorems (CLT), also called fluctuation theorems, are concerned with when there exists a sequence of random variables Y_n and a function $\sigma = \sigma(N)$ such that

$$X_N = \frac{1}{\sigma}\sum_{n=1}^{N} Y_n \Longrightarrow X \tag{1.93}$$

Since $\log Z_U(\theta) = \sum_{n=1}^{N} \log\left(1 - e^{\mathrm{i}(\theta_n - \theta)}\right)$, it seems natural to use the term central limit theorem for weak convergence results on $\log Z_U(\theta)$.

Definition 1.9. *A normal random variable with mean μ and variance σ^2 has probability density*

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \tag{1.94}$$

and a standard normal has mean zero and variance 1.

Definition 1.10. *A standard complex normal random variable is $X + \mathrm{i}Y$ where X, Y are iid standard normal random variables.*

Warning. Some authors adopt the convention that the real and imaginary parts of a complex normal have variance 1/2.

One particular corollary to theorem 1.5 is that if for each $\lambda \in \mathbb{R}$,

$$\lim_{N\to\infty} \mathbb{E}\, e^{\lambda X_N} = e^{\lambda^2/2} \tag{1.95}$$

then $X_N \implies \mathcal{N}(0,1)$ (that is X_N converges in distribution to a standard normal random variable).

1.4.2 Large deviation theory

We present with a quick review of large deviation theory (see, for example, [20, 34]).

Let $\{R_N\}$, $N = 1, 2, \ldots$, be a sequence of $\mathbb{R}^d$ valued random variables.

Definition 1.11. *The $\{R_N\}$ satisfies the Large Deviation Principle (LDP) with rate function $I(\cdot)$ and speed $B(N)$ if, for all measurable $\Gamma \in \mathbb{R}^d$,*

$$\begin{aligned} -\inf_{x\in\Gamma^\circ} I(x) \le \liminf_{N\to\infty} \frac{1}{B(N)} \log \mathbb{P}\{R_N \in \Gamma\} \\ \le \limsup_{N\to\infty} \frac{1}{B(N)} \log \mathbb{P}\{R_N \in \Gamma\} \le -\inf_{x\in\overline{\Gamma}} I(x) \end{aligned} \tag{1.96}$$

where Γ° means the interior of the set, and $\overline{\Gamma}$ means the closure of the set. The speed must tend to infinity as $N \to \infty$ (for otherwise we would be in the central limit regime).

Remark. If $\inf_{x\in\Gamma^\circ} I(x) = \inf_{x\in\overline{\Gamma}} I(x)$, then we have the stronger result

$$\lim_{N\to\infty} \frac{1}{B(N)} \log \mathbb{P}\{R_N \in \Gamma\} = -\inf_{x\in\Gamma} I(x) \tag{1.97}$$

The Gärtner-Ellis theorem is a list of sufficient conditions for R_N to satisfy an LDP.

Assumption 1.1. *For each $\lambda \in \mathbb{R}^d$,*

$$\Lambda(\lambda) := \lim_{N\to\infty} \frac{1}{B(N)} \log \mathbb{E} \exp\left(B(N)\langle \lambda, R_N\rangle\right) \tag{1.98}$$

exists as an extended real number (that is, in $\mathbb{R}^d \cup \{\infty\}$), where $\langle\cdot,\cdot\rangle$ denotes the usual dot product in $\mathbb{R}^d$.

Assumption 1.2. $0 \in \mathcal{D}_\Lambda^\circ$, *where* $\mathcal{D}_\Lambda := \left\{\lambda \in \mathbb{R}^d : \Lambda(\lambda) < \infty\right\}$ *is the effective domain of* $\Lambda(\cdot)$.

Assumption 1.3. $\Lambda(\cdot)$ *is differentiable throughout* $\mathcal{D}_\Lambda^\circ$

Assumption 1.4. $\Lambda(\cdot)$ *is steep, namely,* $\lim_{n\to\infty} |\nabla\Lambda(\lambda_n)| = \infty$ *whenever* $\{\lambda_n\}$ *is a sequence in* $\mathcal{D}_\Lambda^\circ$ *converging to a boundary point of* $\mathcal{D}_\Lambda^\circ$.

Assumption 1.5. $\Lambda(\cdot)$ *is lowersemicontinuous, namely, for all* $\alpha \in [0,\infty)$, *the level set* $\left\{\lambda \in \mathbb{R}^d : \Lambda(\lambda) \le \alpha\right\}$ *is a closed subset of* $\mathbb{R}^d$.

Theorem 1.6. Gärtner-Ellis. *If assumptions 1.1–1.5 hold, then* R_N *satisfies an LDP with speed* $B(N)$ *and rate function* $\Lambda^*(x)$, *where*

$$\Lambda^*(x) := \sup_{\lambda\in\mathbb{R}^d} \left\{\langle\lambda, x\rangle - \Lambda(\lambda)\right\}. \tag{1.99}$$

is the Fenchel-Legendre transform of $\Lambda(\lambda)$.

In many cases we will consider in this thesis assumptions 1.2 and 1.4 fail. The following is a version of theorem 1.6 in the case $d = 1$ which requires weaker assumptions.

Theorem 1.7. *If assumption 1.1 holds, then for* $a < b$,

$$\limsup_{N\to\infty} \frac{1}{B(N)} \log \mathbb{P}\left\{R_N \in [a,b]\right\} \le -\inf_{x\in[a,b]} \Lambda^*(x). \tag{1.100}$$

If, in addition, assumption 1.3 holds, and if

$$(a,b) \subseteq \{\Lambda'(\lambda) : \lambda \in \mathcal{D}^\circ\}, \tag{1.101}$$

then

$$\liminf_{N\to\infty} \frac{1}{B(N)} \log \mathbb{P}\left\{R_N \in (a,b)\right\} \ge -\inf_{x\in(a,b)} \Lambda^*(x). \tag{1.102}$$

1.5 Calculation techniques for random unitary matrices

Many of the calculations required in this thesis are of the form

$$\mathbb{E}_N\left\{\prod_{n=1}^{N} f(\theta_n)\right\} = \frac{1}{(2\pi)^N N!}\int\cdots\int_{-\pi}^{\pi}\prod_{1\le j<k\le N}\left|e^{\mathrm{i}\theta_j}-e^{\mathrm{i}\theta_k}\right|^2\prod_{n=1}^{N} f(\theta_n)\,\mathrm{d}\theta_n \quad (1.103)$$

where $\mathbb{E}_N$ denotes expectation with respect to the CUE, and θ_n are the eigenangles of an $N\times N$ unitary matrix.

For certain functions f, such an integration can be evaluated (after a change of variables) using a form of Selberg's integral (discussed in chapter 17 of [74]):

Lemma 1.8. (Selberg [84]).

$$\int\cdots\int_{-\infty}^{\infty}\prod_{1\le j<k\le N}|x_j-x_k|^{2\gamma}\prod_{n=1}^{N}(a+\mathrm{i}x_n)^{-\alpha}(b-\mathrm{i}x_n)^{-\beta}\mathrm{d}x_n =$$
$$\frac{(2\pi)^N}{(a+b)^{(\alpha+\beta)N-\gamma N(N-1)-N}}\prod_{j=0}^{N-1}\frac{\Gamma(1+(j+1)\gamma)\Gamma(\alpha+\beta-1-(N+j-1)\gamma)}{\Gamma(1+\gamma)\Gamma(\alpha-j\gamma)\Gamma(\beta-j\gamma)} \quad (1.104)$$

where $a, n, \alpha, \beta, \gamma$ are all complex numbers subject to $\mathfrak{Re}(a), \mathfrak{Re}(b), \mathfrak{Re}(\alpha)$ and $\mathfrak{Re}(\beta)$ all greater than zero, $\mathfrak{Re}(\alpha+\beta)>1$ and

$$-\frac{1}{N} < \mathfrak{Re}(\gamma) < \min\left(\frac{\mathfrak{Re}(\alpha)}{N-1}, \frac{\mathfrak{Re}(\beta)}{N-1}, \frac{\mathfrak{Re}(\alpha+\beta-1)}{2(N-1)}\right) \quad (1.105)$$

(there's a slight error in [74] on this condition).

The following lemma is found to be very useful:

Lemma 1.9. *Writing $\mathbb{E}_N$ to denote expectation with respect to Haar measure over $\mathcal{U}(N)$, then if $f(\cdot)$ is a 2π–periodic function,*

$$\mathbb{E}_N\left\{\prod_{n=1}^{N-1} f(\theta_n-\theta_N)\right\} = \frac{1}{N}\,\mathbb{E}_{(N-1)}\left\{\prod_{n=1}^{N-1}\left|1-e^{\mathrm{i}\theta_n}\right|^2 f(\theta_n)\right\} \quad (1.106)$$

Proof. By the definition of expectation, and (1.41), we have

$$\mathbb{E}_N\left\{\prod_{n=1}^{N-1} f(\theta_n-\theta_N)\right\} =$$
$$\frac{1}{N!(2\pi)^N}\int\cdots\int_{-\pi}^{\pi}\prod_{1\le j<k\le N}\left|e^{\mathrm{i}\theta_j}-e^{\mathrm{i}\theta_k}\right|^2\prod_{n=1}^{N-1} f(\theta_n-\theta_N)\prod_{p=1}^{N}\mathrm{d}\theta_p. \quad (1.107)$$

Putting all the $k = N$ terms from the first product into the second product, this equals

$$\frac{1}{N!(2\pi)^N}\int_{-\pi}^{\pi}\left\{\int\cdots\int_{-\pi}^{\pi}\prod_{1\le j<k\le(N-1)}\left|e^{\mathrm{i}\theta_j}-e^{\mathrm{i}\theta_k}\right|^2\times\right.$$
$$\left.\times\prod_{n=1}^{N-1}\left|e^{\mathrm{i}\theta_n}-e^{\mathrm{i}\theta_N}\right|^2 f\left(\theta_n-\theta_N\right)\,\mathrm{d}\theta_n\right\}\mathrm{d}\theta_N \quad (1.108)$$

Changing variables to $\phi_n = \theta_n - \theta_N$, and using the 2π–periodicity of the integrand, one can write

$$\int\cdots\int_{-\pi}^{\pi}\prod_{1\le j<k\le(N-1)}\left|e^{\mathrm{i}\theta_j}-e^{\mathrm{i}\theta_k}\right|^2\prod_{n=1}^{N-1}\left|e^{\mathrm{i}\theta_n}-e^{\mathrm{i}\theta_N}\right|^2 f\left(\theta_n-\theta_N\right)\,\mathrm{d}\theta_n$$
$$=\int\cdots\int_{-\pi}^{\pi}\prod_{1\le j<k\le(N-1)}\left|e^{\mathrm{i}\phi_j}-e^{\mathrm{i}\phi_k}\right|^2\prod_{n=1}^{N-1}\left|1-e^{\mathrm{i}\phi_n}\right|^2 f\left(\phi_n\right)\,\mathrm{d}\phi_n$$
$$=(N-1)!(2\pi)^{N-1}\,\mathbb{E}_{(N-1)}\left\{\prod_{n=1}^{N-1}\left|1-e^{\mathrm{i}\phi_n}\right|^2 f\left(\phi_n\right)\right\} \quad (1.109)$$

which is independent of θ_N. Therefore (1.108) equals

$$\frac{1}{2\pi N}\int_{-\pi}^{\pi}\mathbb{E}_{(N-1)}\left\{\prod_{n=1}^{N-1}\left|1-e^{\mathrm{i}\phi_n}\right|^2 f\left(\phi_n\right)\right\}\,\mathrm{d}\theta_N \quad (1.110)$$

where $\phi_1,\ldots,\phi_{N-1}$ are the eigenangles of an $(N-1)\times(N-1)$ Haar distributed unitary matrix. This completes the proof, since the integrand is independent of θ_N. □

1.5.1 Toeplitz matrices

Let $f(\cdot)$ be a real-valued, 2π-periodic, integrable function, and let

$$\hat{f}_k = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)e^{-\mathrm{i}k\theta}\,\mathrm{d}\theta \quad (1.111)$$

denote its Fourier coefficients. The Nth order Toeplitz determinant with symbol f is defined by

$$D_N[f] = \det(\hat{f}_{j-k})_{1\le j,k\le N}. \quad (1.112)$$

[18] is a recent review of the Toeplitz literature.

The reason we find Toeplitz matrices useful is:

Lemma 1.10. (Heine's[11] identity). *If $\mathbb{E}_N$ denotes expectation taken over $N \times N$ Haar distributed unitary matrices, then*

$$D_N[f] = \mathbb{E}_N \prod_{n=1}^{N} f(\theta_n). \tag{1.113}$$

Proof. (See, for example, [89]). By (1.41)

$$\mathbb{E} \prod_{n=1}^{N} f(\theta_n) = \frac{1}{N!(2\pi)^N} \int \cdots \int_{-\pi}^{\pi} \prod_{1 \le j < k \le N} \left| e^{\mathrm{i}\theta_j} - e^{\mathrm{i}\theta_k} \right|^2 \prod_{n=1}^{N} f(\theta_n) \, \mathrm{d}\theta_n. \tag{1.114}$$

Note that

$$\prod_{1 \le j < k \le N} \left| e^{\mathrm{i}\theta_j} - e^{\mathrm{i}\theta_k} \right|^2 = \prod_{1 \le j < k \le N} \left(e^{\mathrm{i}\theta_j} - e^{\mathrm{i}\theta_k} \right) \left(e^{-\mathrm{i}\theta_j} - e^{-\mathrm{i}\theta_k} \right) \tag{1.115}$$

$$= \det \left\{ e^{\mathrm{i}(j-1)\theta_k} \right\}_{1 \le j,k \le N} \det \left\{ e^{-\mathrm{i}(m-1)\theta_n} \right\}_{1 \le m,n \le N} \tag{1.116}$$

(the determinants are Vandermonde). Expanding each determinant out in terms of a sum over the permutation group, this equals

$$\sum_{\sigma \in S(N)} \operatorname{sgn}(\sigma) \prod_{k=1}^{N} e^{\mathrm{i}(\sigma(k)-1)\theta_k} \sum_{\tau \in S(N)} \operatorname{sgn}(\tau) \prod_{n=1}^{N} e^{-\mathrm{i}(\tau(n)-1)\theta_n}$$

$$= \sum_{\sigma \in S(N)} \sum_{\tau \in S(N)} \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) \prod_{n=1}^{N} e^{\mathrm{i}(\sigma(n)-\tau(n))\theta_n}. \tag{1.117}$$

Hence

$$\mathbb{E} \prod_{n=1}^{N} f(\theta_n) = \frac{1}{N!(2\pi)^N} \int \cdots \int_{-\pi}^{\pi} \sum_{\substack{\sigma \in S(N) \\ \tau \in S(N)}} \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) \prod_{n=1}^{N} e^{\mathrm{i}(\sigma(n)-\tau(n))\theta_n} f(\theta_n) \, \mathrm{d}\theta_n \tag{1.118}$$

$$= \sum_{\sigma, \tau \in S(N)} \frac{\operatorname{sgn}(\sigma) \operatorname{sgn}(\tau)}{N!} \prod_{n=1}^{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\mathrm{i}(\sigma(n)-\tau(n))\theta} f(\theta) \, \mathrm{d}\theta \tag{1.119}$$

$$= \sum_{\sigma, \tau \in S(N)} \frac{\operatorname{sgn}(\sigma) \operatorname{sgn}(\tau)}{N!} \prod_{n=1}^{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\mathrm{i}(\sigma(\tau^{-1}(n))-n)\theta} f(\theta) \, \mathrm{d}\theta \tag{1.120}$$

by the symmetry induced by the product. Writing $\alpha = \sigma(\tau^{-1})$, then

$$\mathbb{E} \prod_{n=1}^{N} f(\theta_n) = \sum_{\alpha \in S(N)} \operatorname{sgn}(\alpha) \prod_{n=1}^{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\mathrm{i}(\alpha(n)-n)\theta} f(\theta) \, \mathrm{d}\theta \tag{1.121}$$

$$= D_N[f] \tag{1.122}$$

as required. □

[11] Though the identity was first written down in 1939 by Szegő [89], he gave the credit to Heine (who lived from 1821 to 1881).

Methods of calculating $D_N[f]$ for certain symbols f (for example, Szegő's theorem, the Fisher-Hartwig conjecture, and the Basor-Forrester method) will be introduced in the text when required (in section 2.3, lemma 2.1 and lemma 6.2 respectively).

1.6 Notations

Let $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ be the unit circle. A function is defined on $\mathbb{T}$ if it is a 2π-periodic function defined in $\mathbb{R}$.

The following definitions concern limits at infinity:

- $f(x) = \mathcal{O}(g(x))$ as $x \to \infty$ if there exists a positive constant A such that $\limsup_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| \leq A$.
- $f(x) = o(g(x))$ as $x \to \infty$ if $\lim_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| = 0$.
- $f(x) = \Omega(g(x))$ as $x \to \infty$ if there exists a positive constant B such that $\limsup_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| \geq B$.
- $f(x) \asymp g(x)$ as $x \to \infty$ if $f = \mathcal{O}(g)$ and $f = \Omega(g)$.
- $f(x) \sim g(x)$ as $x \to \infty$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.
- $f(x) \ll g(x)$ as $x \to \infty$ if $f = o(g)$.
- $f(x) \gg g(x)$ as $x \to \infty$ if $g = o(f)$.
- $f(x) << g(x)$ as $x \to \infty$ if $f = \mathcal{O}(g)$.
- $f(x) >> g(x)$ as $x \to \infty$ if there exists a positive constant C such that $\liminf_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| \geq C$.

1.7 Overview of this thesis

This thesis splits into two parts. Part one (chapters 2 and 3) develop the theory of the characteristic polynomial, $Z_U(\theta)$, in terms of central limit theorems (chapter 2) and large deviations (chapter 3). Part two (chapters 4,5 and 6) use $Z_U(\theta)$ as a model for zeta, obtaining various conjectures in the style of the Keating-Snaith conjecture.

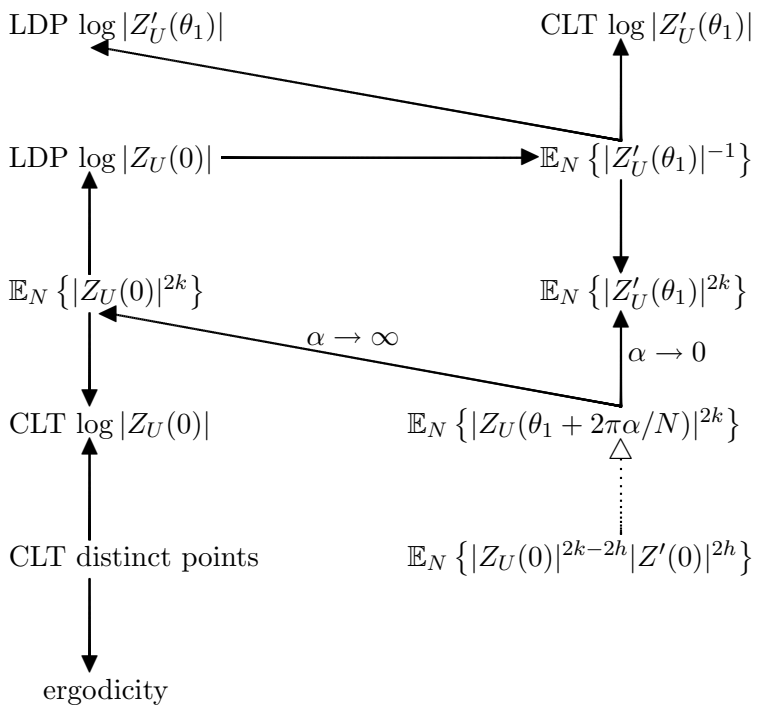

Figure 1.2: Overview of how results on $\mathfrak{Re}\log Z_U(\theta)$ relate to each other

1.7.1 Things not considered here

I have concentrated solely on the Riemann zeta function and the unitary group in this thesis. However, the characteristic polynomial philosophy should apply to many other interesting problems. For example:

- The characteristic polynomial can be formed in any ensemble, and its properties studied. Results for the imaginary part of the logarithm, for example, would then connect up with the work of Wieand [96] on the counting function in the orthogonal and symplectic group ensemble, as well as the permutation group, in an analogous manner to how the results in this thesis connect with her results on the counting function over $\mathcal{U}(N)$. The characteristic polynomial in the permutation group has been studied in [51], and in the orthogonal and symplectic group ensembles in [70].

- As is mentioned in [60], there is a connection between the functional central limit theorem in section 2.3 and Brownian motion.

- One can use random matrix theory to study other L–functions. For example,

Rudnick and Sarnak [83] showed that the zeros of principal L–functions are correlated like the eigenangles of the CUE[12]. This connection has been generalized by Katz and Sarnak in [66], where they consider the low-lying zeros of various families of L–functions, conjecturing that they are modeled by either the unitary, orthogonal or symplectic group ensembles, depending on the family. See also [25, 19, 70].

- It might be possible to model the Selberg zeta function using these methods.
- There are many connections between random matrix theory and quantum chaos. For example, deviations from random matrix theory in the counting function of energy levels is of interest.

[12] for certain test functions—the restriction being essentially the same as the restriction in Montgomery's theorem (theorem 1.1 in this thesis)

Chapter 2

Central Limit Theorems

We define $W_U(\theta) = \log Z_U(\theta)/\sigma$, where $\sigma = \sqrt{\frac{1}{2}\log N}$. Recall that the CLT for $\log Z_U(0)/\sigma$ (theorem 1.3) is in good agreement with Selberg's central limit theorem for the Riemann zeta function, (1.71), if one accepts the connection $N = \log \frac{T}{2\pi}$ between random matrix theory and zeta. For this reason (and others) it is of interest to obtain more detailed fluctuation theorems for $\log Z_U(\theta)$. Such results are the subject of this chapter.

Our first generalization of theorem 1.3 is to show that $W_U(\theta)$, when evaluated at k distinct points, weakly converges to k i.i.d. standard complex normal random variables. We then use this to show that the distribution of $W_U(0)$ when considered as a random variable over the CUE is asymptotically the same as the value distribution of $W_U(\theta)$, when thought of as a random variable over θ, for a typical matrix U. This ergodicity theorem is the main result of this chapter.

Next we show how our results for $\mathfrak{Im} W_U(\theta)$ can be used to explain the mysterious covariance structure which has been observed by Costin and Lebowitz [33] and Wieand [96, 97] in the eigenvalue counting function. We then discuss the connection between the counting functions for random unitary matrices and for the Riemann zeta function, in particular focussing on the number variance, summarising previous results where long-range deviations from random matrix theory have been studied.

And finally, we also obtain a functional central limit theorem for $\log Z_U(\theta)$.

As always, in this chapter we take the underlying probability space for $\mathbb{E}$ and $\mathbb{P}$ to be $\mathcal{U}(N)$ with Haar measure.

Many of the results in the chapter have previously been published in [60].

2.1 Central limit theorems for $W_U(\theta)$

In order to prove the central limit theorems, we need to understand the asymptotic behaviour of the joint moment generating function:

Lemma 2.1. *For any $d(N) \to \infty$ as $N \to \infty$, for any fixed $s, t \in \mathbb{R}^k$ with N sufficiently large such that $s_j > -d(N)$ for all j, and for any fixed r_j distinct in $\mathbb{T}$, we have*

$$\mathbb{E}\exp\left(\sum_{j=1}^{k} s_j \mathfrak{Re}\log Z_U(r_j)/d + t_j \mathfrak{Im}\log Z_U(r_j)/d\right) \tag{2.1}$$

$$\sim \prod_{j=1}^{k} \mathbb{E}\exp\left(s_j \mathfrak{Re}\log Z_U(r_j)/d + t_j \mathfrak{Im}\log Z_U(r_j)/d\right) \tag{2.2}$$

$$\sim \exp\left(\sum_{j=1}^{k} \frac{\log N}{4d^2}(s_j^2 + t_j^2)\right). \tag{2.3}$$

Proof. This follows from Heine's identity (lemma 1.10) and a result of Basor [6] on the asymptotic behaviour of Toeplitz determinants with Fisher-Hartwig symbols.

Heine's identity says that

$$\mathbb{E}\exp\left(\sum_{j=1}^{k} s_j \mathfrak{Re}\log Z_U(r_j)/d + t_j \mathfrak{Im}\log Z_U(r_j)/d\right) = D_N[f] \tag{2.4}$$

where $D_N[f]$, defined in (1.112), is the Toeplitz determinant of the symbol

$$f(\theta) = \prod_{j=1}^{k} \left|1 - e^{\mathrm{i}(\theta - r_j)}\right|^{s_j/d} \exp\left(\frac{t_j}{d}\mathfrak{Im}\log(1 - e^{\mathrm{i}(\theta - r_j)})\right) \tag{2.5}$$

$$= \prod_{j=1}^{k} (1 - e^{\mathrm{i}(\theta - r_j)})^{\alpha_j + \beta_j}(1 - e^{\mathrm{i}(r_j - \theta)})^{\alpha_j - \beta_j}, \tag{2.6}$$

where $\alpha_j = s_j/2d$ and $\beta_j = -\mathrm{i}t_j/2d$. This is an example of a Fisher-Hartwig symbol. A review of the literature on Toeplitz determinants of Fisher-Hartwig symbols can be found in [18]. In particular, Basor [6] proves that as $N \to \infty$, if the r_j are distinct in $\mathbb{T}$, the $\alpha_j > -1/2$ are real and the β_j are purely imaginary, then

$$D_N[f] \sim F(\alpha_1, \beta_1, r_1, \ldots, \alpha_k, \beta_k, r_k) \prod_{j=1}^{k} N^{\alpha_j^2 - \beta_j^2}, \tag{2.7}$$

where

$$F(\alpha_1, \beta_1, r_1, \ldots, \alpha_k, \beta_k, r_k) = \prod_{\substack{1 \le m,n \le k \\ m \ne n}} \left(1 - e^{\mathrm{i}(r_m - r_n)}\right)^{-(\alpha_m - \beta_m)(\alpha_n + \beta_n)} \prod_{j=1}^{k} \frac{G(1+\alpha_j+\beta_j)G(1+\alpha_j-\beta_j)}{G(1+2\alpha_j)}, \quad (2.8)$$

where $\left|\arg\left(1 - e^{\mathrm{i}(r_m - r_n)}\right)\right| \le \pi/2$. By closer inspection of the proof of [6], it can be seen that (2.7) holds uniformly for $|\alpha_j| < 1/2 - \delta$, and $|\beta_j| < \gamma$, for any fixed $\delta, \gamma > 0$[1]. Uniformity in β is worked out carefully in [96, 97] in the case $\alpha_j = 0$ for each j, and uniformity in α is briefly discussed in [6]. The statement of the lemma follows from noting that $F(0, 0, r_1, \ldots, 0, 0, r_k) = 1$. □

Theorem 2.2. *If $r_1, \ldots, r_k \in \mathbb{T}$ are distinct, then $(W_U(r_1), \ldots, W_U(r_k))$ converges in distribution to k iid standard complex normal random variables.*

Proof. Setting $d = \sigma$ in lemma 2.1, we see that the joint moment generating function of $(W_U(r_1), \ldots, W_U(r_k))$ converges to that of k iid standard complex normal random variables, which, by theorem 1.5, is sufficient to prove convergence in distribution. □

Remark. Diaconis and Evans [35] have given an alternative proof of this theorem, though they define a standard complex normal random variable to be such that the real and imaginary parts have variance equal to 1/2 rather than 1.

Remark. The case $k = 1$ of this theorem is theorem 1.3, as found in [69] (but the method of proof there is very different. They use Selberg's integral, lemma 1.8, to calculate the exact moment generating function of $\log Z_U(r_1)$, (1.73)).

One application of theorem 2.2 is that for a typical matrix U, the value distribution of $W_U(\theta)$ when θ is chosen uniformly from $\mathbb{T}$ is asymptotically the same as the value distribution of $W_U(0)$ when U is chosen according to Haar measure (which is a standard complex normal random variable, by theorem 1.3 or theorem 2.2 above):

Theorem 2.3. *Denote by m the uniform probability measure on $\mathbb{T}$ (so that $m(\mathrm{d}\theta) = \mathrm{d}\theta/2\pi$). The sequence of laws $\left\{m \circ (W_U(\theta))^{-1}\right\}$ converges weakly in probability to a standard complex normal variable.*

[1] This was pointed out to us by Harold Widom, in an email to Neil O'Connell.

What do we mean by this? For any set $A \subseteq \mathbb{C}$, $(W_U(\theta))^{-1}(A)$ is the set of all $\theta \in (-\pi, \pi]$ such that $W_U(\theta) \in A$. So, $m \circ (W_U(\theta))^{-1}$ is the probability distribution of $W_U(\theta)$ over θ, for a given matrix U.

Recall theorem 1.5, which states that weak convergence of a family of probability measures, $\{\mu_n(\cdot)\}$, to $\mu(\cdot)$ is implied by the pointwise convergence of the moment generating functions, as $n \to \infty$. However, for each N (the matrix size) we don't have just one measure, but an infinite set, as $m \circ (W_U(\theta))^{-1}$ depends on the matrix, not just on the matrix size. Weak convergence in probability means that the measure of those matrices which don't converge weakly to a standard complex normal is vanishingly small. We will prove theorem 2.3 by showing that for each s, t, for any $\epsilon > 0$,

$$\lim_{N\to\infty} \mathbb{P}_N \left\{ \left| \int_{\mathbb{T}} e^{s\mathfrak{Re}W_U(\theta) + t\mathfrak{Im}W_U(\theta)} m(\mathrm{d}\theta) - e^{(s^2+t^2)/2} \right| > \epsilon \right\} = 0, \tag{2.9}$$

where $\mathbb{P}_N$ denotes the probability with respect to Haar measure on $\mathcal{U}(N)$.

Proof of theorem 2.3. Set $X_U(\theta) = \mathfrak{Re}W_U(\theta)$, $Y_U(\theta) = \mathfrak{Im}W_U(\theta)$ and

$$\phi_U(s,t) = \int_{\mathbb{T}} \exp\left(sX_U(\theta) + tY_U(\theta)\right) m(\mathrm{d}\theta). \tag{2.10}$$

By theorem 1.3 (which is the case $k = 1$ of theorem 2.2),

$$\mathbb{E}\,\phi_U(s,t) = \mathbb{E}\exp\left(sX_U(0) + tY_U(0)\right) \tag{2.11}$$

$$\to e^{(s^2+t^2)/2}. \tag{2.12}$$

Due to the rotation invariance of the CUE, we have

$$\mathbb{E}\left\{\phi_U(s,t)^2\right\} = \int_{\mathbb{T}} \mathbb{E}\exp\left(sX_U(\theta) + tY_U(\theta) + sX_U(0) + tY_U(0)\right) m(\mathrm{d}\theta) \tag{2.13}$$

and for fixed $\theta \neq 0$, theorem 2.2 gives

$$\mathbb{E}\exp\left(sX_U(\theta) + tY_U(\theta) + sX_U(0) + tY_U(0)\right) \to e^{s^2+t^2} \tag{2.14}$$

However, since pointwise convergence of the integrand does not imply convergence of the integral, this is insufficient to show $\mathbb{E}\left\{\phi_U(s,t)^2\right\} \to e^{s^2+t^2}$ as $N \to \infty$. The Cauchy-Schwartz inequality gives

$$\begin{aligned}\mathbb{E}\left\{\exp\left(sX_U(\theta) + tY_U(\theta) + sX_U(0) + tY_U(0)\right)\right\} \leq \\ \sqrt{\mathbb{E}\left\{\exp\left(2sX_U(\theta) + 2tY_U(\theta)\right)\right\}}\sqrt{\mathbb{E}\left\{\exp\left(2sX_U(0) + 2tY_U(0)\right)\right\}} \\ = \mathbb{E}\left\{\exp\left(2sX_U(0) + 2tY_U(0)\right)\right\}\end{aligned} \tag{2.15}$$

and for each $\epsilon > 0$ there exists an $N_0(\epsilon)$ such that for all $N > N_0$ the last line is bounded above by $e^{2s^2+2t^2} + \epsilon$ for all θ by theorem 1.3. Thus, by the bounded convergence theorem, it is indeed true that

$$\mathbb{E}\left\{\phi_U(s,t)^2\right\} \to e^{s^2+t^2}. \tag{2.16}$$

Hence

$$\mathbb{P}_N(|\phi_U(s,t) - e^{(s^2+t^2)/2}| > \epsilon) \leq \frac{\text{Var } \phi_U(s,t)}{\epsilon^2} \tag{2.17}$$

$$\to 0 \quad \text{as } N \to \infty, \tag{2.18}$$

for any $\epsilon > 0$, by Chebyshev's inequality. Thus, for each s, t, the sequence $\phi_U(s,t)$ converges in probability to $e^{(s^2+t^2)/2}$, which is the moment generating function of a standard complex random variable. This therefore proves the theorem. □

Remark. Theorem 2.3 says that the supremum of the vertical distance[2] between the graphs of the cumulative distribution function for each matrix and the cumulative distribution function of a standard complex normal random variable tends to zero as $N \to \infty$, with probability one.

We will now discuss how these central limit theorems compare with what is known for the Riemann zeta function

The first connection is Selberg's distribution theorem for $\log \zeta(1/2 + \mathrm{i}t)$, (1.71). However, our results suggest that Selberg's result can be strengthened in at least two ways. For simplicity, set

$$Y_T(t) = \frac{\mathfrak{Im} \log \zeta(1/2 + \mathrm{i}t)}{\sqrt{\frac{1}{2} \log \log T}}. \tag{2.19}$$

First of all, theorem 2.3 shows that the distribution of almost all matrices (which is an average over the N zeros of $\log Z_U(\theta)/\sigma$) is asymptotically the same as the distribution obtained from the full unitary group. Therefore, we might expect the range over which the zeta function is averaged, $[T, 2T]$, in (1.71) might be able to be reduced to $[T, T+H]$, with $H \gg 1$. (Note that theorem 2.3 is a statement about "almost-all" matrices, which is why we take $H \gg 1$ rather than $H = 2\pi$, as this equates to an average over many matrices, hence getting rid of the "almost all").

[2]This is known as the Kolmogoroff-Smirnov discrepancy.

There are some known results in this direction: For example, when calculating the empirical distribution of $\mathfrak{Re}\log\zeta(1/2+\mathrm{i}t)$ around the 10^{20}th zero (which is plotted in figure 1.1), Odlyzko [78] averaged over only 10^6 zeros starting at zero number $10^{20}+15,316,087$, which equates to $T=1.52\times10^{19}$ and $H=148,437$. Also, Selberg [85] has shown that for $T^a\le H\le T$, $a>1/2$ ($a>0$ under the assumption of RH)

$$\lim_{T\to\infty}\frac{1}{H}\int_T^{T+H}\{Y_T(t)\}^{2k}\,\mathrm{d}t=\frac{(2k)!}{2^k k!} \tag{2.20}$$

for k an integer. These are the even integer moments of a standard normal random variable, thus implying that for $a<b$,

$$\lim_{T\to\infty}\frac{1}{H}\operatorname{meas}\{T\le t\le T+H\ :\ Y_T(t)\in(a,b)\}=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}\,\mathrm{d}x. \tag{2.21}$$

Secondly, theorem 2.2 suggests that the distribution of zeta between two blocks $[T,T+H]$ and $[T+H,T+2H]$ should be independent, for any $1\ll H\le T$. Tsang [92] has shown that for $T^a\le H\le T$, $a>1/2$, and $0<h<1$,

$$\lim_{T\to\infty}\frac{1}{H}\int_T^{T+H}\{Y_T(t+h)-Y_T(t)\}^{2k}\,\mathrm{d}t=\frac{(2k)!}{k!} \tag{2.22}$$

which is consistent with the distribution of $Y_T(t+h)$ being asymptotically independent from $Y_T(t)$, although one cannot deduce such a result from (2.22) alone.

2.2 The counting function

For $-\pi<s<t\le\pi$, let $C_U(s,t)$ denote the number of eigenangles of U that lie in the interval (s,t), and define

$$\widetilde{C}_U(s,t):=\frac{C_U(s,t)-(t-s)N/2\pi}{\frac{1}{\pi}\sqrt{\log N}}. \tag{2.23}$$

Wieand [96, 97] proves that for fixed s,t, $\widetilde{C}_U(s,t)$ converges in distribution to a standard normal random variable. In fact she goes much further by proving weak convergence of $\widetilde{C}_U(s,t)$ to a certain Gaussian process $C(s,t)$.

Definition 2.1. *A centered Gaussian process is a collection of normal random variables $\{X(\alpha),\ \alpha\in I\}$, with mean zero and the property that for any $\alpha_1,\dots,\alpha_m\in I$, the joint distribution of $(X(\alpha_1),\dots,X(\alpha_m))$ is multivariate normal.*

Theorem 2.4. (Wieand [96, 97]). *For $-\pi < s < t \leq \pi$, the finite dimensional distributions of the process $\widetilde{C}_U(s,t)$ converge as $N \to \infty$ to those of a centered Gaussian process $C(s,t)$ with covariance structure*

$$\mathbb{E}\left\{C(s,t)C(s',t')\right\} = \begin{cases} 1 & \textit{if } s = s', t = t' \\ -1 & \textit{if } s = t', t = s' \\ \frac{1}{2} & \textit{if } s = s' \textit{ or if } t = t' \textit{ but not both} \\ -\frac{1}{2} & \textit{if } s = t' \textit{ or if } t = s' \textit{ but not both} \\ 0 & \textit{otherwise} \end{cases} \tag{2.24}$$

(A similar process result had previously been found by Costin and Lebowitz [33] for GUE matrices, and Soshnikov [87] considers a process result for counting the number of eigenangles in an interval with a given minimum displacement). To quote from [96], what these correlations are saying is that "if an interval contains significantly more than the average number of eigenangles, than an interval next to it will usually have fewer than the average number, [whereas if an] interval contains more than the average number, a subset [sharing a common endpoint] usually will as well". The surprising thing about these correlations is that they imply that even if an interval contains more than the average number of eigenangles, then any subset not sharing a common endpoint will usually still contain *its* average number. Also, no matter how close together two intervals are, unless they share an endpoint, then they are not correlated.

An explanation for this covariance structure can be found using theorem 2.2:

First of all, note that $C(s,t) = Y(t) - Y(s)$ where $Y(s)$ is a centered Gaussian random variable with covariance structure

$$\mathbb{E}\left\{Y(s)Y(t)\right\} = \begin{cases} \frac{1}{2} & \text{if } s = t \\ 0 & \text{otherwise} \end{cases} \tag{2.25}$$

Secondly, by rearranging the Fourier expansion of the indicator function,

$$\mathbb{1}_{\{\theta \in (s,t)\}} = \frac{t-s}{2\pi} + \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \frac{\mathrm{i}}{2\pi k}(e^{-\mathrm{i}kt} - e^{-\mathrm{i}ks})e^{\mathrm{i}k\theta} \tag{2.26}$$

$$= \frac{t-s}{2\pi} + \frac{1}{\pi}\mathfrak{Im}\log(1 - e^{\mathrm{i}(\theta - t)}) - \frac{1}{\pi}\mathfrak{Im}\log(1 - e^{\mathrm{i}(\theta - s)}). \tag{2.27}$$

Therefore

$$C_U(s,t) = \sum_{n=1}^{N} \mathbb{1}_{\{\theta_n \in (s,t)\}} \tag{2.28}$$

$$= \frac{N}{2\pi}(t-s) + \frac{1}{\pi}\mathfrak{Im}\log Z_U(t) - \frac{1}{\pi}\mathfrak{Im}\log Z_U(s), \tag{2.29}$$

and so

$$\widetilde{C}_U(s,t) = \frac{\mathfrak{Im}\log Z_U(t)}{\sqrt{\log N}} - \frac{\mathfrak{Im}\log Z_U(s)}{\sqrt{\log N}}. \tag{2.30}$$

Finally, observe that theorem 2.2 applies, and says that the finite dimensional distributions of $\frac{\mathfrak{Im}\log Z_U(\theta)}{\sqrt{\log N}}$ converges to those of $Y(\theta)$, given in (2.25). (Note that the scaling is $\sqrt{\log N}$ here, not the usual $\sqrt{\frac{1}{2}\log N}$).

2.2.1 The variance of the counting function

The remaining part of this section concerns the calculation of the variance of $C_U(0,\theta)$ for fixed θ and also for θ is the scale of the mean density, and compares these results with the Riemann zeta function. The variance of the counting function can be calculated exactly at finite N, making use of the identity 2.29 in terms of the correlation between $\mathfrak{Im}\log Z_U(\theta)$ and $\mathfrak{Im}\log Z_U(0)$ as follows

$$\operatorname{var}\{C_U(0,\theta)\} = \frac{1}{\pi^2}\mathbb{E}\left\{(\mathfrak{Im}\log Z_U(\theta) - \mathfrak{Im}\log Z_U(0))^2\right\} \tag{2.31}$$

$$= \frac{2}{\pi^2}\mathbb{E}\left\{(\mathfrak{Im}\log Z_U(0))^2\right\} - \frac{2}{\pi^2}\mathbb{E}\left\{\mathfrak{Im}\log Z_U(\theta)\mathfrak{Im}\log Z_U(0)\right\} \tag{2.32}$$

Lemma 2.5.

$$\mathbb{E}\left\{\mathfrak{Im}\log Z_U(\theta)\mathfrak{Im}\log Z_U(0)\right\} = \frac{1}{2}\sum_{n=1}^{N-1}\frac{\cos(n\theta)}{n} + \frac{1}{2}\sum_{n=N}^{\infty}\frac{N\cos(n\theta)}{n^2}$$

$$= -\tfrac{1}{2}\log\left|2\sin(\tfrac{1}{2}\theta)\right| + \tfrac{1}{2}\operatorname{Ci}(N|\theta|) + \tfrac{1}{2}\cos(N\theta) - \tfrac{1}{4}N\pi|\theta| + \tfrac{1}{2}N\theta\operatorname{Si}(N\theta) + \mathcal{O}\left(\frac{1}{N}\right) \tag{2.33}$$

uniformly for $-\pi \le \theta \le \pi$, *where* $\operatorname{Si}(\cdot)$ *is the sine integral and* $\operatorname{Ci}(\cdot)$ *is the cosine integral.*

Proof.[3] Taylor expanding $\mathfrak{Im}\log\left(1 - e^{\mathrm{i}(\theta_n-\theta)}\right)$ (which is valid other than for a

[3]This lemma can also be deduced from work of Rains, [80], or, more recently, of Bump, Diaconis and Keller [21]. A similar statement can be found in [43].

measure zero set) gives

$$\mathfrak{Im} \log Z_U(\theta) = \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{\mathrm{i}}{2k} e^{-\mathrm{i}k\theta} \operatorname{Tr} U^k. \tag{2.34}$$

It is well known (see, e.g., [35, 36, 48, 80]) that

$$\mathbb{E}\left\{\operatorname{Tr} U^m \operatorname{Tr} U^n\right\} = \begin{cases} \min(|n|, N) & \text{if } m = -n \\ 0 & \text{otherwise} \end{cases}. \tag{2.35}$$

(This is essentially just the unscaled form factor, the Fourier transform of Dyson's two-point correlation function).

Thus

$$\mathbb{E}\{\mathfrak{Im} \log Z_U(\theta) \mathfrak{Im} \log Z_U(0)\} = \sum_{\substack{m=-\infty \\ k\neq 0}}^{\infty} \sum_{\substack{n=-\infty \\ k\neq 0}}^{\infty} \frac{-1}{4mn} e^{-\mathrm{i}m\theta} \mathbb{E}\left\{(\operatorname{Tr} U^m)(\operatorname{Tr} U^n)\right\} \tag{2.36}$$

$$= \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{1}{4n^2} e^{-\mathrm{i}n\theta} \min(|n|, N) \tag{2.37}$$

$$= \frac{1}{2} \sum_{n=1}^{N-1} \frac{\cos(n\theta)}{n} + \frac{1}{2} \sum_{n=N}^{\infty} \frac{N \cos(n\theta)}{n^2} \tag{2.38}$$

Using Euler Maclaurin summation, one can show that for $-\pi \leq \theta \leq \pi$,

$$N \sum_{n=N}^{\infty} \frac{\cos(n\theta)}{n^2} = \cos(N\theta) - \tfrac{1}{2} N\pi|\theta| + N\theta \operatorname{Si}(N\theta) + \mathcal{O}\left(\frac{1}{N}\right) \tag{2.39}$$

where

$$\operatorname{Si}(z) = \int_0^z \frac{\sin x}{x} \, \mathrm{d}x \tag{2.40}$$

is the sine integral. Using Euler Maclaurin again,

$$\sum_{n=1}^{N-1} \frac{\cos(n\theta)}{n} = \sum_{n=1}^{\infty} \frac{\cos(n\theta)}{n} - \sum_{n=N}^{\infty} \frac{\cos(n\theta)}{n} \tag{2.41}$$

$$= -\log\left|2\sin(\tfrac{1}{2}\theta)\right| + \operatorname{Ci}(N|\theta|) + \mathcal{O}\left(\frac{1}{N}\right) \tag{2.42}$$

where $\sum_{n=1}^{\infty} \frac{\cos(n\theta)}{n} = -\log\left|2\sin(\frac{1}{2}\theta)\right|$ (it is a Claussen sum) and the cosine integral is

$$\operatorname{Ci}(z) = \gamma + \log z + \int_0^z \frac{\cos x - 1}{x} \, \mathrm{d}x \quad |\arg(z)| < \pi \tag{2.43}$$

$$= -\int_z^{\infty} \frac{\cos x}{x} \, \mathrm{d}x. \tag{2.44}$$

This completes the proof. □

Remark. The Taylor expansion of the sine and cosine integrals are

$$\mathrm{Si}(z) = z + \mathcal{O}(z^3), \tag{2.45}$$

$$\mathrm{Ci}(z) = \log z + \gamma - \tfrac{1}{4}z^2 + \mathcal{O}(z^4). \tag{2.46}$$

and so, letting $\theta \to 0$,

$$\mathbb{E}\left\{\mathfrak{Im} \log Z_U(0)^2\right\} = \tfrac{1}{2}(\log N + 1 + \gamma) + \mathcal{O}\left(\frac{1}{N}\right) \tag{2.47}$$

(which is, of course, well known [2, 69]).

The sine and cosine integrals have an asymptotic expansion as $z \to +\infty$ of

$$\mathrm{Si}(z) \sim \tfrac{1}{2}\pi - \frac{\cos z}{z} - \frac{\sin z}{z^2} + \cdots, \tag{2.48}$$

$$\mathrm{Ci}(z) \sim \frac{\sin z}{z} - \frac{\cos z}{z^2} + \cdots, \tag{2.49}$$

and so the variance of the counting function is, for fixed θ,

$$\mathrm{var}\left\{C_U(0,\theta)\right\} \sim \frac{1}{\pi^2}\left(\log N + 1 + \gamma + \log|2\sin(\tfrac{1}{2}\theta)|\right). \tag{2.50}$$

In fact Rains has gone much further, computing an arbitrarily precise asymptotic expansion for the variance:

Theorem 2.6. (Rains [80]). *For any $d \geq 1$,*

$$\mathrm{var}\left\{C_U(0,\theta)\right\} = \frac{1}{\pi^2}\Bigg[\log N + 1 + \gamma + \log|2\sin(\tfrac{1}{2}\theta) + \\ + \sum_{k=2}^{d} \frac{(1-\frac{1}{k})B_k + \mathfrak{Re}(\mathrm{i}^k e^{\mathrm{i}N\theta})2^{-k}(1-k)\cot^{(k-1)}(\frac{1}{2}\theta)}{N^k}\Bigg] + \mathcal{O}\left(N^{-(d+1)}\right) \tag{2.51}$$

where B_k is the kth Bernoulli number.

If we rescale in the mean density, so $\theta = 2\pi\alpha/N$, then $C_U(0, 2\pi\alpha/N)$ weakly converges to an (as yet unknown) discrete random variable with mean α and variance

$$\mathrm{var}\{C_U(0,2\pi\alpha/N)\} = \frac{1}{\pi^2}\Big(1 + \gamma + \log(2\pi\alpha) - \mathrm{Ci}(2\pi\alpha) - \cos(2\pi\alpha) + \pi^2\alpha \\ - 2\pi\alpha\,\mathrm{Si}(2\pi\alpha)\Big) + \mathcal{O}\left(\frac{1}{N}\right). \tag{2.52}$$

(Again this can be found in [80]).

It is interesting to compare this with the number variance of the Riemann zeros. Recall that $N(T) = \overline{N}(T) + S(T)$, with $S(T)$ given by (1.13). $S(T)$ and $\frac{1}{\pi}\mathfrak{Im}\log Z_U(\theta)$ play analogous roles, being both imaginary parts of the logarithm and the error term in the counting function.

The number variance of the Riemann zeros (in the scale of the mean density) is defined as

$$V_T(\alpha) := \frac{1}{T}\int_0^T \left\{ S\left(t + \frac{2\pi\alpha}{\log(T/2\pi)}\right) - S(t)\right\}^2 \mathrm{d}t. \tag{2.53}$$

Using techniques developed in [44], Fujii [41] proved:

Theorem 2.7. (Fujii [41]). *Assuming RH and Montgomery's conjecture, for* $0 < \alpha \ll \log T$, $V_T(\alpha)$ *equals*

$$\frac{1}{\pi^2}\left\{\log(2\pi\alpha) + \gamma - \mathrm{Ci}(2\pi\alpha) + 1 - \cos(2\pi\alpha) + \pi^2\alpha - 2\pi\alpha\,\mathrm{Si}(2\pi\alpha)\right\} + o(1) \tag{2.54}$$

We see immediately that this result agrees exactly with the random matrix calculation, (2.52), since $N = \log\frac{T}{2\pi}$, so the comparison is between the number variance in the scale of the mean density of zeros.

However, as α grows to be on the order of the mean density, random matrix theory only correctly models the leading order term in the fluctuations in the counting function of Riemann zeros, but not the sub-leading terms. (c.f. the conjecture of Berry, [9]).

Theorem 2.8. (Fujii [42]) *For fixed constants* $A, B > 0$, *if* $A \le h \le BT$ *then, assuming the Riemann Hypothesis,*

$$\frac{1}{T}\int_0^T (S(t+h) - S(t))^2 \mathrm{d}t = V_T\left(\frac{\log(T/2\pi)}{2\pi}h\right) \tag{2.55}$$

$$= \frac{1}{\pi^2}\left(\log\log T - \log|\zeta(1+\mathrm{i}h)|\right) + \mathcal{O}(1) \tag{2.56}$$

For $h = \mathcal{O}(1)$ and $N = \log\frac{T}{2\pi}$, we get leading-order agreement with (2.50), but the sub-leading term involves primes. It is obvious that in this region the RMT–zeta function analogy will break down: In the CUE α cannot grow larger than $2\pi N$ due to periodicity, whereas for the Riemann zeta function, α may grow much larger than the mean density (and in the above theorem it does!).

The right hand side of (2.54) grows like $\frac{1}{\pi^2}\log\alpha$ as $\alpha \to \infty$, which is correct only so long as $\alpha << \log T$ (the mean density) since the right-hand side of (2.56)

fluctuates around $\frac{1}{\pi^2}\log\log T$. (Note that $\log|\zeta(1+\mathrm{i}h)| = \mathcal{O}(\log\log\log h)$ as $h\to\infty$, and since $h < BT$ this is subdominant to the $\log\log T$ term). This phase transition in the growth rate had previously been accounted for in a conjecture of Berry [9]:

Conjecture 2.9. (Berry [9].) *For any $\tau\in(0,1)$ and any $\alpha>0$,*

$$V_T(\alpha) \approx \frac{1}{\pi^2}\left\{\log(2\pi\alpha)+\gamma-\mathrm{Ci}(2\pi\alpha)+1-\cos(2\pi\alpha)+\pi^2\alpha-2\pi\alpha\,\mathrm{Si}(2\pi\alpha)\right\}$$
$$+\frac{1}{\pi^2}\left\{2\sum_{r=1}^{\infty}\sum_{\substack{p\text{ prime}\\ p^r<(T/2\pi)^\tau}}\frac{\sin^2\left(\frac{\pi\alpha r\log p}{\log T/2\pi}\right)}{r^2p^r}+\mathrm{Ci}(2\pi\alpha\tau)-\log(2\pi\alpha\tau)-\gamma\right\} \tag{2.57}$$

In particular for fixed τ, if $0<\alpha\ll\log T$, the second term in braces can be shown to be subdominant to the first term. And if $\log T\ll\alpha\ll T$ then the conjecture for $V_T(\alpha)$ can be shown to be asymptotic to $\frac{1}{\pi^2}\log\log T$.

Noticing that

$$V_T(\alpha) = \alpha + 2\int_0^\alpha(\alpha-x)\left(R_2^{(T)}(x)-1\right)\mathrm{d}x \tag{2.58}$$

where $R_2^{(T)}(x)$ is the scaled two-point correlation function at finite height T, Berry's conjecture has been sharpened slightly by Bogomolny and Keating [15] who heuristically showed that if $L=\frac{1}{2\pi}\log\frac{T}{2\pi}$ then

$$R_2^{(T)}(x) = 1 + R_2^{(\mathrm{d})}(x) + R_2^{(\mathrm{off})}(x) \tag{2.59}$$

with

$$R_2^{(\mathrm{d})}(x) = \frac{1}{2\pi^2L^2}\mathfrak{Re}\left\{\frac{\mathrm{d}^2}{\mathrm{d}s^2}\log\zeta(s) - \sum_{p\text{ prime}}\frac{\log^2 p}{(p^s-1)^2}\bigg|_{s=1+\mathrm{i}x/L}\right\} \tag{2.60}$$

and

$$R_2^{(\mathrm{off})}(x) = \frac{1}{2\pi^2L^2}\left|\zeta(1+\tfrac{\mathrm{i}x}{L})\right|^2\mathfrak{Re}\left\{\exp(2\pi\mathrm{i}x)\prod_p\left(1-\frac{(p^{\mathrm{i}x/L}-1)^2}{(p-1)^2}\right)\right\}. \tag{2.61}$$

(The detailed calculations are given in [68, 10]). Of course, as $T\to\infty$, it can be seen that $R_2^{(T)}(x)\to 1-\frac{\sin^2(\pi x)}{\pi^2x^2}$.

Finally, we compare $\mathbb{E}\{\mathfrak{Im}\log Z_U(0)^2\}$ with

Theorem 2.10. (Goldston [44]). *Assuming RH and Montgomery's conjecture,*

$$\frac{1}{T}\int_0^T S(t)^2\mathrm{d}t = \frac{1}{2\pi^2}\left\{\log\log T+\gamma+1-\sum_{m=2}^{\infty}\sum_p\left(\frac{m-1}{m^2}\right)\frac{1}{p^m}\right\}+o(1) \tag{2.62}$$

(The sign error in [44] is corrected here). This is in agreement with (2.47) only at leading-order. Primes appear in the sub-leading term.

2.3 Functional central limit theorem

Szegő's theorem for real-valued functions on $\mathbb{T}$ states that if

$$A(h) = \sum_{k=1}^{\infty} k|\hat{h}_k|^2 < \infty, \tag{2.63}$$

then

$$D_N[e^h] = \exp\Big(N\hat{h}_0 + A(h) + o(1)\Big) \tag{2.64}$$

as $N \to \infty$. Combining this with Heine's identity (lemma 1.10), we see that if $\hat{h}_0 = 0$ and $A(h) < \infty$, then $\operatorname{Tr} h(U)$ is asymptotically normal with zero mean and variance $2A(h)$. (Many papers have been written about Szegő's theorem, see [18, 63, 64, 88], and the references contained within those papers, for proofs and generalizations).

Now, $\mathfrak{Re} \log Z_U(\theta) = \operatorname{Tr} h(U)$, where $h\left(e^{\mathrm{i}t}\right) = \mathfrak{Re} \log(1 - e^{\mathrm{i}(t-\theta)})$, but the Fourier coefficients $\hat{h}_k$ are of order $1/k$ in this case and $A(h) = +\infty$, so we cannot use Szegő's theorem to prove results such as theorem 2.2, for example. (Indeed, the proof of lemma 2.1 required knowledge of how Toeplitz determinants of Fisher-Hartwig symbols behave as $N \to \infty$, and Fisher-Hartwig symbols are such that Szegő's theorem does not apply to them).

However, Szegő's theorem can be applied to obtain a functional central limit theorem for $\log Z_U(\theta)$. That means there exists a random function $F(\theta)$ and a metric space of functions H_0^a, (both identified below) such that

$$\lim_{N\to\infty} \mathbb{E}\{q(\log Z_U(\theta))\} = E\{q(F(\theta))\} \tag{2.65}$$

for all bounded continuous functions $q : H_0^a \longrightarrow \mathbb{R}$ (so $q(\cdot)$ is a function of a function).

A generalized real-valued function $f(\theta)$ on $\mathbb{T}$ can be identified uniquely with the infinite sequence of its Fourier coefficients $(\hat{f}_0, \hat{f}_1, \dots)$. (Note that for $f(\cdot)$ to be real, $\hat{f}_{-k} = \hat{f}_k^*$). Denote by H_0^a the space of generalized real-valued functions $f(\theta)$ on $\mathbb{T}$ with $\hat{f}_0 = 0$ and $\|f(\theta)\|_a < \infty$, where $\|f(\theta)\|_a^2 = \langle f(\theta), f(\theta)\rangle_a$, the inner product being

$$\langle f(\theta), g(\theta)\rangle_a = \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} |k|^{2a} \hat{f}_k \hat{g}_k^* \tag{2.66}$$

$$= 2\mathfrak{Re} \sum_{k=1}^{\infty} k^{2a} \hat{f}_k \hat{g}_k^*. \tag{2.67}$$

Let $R_1, R_2, \ldots$ be a sequence of i.i.d. standard complex normal random variables, and $R_{-k} = R_k^*$, and define

$$F(\theta) = \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \left(\frac{R_k}{2\sqrt{2|k|}} \right) e^{ik\theta}. \tag{2.68}$$

Note that

$$\mathfrak{Re} \log Z_U(\theta) = \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{-\operatorname{Tr} U^{-k}}{2|k|} e^{\mathrm{i}k\theta}. \tag{2.69}$$

Lemma 2.11. *The finite dimensional distributions of $\mathfrak{Re} \log Z_U(\theta)$ converge weakly as $N \to \infty$ to the finite dimensional distribution of $F(\theta)$ in H_0^a.*

Proof. This lemma follows from theorem 1.5 and Szegő's theorem. It amounts to proving that for any $m > 0$,

$$\left(\frac{-\operatorname{Tr} U^{-1}}{2}, \ldots, \frac{-\operatorname{Tr} U^{-m}}{2m} \right) \Longrightarrow \left(\frac{R_1}{2\sqrt{2}}, \ldots, \frac{R_m}{2\sqrt{2m}} \right) \tag{2.70}$$

Let $f(\theta)$ be such that $\hat{f}_k = 0$ for $|k| > m$ and $\hat{f}_0 = 0$. (So $f(\theta) \in H_0^a$ for any a).

Let E denote expectation over the complex normal random variables. Writing $R_k = X_k + \mathrm{i}Y_K$ with X_k and Y_k iid standard normal random variables for $k = 1, 2, \ldots$, then

$$\langle f(\theta), F(\theta) \rangle_a = \sum_{\substack{k=-m \\ k\neq 0}}^{m} |k|^{2a} \hat{f}_k \frac{R_k^*}{2\sqrt{2}|k|} \tag{2.71}$$

$$= \frac{1}{\sqrt{2}} \sum_{k=1}^{m} \left[k^{2a-1/2} \mathfrak{Re}(\hat{f}_k) X_k + k^{2a-1/2} \mathfrak{Im}(\hat{f}_k) Y_k \right] \tag{2.72}$$

so, by the independence of X_k and Y_k,

$$E \exp\left(\langle f(\theta), F(\theta) \rangle_a \right) = \prod_{k=1}^{m} e^{k^{4a-1}(\mathfrak{Re}\hat{f}_k)^2/4} e^{k^{4a-1}(\mathfrak{Im}\hat{f}_k)^2/4} \tag{2.73}$$

$$= \exp\left(\tfrac{1}{8} \langle f(\theta), f(\theta) \rangle_{2a-1/2} \right). \tag{2.74}$$

Next, note that

$$\langle f(\theta), \mathfrak{Re} \log Z_U(\theta) \rangle_a = \frac{-1}{2} \sum_{\substack{k=-m \\ k\neq 0}}^{m} |k|^{2a-1} \hat{f}_k \operatorname{Tr} U^k \tag{2.75}$$

$$= \sum_{n=1}^{N} \frac{-1}{2} \sum_{\substack{k=-m \\ k\neq 0}}^{m} |k|^{2a-1} \hat{f}_k e^{\mathrm{i}k\theta_n} \tag{2.76}$$

so Heine's identity gives that

$$\mathbb{E}\exp\left(\langle f(\theta), \mathfrak{Re}\log Z_U(\theta)\rangle_a\right) = D_N[e^h] \tag{2.77}$$

with $\hat{h}_k = \frac{-1}{2}|k|^{2a-1}\hat{f}_k$, and thus Szegő's theorem implies

$$\mathbb{E}\exp\left(\langle f(\theta), \mathfrak{Re}\log Z_U(\theta)\rangle_a\right) \to \exp\left(\sum_{k=1}^{m} k\tfrac{1}{4}k^{4a-2}|\hat{f}_k|^2\right) \tag{2.78}$$

$$= \exp\left(\tfrac{1}{8}\langle f(\theta), f(\theta)\rangle_{2a-1/2}\right) \tag{2.79}$$

which equals (2.74), so (2.70) is true by theorem 1.5, and this proves the lemma. □

Remark. In the case $a = 0$, Gangardt [43] has essentially reproven this lemma using a technique from theoretical physics called bosonization.

Remark. If $a < 1/2$ then (2.74) remains finite as $m \to \infty$. However, there is no guarantee that the conclusion of theorem 1.5 remains true in that case, so one cannot conclude that $\mathfrak{Re}\log Z_U(\theta) \Longrightarrow F(\theta)$. The following theorem shows the conclusion is valid for $a < 0$.

Theorem 2.12. *For fixed $a < 0$,*

$$\mathfrak{Re}\log Z_U(\theta) \Longrightarrow F(\theta) \tag{2.80}$$

in H_0^a.

Proof. Prohorov's theorem, given in section 6 of [11], says that if the finite dimensional distributions of X_N converge weakly to the finite dimensional distribution of X, and $\{X_N\}$ is tight (which means for for every $\epsilon > 0$ there exists a compact set K such that $\mathbb{P}\{X_N \in K\} > 1 - \epsilon$ for all N) then X_N converges weakly to X.

Lemma (2.11) shows that the finite dimensional distributions of $\mathfrak{Re}\log Z_U(\theta)$ converge to those of $F(\theta)$. Thus, all that is required to prove this theorem is to show that $\{\mathfrak{Re}\log Z_U(\theta)\}$ is tight in H_0^a.

First, using (2.35),

$$\mathbb{E}\,\|\mathfrak{Re}\log Z_U(\theta)\|_a^2 = \frac{1}{2}\sum_{k=1}^{\infty} k^{2a-2}\,\mathbb{E}\,|\operatorname{Tr} U^{-k}|^2 \tag{2.81}$$

$$= \frac{1}{2}\sum_{k=1}^{\infty} k^{2a-2}\min(k,N) \tag{2.82}$$

$$\leq \frac{1}{2}\sum_{k=1}^{\infty} k^{2a-1} \tag{2.83}$$

Note that (2.82) is finite for $a < 1/2$, and that it diverges as $N \to \infty$ for $0 \leq a < 1/2$. It is only for $a < 0$ that the uniform (in N) bound

$$\mathbb{E}\,\|\mathfrak{Re}\log Z_U(\theta)\|_a^2 \leq \tfrac{1}{2}\zeta(1-2a) \tag{2.84}$$

holds.

Second, for $a < 0$ and for each $\epsilon > 0$, choose

$$x > \sqrt{\frac{\zeta(1-2a)}{2\epsilon}} \tag{2.85}$$

and set

$$K = \{f \in H_0^a \;:\; \|f(\theta)\|_a \leq x\}\,. \tag{2.86}$$

Note that for all $N \geq 1$,

$$\mathbb{P}\,\{\mathfrak{Re}\log Z_U(\theta) \in K\} = 1 - \mathbb{P}\,\{\|\mathfrak{Re}\log Z_U(\theta)\|_a > x\} \tag{2.87}$$

$$\geq 1 - \frac{\mathbb{E}\,\|\mathfrak{Re}\log Z\|_a^2}{x^2} \tag{2.88}$$

by Chebyshev's inequality. But by (2.84), and the choice of x,

$$\mathbb{P}\,\{\mathfrak{Re}\log Z_U(\theta) \in K\} \geq 1 - \frac{\zeta(1-2a)}{2x^2} \tag{2.89}$$

$$\geq 1 - \epsilon \tag{2.90}$$

for all N. This proves tightness of the sequence $\{\mathfrak{Re}\log Z_U(\theta)\}$, which proves the theorem. □

Remark. Let $\mathcal{H}$ be the Hilbert transform:

$$\widehat{\mathcal{H}f}_k = \begin{cases} i\hat{f}_k & k > 0 \\ -i\hat{f}_k & k < 0 \end{cases}. \tag{2.91}$$

Since $\mathfrak{Im} \log Z_U(\theta) = \mathcal{H}(\mathfrak{Re} \log Z_U(\theta))$, then if $\mathfrak{Re} \log Z_U(\theta) \Longrightarrow F(\theta)$ one immediately obtains as a corollary

$$(\mathfrak{Re} \log Z_U(\theta), \mathfrak{Im} \log Z_U(\theta)) \Longrightarrow (F(\theta), \mathcal{H}(F(\theta))) \tag{2.92}$$

in the same topology.

Chapter 3

Large Deviations

In this chapter we obtain a range of large and moderate deviations results for $\log Z_U(\theta)$.

The central limit theorem for a finite set of distinct points (theorem 2.2) occurs when $\log Z_U(\theta)$ is scaled by $A = \sqrt{\log N}$. Deviations from the central limit theorem are found to have a quadratic rate function (as is anticipated from a normal distribution) for scalings $\sqrt{\log N} \ll A \ll \log N$.

If the real and imaginary parts are considered separately and at just one point (say $\theta = 0$) then it is possible to go much further. $\mathfrak{Im} \log Z_U(0)/A$ has a quadratic rate function for $\sqrt{\log N} \ll A \ll N$, but not at $A = N$. When $A = N$, an explicit expression for the rate function, which occurs at speed $B = N^2$, is given. Large deviations of $\mathfrak{Re} \log Z_U(0)/A$ to the right have a quadratic rate function for $\sqrt{\log N} \ll A \ll N$, but not at $A = N$. An explicit formula for the rate function when $A = N$, which again occurs at $B = N^2$, is given. But large deviations of $\mathfrak{Re} \log Z_U(0)/A$ to the left are quadratic only for $\sqrt{\log N} \ll A \leq \log N$. For $A \geq \log N$ they become linear (and occur at speed $B = A$), the phase transition being caused by one eigenangle coming too close to 0.

Large deviation estimates for the real and imaginary parts in certain scalings can be exponentially improved by estimating the probability density using the saddle-point method. The deviations of $\mathfrak{Re} \log Z_U(0)$ to the left can be similarly improved (and indeed calculated exactly) using calculus of residues, the linear rate function expressing itself as a simple exponential decay in the left tail of the probability density.

Similar calculations applied to the Riemann zeta function lead to a conjecture on the probability of $\log|\zeta(1/2+\mathrm{i}t)|$ being very large and negative. The difficulty of using random matrix theory to conjecture the true rate of growth of $|\zeta(1/2+\mathrm{i}t)|$ is also discussed.

Inside the unit circle, $\log Z_U(\theta)$ becomes a continuous function, and the LDP at scaling $A = N$ and speed $B = N^2$ can be calculated using a result due to Hiai and Petz [58]. Their result also enables us to write down the most likely distribution of eigenangles, given that such a large deviation has occurred.

The different scalings for left and right deviations of $\mathfrak{Re}\log Z_U(\theta)$ makes itself manifest in the moments: the fixed moments are Gaussian (as $N \to \infty$), as they must be for the central limit theorem to hold. But the high moments are far from Gaussian, being dominated by the large left tail of the distribution of $\mathfrak{Re}\log Z_U(\theta)$.

Some of the results in this chapter have previously appeared in [60].

3.1 Large deviations for $\log Z_U(\theta)$ evaluated at distinct points

Theorem 3.1. *For $\sqrt{\log N} \ll A \ll \log N$, and for any distinct $r_1, \ldots, r_k \in \mathbb{T}$, the sequence*

$$\left(\frac{\mathfrak{Re}\log Z_U(r_1)}{A}, \frac{\mathfrak{Im}\log Z_U(r_1)}{A}, \ldots, \frac{\mathfrak{Re}\log Z_U(r_k)}{A}, \frac{\mathfrak{Im}\log Z_U(r_k)}{A}\right) \tag{3.1}$$

satisfies the LDP in $\mathbb{R}^{2k}$ with speed $B = A^2/\log N$ and rate function

$$I(x_1, y_1 \ldots, x_k, y_k) = \sum_{j=1}^{k} x_j^2 + y_j^2. \tag{3.2}$$

Proof. By lemma 2.1, if $B/A \ll 1$,

$$\log \mathbb{E}\exp\left(\sum_{j=1}^{k} s_j \mathfrak{Re}\log Z_U(r_j)B/A + t_j \mathfrak{Im}\log Z_U(r_j)B/A\right) \sim \left(\sum_{j=1}^{k}(s_j^2+t_j^2)/4\right)\frac{B^2\log N}{A^2} \tag{3.3}$$

so choosing the speed $B = \frac{A^2}{\log N}$, we see that $\Lambda(s_1, t_1, \ldots, s_k, t_k) = \frac{1}{4}\sum_{j=1}^{k}(s_j^2+t_j^2)$, and by theorem 1.6 (with $d = 2k$), the rate function is the convex-dual (Fenchel-Legendre transform) of this. □

Remark. Closer inspection of the proof of lemma 2.1 actually allows $A = \log N$ as a scaling in theorem 3.1, so long as $s_j > -1$ for all j (which means $x_j > -1/2$).

Large deviations results for the counting function can also be deduced, using the identity (2.29).

Theorem 3.2. *For $\sqrt{\log N} \ll A \ll \log N$, and $-\pi < s < t \leq \pi$, the sequence $(C_U(s,t) - (t-s)N/2\pi)/A$ satisfies the LDP in $\mathbb{R}$ with speed $B = A^2/\log N$ and rate function $L(x) = \pi^2 x^2/2$.*

Proof. By (2.29),

$$\frac{C_U(s,t) - (t-s)\frac{N}{2\pi}}{A} = \frac{\mathfrak{Im} \log Z_U(t) - \mathfrak{Im} \log Z_U(s)}{\pi A} \tag{3.4}$$

and by lemma 2.1, if $B = A^2/\log N$,

$$\lim_{N\to\infty} \log \mathbb{E} \exp\left(\lambda B \left(\frac{\mathfrak{Im} \log Z_U(t) - \mathfrak{Im} \log Z_U(s)}{\pi A}\right)\right) = \frac{\lambda^2}{2\pi^2} \tag{3.5}$$

so long as s, t are distinct in $\mathbb{T}$. The result now follows from theorem 1.6 with $d = 1$.

□

If one specializes to evaluating $\log Z_U(\theta)$ at only one point[1], then it is possible to go further, due to the existence of an exact moment generating function (1.73), rather than just an asymptotic one, (lemma 2.1). For simplicity (rather than necessity) we will consider the real and imaginary parts separately.

3.2 Large deviations for $\mathfrak{Im} \log Z_U(0)$

Theorem 3.3. *For scalings $\sqrt{\log N} \ll A \ll N$, $\mathfrak{Im} \log Z_U(0)/A$ satisfies the LDP with speed $B = -A^2/W_{-1}(-A/N)$ and rate function $J(y) = y^2$. Here W_{-1} is the -1–branch of the Lambert–W function, defined in appendix B.*

Proof. For a given scaling sequence $A(N)$ we wish to find $B(N)$ such that

$$\lim_{N\to\infty} \frac{1}{B} \log L_N(-\mathrm{i}tB/A) \tag{3.6}$$

exists as a non-trivial pointwise limit, where L_N is given by (1.80). Applying results from appendix D,

$$\log L_N(-\mathrm{i}tN/\chi) = \tfrac{1}{4} t^2 N^2 \frac{\log \chi}{\chi^2} + \mathcal{O}_t\left(\frac{N^2}{\chi^2}\right) \tag{3.7}$$

[1]Since the law of $\log Z_U(\theta)$ is independent of θ, we will set $\theta = 0$ without loss of generality.

for all fixed $t \in \mathbb{R}$, so long as $\chi \gg 1$.

Thus, putting $\chi = NA/B$, it is seen that a non-trivial limit of (3.6) occurs if $B = N^2 \log \chi / \chi^2$, that is, if

$$B = \frac{A^2}{-W_{-1}\left(-\frac{A}{N}\right)}, \tag{3.8}$$

where W_{-1} is the -1–branch of the Lambert W–function, described in appendix B. (We need to be on the -1–branch in order for $B \sim 1$ when $A = \sqrt{\log N}$, which must be the case due to the central limit theorem). Note that the restriction $\chi \gg 1$ implies $A \ll N$.

With this B, theorem 1.6 implies the rate function is

$$J(y) = \sup_{t \in \mathbb{R}} \left\{ ty - \tfrac{1}{4}t^2 \right\} \tag{3.9}$$

$$= y^2 \tag{3.10}$$

as required. □

Remark. If $\sqrt{\log N} \ll A \ll \log N$, then $B \sim \frac{A^2}{\log N}$, as in theorem 3.1.

Theorem 3.3 fails to cover the case $A = N$ (when $\chi = 1$). It is found that the rate function ceases to be quadratic there:

Theorem 3.4. *The sequence $\mathfrak{Im} \log Z_U(0)/N$ satisfies the LDP with speed N^2 and rate function given by the convex dual (Fenchel-Legendre transform) of*

$$\Lambda(t) = \tfrac{1}{8}t^2 \log\left(1 + \frac{4}{t^2}\right) - \tfrac{1}{2}\log\left(1 + \tfrac{1}{4}t^2\right) + t \arctan\left(\tfrac{1}{2}t\right) \tag{3.11}$$

Proof. $\log \mathbb{E} \exp(tN\mathfrak{Im} \log Z_U(0)) = \log L_N(-\mathrm{i}Nt)$, and the asymptotics (given in appendix D) imply that

$$\Lambda(t) = \lim_{N \to \infty} \frac{1}{N^2} \log L_N(-\mathrm{i}Nt) \tag{3.12}$$

$$= \tfrac{1}{8}t^2 \log\left(1 + \frac{4}{t^2}\right) - \tfrac{1}{2}\log\left(1 + \tfrac{1}{4}t^2\right) + t \arctan\left(\tfrac{1}{2}t\right) \tag{3.13}$$

Theorem 1.6 implies that $\mathfrak{Im} \log Z_U(0)/N$ satisfies an LDP at speed N^2 with rate function

$$J(y) := \sup_{t \in \mathbb{R}} \left\{ ty - \tfrac{1}{8}t^2 \log\left(1 + \frac{4}{t^2}\right) + \tfrac{1}{2}\log\left(1 + \tfrac{1}{4}t^2\right) - t \arctan\left(\tfrac{1}{2}t\right) \right\} \tag{3.14}$$

for all $y \in \mathbb{R}$. □

Remark. $J(y) = \infty$ for $|y| \geq \frac{1}{2}\pi$. In fact, since $|\mathfrak{Im} \log Z_U(0)| \leq N\pi/2$, the scaling $A = N$ is the maximal non-trivial scaling.

3.3 Large deviations for $\mathfrak{Re}\log Z_U(0)$

In this section we will use theorem 1.7 to calculate the LDPs for $\mathfrak{Re}\log Z_U(0)$. However, we will find that the conditions for that theorem fail at certain points, and to overcome those restrictions we need the following result:

Theorem 3.5. *For any $A(N) \gg \log N$, and $a < b < 0$,*

$$\lim_{N\to\infty} \frac{1}{A} \log \mathbb{P}\left\{ \frac{\mathfrak{Re}\log Z_U(0)}{A} \in (a,b) \right\} = b \tag{3.15}$$

Also, for any $a < b < -1/2$,

$$\lim_{N\to\infty} \frac{1}{\log N} \log \mathbb{P}\left\{ \frac{\mathfrak{Re}\log Z_U(0)}{\log N} \in (a,b) \right\} = b + 1/4 \tag{3.16}$$

Rephrasing this: For scalings $A \gg \log N$, $\mathfrak{Re}\log Z_U(0)$ satisfies a partial LDP at speed A with rate function $I(x) = |x|$ for $x < 0$. And if $A = \log N$, $\mathfrak{Re}\log Z_U(0)$ satisfies a partial LDP at speed $\log N$ with rate function $I(x) = |x| - 1/4$ for $x < -1/2$.

Proof. If $\limsup_{N\to\infty} x/\log N < -1/2$, then

$$p(x) \sim e^x \exp\left(3\zeta'(-1) + \tfrac{1}{12}\log 2 - \tfrac{1}{2}\log\pi\right) N^{1/4}. \tag{3.17}$$

where $p(x)$ is the probability density function of $\mathfrak{Re}\log Z_U(0)$. This will be proved in theorem 3.9.

Therefore, for $a < b < -1/2$,

$$\begin{aligned} \mathbb{P}\left\{ \frac{\mathfrak{Re}\log Z_U(0)}{\log N} \in (a,b) \right\} &= \int_{a\log N}^{b\log N} p(x)\,\mathrm{d}x \\ &\sim \frac{1}{\log N} \exp\left(3\zeta'(-1) + \tfrac{1}{12}\log 2 - \tfrac{1}{2}\log\pi\right) N^{1/4} \left(N^b - N^a\right) \end{aligned} \tag{3.18}$$

and the result follows from taking logarithms of both sides. Similarly for $A(N) \gg \log N$ with $a < b < 0$. □

These are partial LDPs (partial because they are valid only for left deviations) occurring at speed $B = A$ for $A \geq \log N$. The next two theorems show that the deviations to the right occur at a different speed for $A \gg \log N$. It is easy to see that left and right deviations must have, at some point, very different behaviour: We have shown that for any A there exists a non-trivial LDP to the left, whereas,

since $\mathfrak{Re}\log Z_U(0) \leq N\log 2$, the scaling $A = N$ must be the maximal non-trivial scaling for deviations to the right. (This difference is discussed in more detail in section 3.3.1.)

Theorem 3.6. *For scalings $\sqrt{\log N} \ll A \ll N$, the sequence $\mathfrak{Re}\log Z_U(0)/A$ satisfies the LDP with speed $B = -A^2/W_{-1}(-A/N)$ and rate function given by*

$$I(x) = \begin{cases} x^2 & \text{if } \sqrt{\log N} \ll A \ll \log N \\ \begin{cases} x^2 & x \geq -1/2 \\ -x - 1/4 & x < -1/2 \end{cases} & \text{if } A = \log N \\ \begin{cases} x^2 & x \geq 0 \\ 0 & x < 0 \end{cases} & \text{if } \log N \ll A \ll N \end{cases} \tag{3.19}$$

Proof. For a given scaling sequence $A(N)$ we wish to find $B(N)$ such that

$$\Lambda(s) = \lim_{N\to\infty} \frac{1}{B}\log \mathbb{E}\exp\left(\frac{sB\mathfrak{Re}\log Z_U(0)}{A}\right) \tag{3.20}$$

$$= \begin{cases} \lim_{N\to\infty} \frac{1}{B}\log M_N(sB/A) & \text{if } \liminf_{N\to\infty} \frac{sB}{A} > -1 \\ \infty & \text{otherwise} \end{cases} \tag{3.21}$$

exists as a non-trivial pointwise limit.

For $\chi(N) \gg 1$ as $N \to \infty$, we have for each s,

$$\frac{1}{B}\log \mathbb{E}\exp(sN\mathfrak{Re}\log Z_U(0)/\chi) = \begin{cases} \frac{1}{4}s^2\frac{N^2\log\chi}{B\chi^2} + \mathcal{O}_s\left(\frac{N^2}{\chi^2}\right) & \text{if } Ns/\chi > -1 \\ \infty & \text{if } Ns/\chi \leq -1 \end{cases} \tag{3.22}$$

which follows from results summarized in appendix C.

So, as in the proof of theorem 3.3, we need B to be as in (3.8) (which will be valid for $\sqrt{\log N} \ll A \ll N$).

If we set $\delta = \liminf_{N\to\infty} \frac{\chi}{N}$, then we have

$$\Lambda(s) = \begin{cases} \frac{1}{4}s^2 & \text{for } s > -\delta \\ \infty & \text{for } s < -\delta \end{cases} \tag{3.23}$$

If $\sqrt{\log N} \ll A \ll \log N$ then $\delta = +\infty$ and theorem 1.6 implies that $I(x) = x^2$ for all $x \in \mathbb{R}$.

If $A = \log N$, then $\delta = 1/2$, and theorem 1.7 applies only for $x > -1/2$, where we have $I(x) = x^2$. However, since $B \sim \log N$ at this scaling, theorem 3.5 implies that, for $x < -1/2$, $I(x) = |x| - 1/4$.

Finally, if $\log N \ll A \ll N$, then $\delta = 0$, and $I(x) = x^2$ for $x > 0$ by theorem 1.7 and $I(x) = 0$ for $x < 0$ by theorem 3.5 (since $B \gg A$ for $A \gg \log N$).

This completes the proof of theorem 3.6. □

The restriction $\chi \gg 1$ means that theorem 3.6 does not cover the case $A = N$. (The deviations to the left are dealt with by theorem 3.5, where they occur at speed $B = N$). The following theorem will deal with deviations to the right, which occur at speed N^2.

Theorem 3.7. *The sequence $\mathfrak{Re}\log Z_U(0)/N$ satisfies the LDP with speed N^2 and rate function given by the convex dual of*

$$\Lambda(s) = \begin{cases} \frac{1}{2}(1+s)^2\log(1+s) - \left(1+\frac{1}{2}s\right)^2\log\left(1+\frac{1}{2}s\right) - \frac{1}{4}s^2\log 2s & \textit{for } s \geq 0 \\ \infty & \textit{for } s < 0 \end{cases} \tag{3.24}$$

Proof.

$$\log \mathbb{E}\exp(sN\mathfrak{Re}\log Z_U(0)) = \begin{cases} \log M_N(Ns) & \text{for } Ns > -1 \\ \infty & \text{otherwise} \end{cases} \tag{3.25}$$

(the asymptotics of $M_N(Ns)$ are given in appendix C), and so for $s \geq 0$,

$$\Lambda(s) = \lim_{N\to\infty} \frac{1}{N^2}\log M_N(Ns) \tag{3.26}$$

$$= \tfrac{1}{2}(1+s)^2\log(1+s) - \left(1+\tfrac{1}{2}s\right)^2\log\left(1+\tfrac{1}{2}s\right) - \tfrac{1}{4}s^2\log 2s, \tag{3.27}$$

and $\Lambda(s) = \infty$ for $s < 0$.

If $x > 0$, then theorem 1.7 implies that the rate function, $I(x)$, is given by the convex dual of $\Lambda(s)$. If $x < 0$, then theorem 3.5 implies that $I(x) = 0$. Thus for $x \in \mathbb{R}$, $I(x)$ is given by the convex dual of $\Lambda(s)$, and this completes the proof of theorem 3.7. □

Remark. $I(x) = \infty$ for $x \geq \log 2$. In fact, $\mathfrak{Re}\log Z_U(0) \leq N\log 2$ for all matrices.

Remark. For all $\sqrt{\log N} \ll A \leq N$ even thought the conditions of theorem 1.7 are not satisfied at some points, it turns out that $I(x)$ is still the convex dual of $\Lambda(s)$.

Scaling $A(N)$	Speed $B(N)$	Rate function $I(x)$
$\sqrt{\log N} \ll A \ll \log N$	$\frac{A^2}{\log N}$	x^2
$A = \log N$	$\log N$	$\begin{cases} x^2 & \text{if } x \geq -1/2 \\ \lvert x\rvert - 1/4 & \text{if } x \leq -1/2 \end{cases}$
$\log N \ll A \ll N$	$\frac{-A^2}{W_{-1}(-A/N)}$	$\begin{cases} x^2 & \text{if } x \geq 0 \\ 0 & \text{if } x \leq 0 \end{cases}$
	A	$\begin{cases} \infty & \text{if } x > 0 \\ \lvert x\rvert & \text{if } x \leq 0 \end{cases}$
$A = N$	N^2	$\begin{cases} \infty & \text{if } x \geq \log 2 \\ I_c(x) & \text{if } 0 \leq x \leq \log 2 \\ 0 & \text{if } x \leq 0 \end{cases}$
	N	$\begin{cases} \infty & \text{if } x > 0 \\ \lvert x\rvert & \text{if } x \leq 0 \end{cases}$
$A \gg N$	A	$\begin{cases} \infty & \text{if } x > 0 \\ \lvert x\rvert & \text{if } x \leq 0 \end{cases}$

where, for $0 \leq x \leq \log 2$,

$$I_c(x) = \sup_{\lambda > 0} \Big\{ \lambda x - \tfrac{1}{2}(1+\lambda)^2 \log(1+\lambda) + (1+\tfrac{1}{2}\lambda)^2 \log(1+\tfrac{1}{2}\lambda) + \tfrac{1}{4}\lambda^2 \log 2\lambda \Big\}$$

Table 3.1: Speed and non-trivial rate function for $\mathfrak{Re} \log Z_U(0)$, at each possible scaling

3.3.1 The phase transition

There is a phase transition in theorem 3.6, where the rate function of $\frac{\log Z_U(0)}{\log N}$ changes from being quadratic to linear. In this section we will argue that the linear part is caused by an eigenangle coming close to 0.

If $x < 0$, then theorem 3.6 says that if $A \gg \log N$,

$$\frac{1}{A} \log \mathbb{P}\{\mathfrak{Re} \log Z_U(0) < xA\} \sim -|x|. \tag{3.28}$$

For any $\epsilon > 0$, the lower bound

$$\mathbb{P}(\mathfrak{Re} \log Z_U(0) < xA) \geq \mathbb{P}\left\{\log|1 - e^{\mathrm{i}\theta_1}| < (x-\epsilon)A \ , \ \sum_{n=2}^{N} \log|1 - e^{\mathrm{i}\theta_n}| < \epsilon A\right\} \tag{3.29}$$

holds.

Note that for $A \gg \sqrt{\log N}$

$$\mathbb{P}\left\{\sum_{n=2}^{N} \log|1 - e^{\mathrm{i}\theta_n}| < \epsilon A\right\} \to 1, \tag{3.30}$$

(this follows from the central limit theorem) and that θ_1 is uniformly distributed on $\mathbb{T}$. Therefore, assuming the two events on the right hand side of (3.29) are approximately independent, we have

$$\begin{aligned}\liminf_{N\to\infty} \frac{1}{A} \log \mathbb{P}\{\mathfrak{Re} \log Z_U(0) < xA\} &\geq \liminf_{N\to\infty} \frac{1}{A} \log \mathbb{P}\left\{\log|1 - e^{\mathrm{i}\theta_1}| < (x-\epsilon)A\right\} \\ &= -(|x| + \epsilon), \end{aligned} \tag{3.31}$$

and since ϵ is arbitrary, we obtain

$$\liminf_{N\to\infty} \frac{1}{A} \log \mathbb{P}\{\mathfrak{Re} \log Z_U(0) < xA\} \geq -|x|. \tag{3.32}$$

For $A \gg \log N$ this is the correct answer, and so we conclude that the linear rate function for large deviations to the left is most likely caused by just one eigenangle coming too close to 0.

Observe that if $\sqrt{\log N} \ll A \ll \log N$, the speed is $B = A^2 \log N \ll A$, and so all our estimate achieves is the trivial bound

$$\liminf_{N\to\infty} \frac{1}{B} \log \mathbb{P}\{\mathfrak{Re} \log Z_U(0) < xA\} \geq -\infty. \tag{3.33}$$

From this we conclude that for moderate deviations (in the scaling $\sqrt{\log N} \ll A \ll \log N$) are not caused by just one eigenangle, but by a conspiracy among many eigenangles.

3.4 Refined large deviations asymptotics for $\mathfrak{Im} \log Z_U(0)$

The probability density of $\mathfrak{Im} \log Z_U(0)$ is given by

$$q(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\mathrm{i}yx} L_N(y) \, \mathrm{d}y. \tag{3.34}$$

This integral will be calculated using the saddle-point method, which involves deforming the contour of integration to pass through the point where the derivative of $-\mathrm{i}yx + \log L_N(y)$ vanishes.

Making the ansatz that such a saddle-point will occur for $|y_s| \ll 1$, then (using Appendix D on the asymptotics of $\log L_N(y)$, and noting that in this case it is legitimate to differentiate the error terms)

$$y_s = \frac{-2\mathrm{i}x}{\log N + 1 + \gamma} + \mathcal{O}\left(\frac{x^3}{(\log N)^3}\right). \tag{3.35}$$

The assumption on y_s will be valid so long as $|x| \ll \log N$. Taylor expanding around $y = y_s + \epsilon$,

$$\begin{aligned} &-\mathrm{i}yx + \log L_N(\mathrm{i}y) = \\ &\frac{-x^2}{\log N + 1 + \gamma} - \frac{1}{4}(\log N + 1 + \gamma)\epsilon^2 + \mathcal{O}(\epsilon^4) + \mathcal{O}\left(\frac{x^2}{(\log N)^2}\right) + \mathcal{O}\left(\frac{1}{N}\right) \end{aligned} \tag{3.36}$$

where the constant in the $\mathcal{O}(\epsilon^4)$ term can be made independent of N. Therefore,

$$q(x) \sim \frac{1}{2\pi} \int_{-\epsilon_\infty}^{\epsilon_\infty} \exp\left(-\mathrm{i}(y_s + \epsilon)x + L_N(y_s + \epsilon)\right) \, \mathrm{d}\epsilon \tag{3.37}$$

where ϵ_∞ is chosen so that it is much bigger than the scale of the standard deviation, but small enough so that the error terms don't contribute. Thus, for $1/\sqrt{\log N} \ll \epsilon_\infty \ll 1$ we have

$$q(x) \sim \frac{1}{\sqrt{\pi(\log N + 1 + \gamma)}} \exp\left(\frac{-x^2}{\log N + 1 + \gamma}\right) \tag{3.38}$$

which is the density function of a normal random variable with mean zero and variance $\frac{1}{2}(\log N + 1 + \gamma)$.

Under the assumption that $1 \ll |y_s| \ll \sqrt{N}$, then appendix D gives that the saddle point is the solution of

$$\mathrm{i}x + \tfrac{1}{2} y_s(\log N - \log(\mathrm{i}y_s) + \log 2 + 1) + \mathcal{O}\left(\frac{1}{y_s}\right) + \mathcal{O}\left(\frac{y_s^4}{N^2}\right) = 0, \tag{3.39}$$

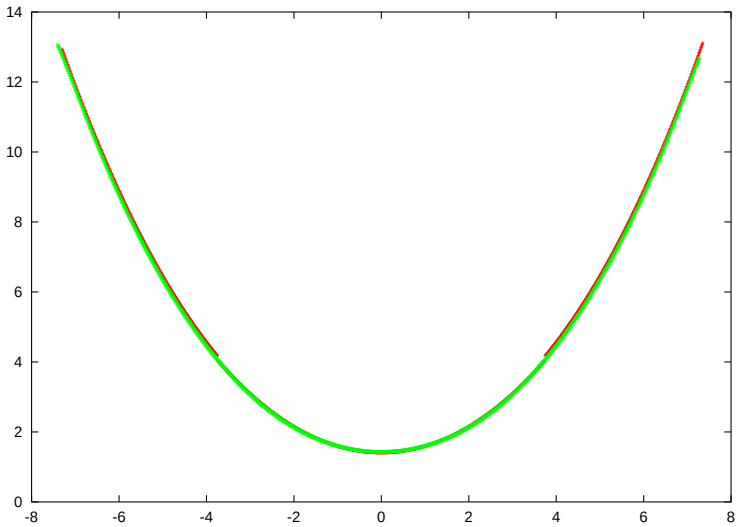

Figure 3.1: Graph of the negative log of the probability density of $\mathfrak{Im} \log Z_U(0)$ (from N. Snaith) (green) against its leading order asymptotics (red), for $N = 42$.

the solution being

$$y_s = \frac{2\mathrm{i}x}{W} + \mathcal{O}\left(\frac{W}{x}\right) + \mathcal{O}\left(\frac{x^4}{N^2W^4}\right) \tag{3.40}$$

where $W = W_{-1}\left(\frac{-x}{eN}\right)$ is the -1 branch of the Lambert W–function (see appendix B). Our ansatz on y_s is valid so long as $\log N \ll x \ll \sqrt{N}$. Taylor expanding about this point, we obtain for $y = y_s + \epsilon$

$$\begin{aligned} -\mathrm{i}yx + \log L_N(\mathrm{i}y) = &\left[\frac{x^2}{W} + \frac{x^2}{2W^2} + \tfrac{1}{6}\log(-W) - \tfrac{1}{6}\log x + 2\zeta'(-1)\right] \\ &+ \frac{\mathrm{i}W}{12x}\epsilon + \left(\tfrac{1}{4}W + \tfrac{1}{4} - \frac{W^2}{48x^2}\right)\epsilon^2 \end{aligned} \tag{3.41}$$

(omitting the various error terms) so integrating over ϵ from $-\epsilon_\infty$ to ϵ_∞, with $1/\sqrt{\log N} \ll \epsilon_\infty \ll 1$, we have

$$q(x) \sim \frac{1}{\sqrt{\pi}\sqrt{-W-1}} \exp\left(\frac{x^2}{W} + \frac{x^2}{2W^2} + \tfrac{1}{6}\log(-W) - \tfrac{1}{6}\log x + 2\zeta'(-1)\right) \tag{3.42}$$

Combining the above results, we have proven

Theorem 3.8. *If* $|x| \ll \log N$ *then*

$$q(x) \sim \frac{1}{\sqrt{\pi(\log N + 1 + \gamma)}} \exp\left(\frac{-x^2}{\log N + 1 + \gamma}\right) \tag{3.43}$$

and if $\log N \ll |x| \ll \sqrt{N}$*, then writing* W *for* $W_{-1}\left(\frac{-x}{Ne}\right)$*,*

$$q(x) \sim \frac{1}{\sqrt{\pi}} \exp\left(\frac{-x^2}{-W} + \frac{x^2}{2W^2} - \tfrac{1}{3}\log(-W) - \tfrac{1}{6}\log x + 2\zeta'(-1)\right). \tag{3.44}$$

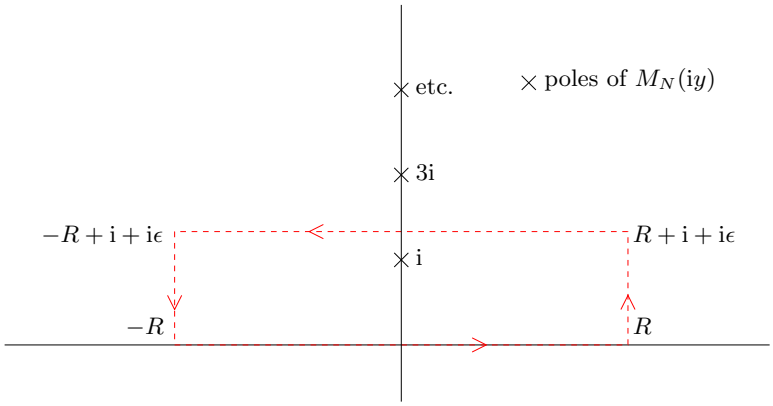

Figure 3.2: The contour C

3.5 Refined large deviations asymptotics for $\mathfrak{Re}\log Z_U(0)$

3.5.1 Refined large deviations estimates to the left

By Fourier inversion, the probability density of $\mathfrak{Re}\log Z_U(0)$ is given by

$$p(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\mathrm{i}yx} M_N(\mathrm{i}y)\,\mathrm{d}y, \tag{3.45}$$

where $M_N(\mathrm{i}y) = \mathbb{E}\, e^{\mathrm{i}y\mathfrak{Re}\log Z_U(0)}$ is given by (1.77).

Theorem 3.9. *If* $\limsup_{N\to\infty} x/\log N < -1/2$, *then*

$$p(x) \sim e^x \exp\left(3\zeta'(-1) + \tfrac{1}{12}\log 2 - \tfrac{1}{2}\log\pi\right) N^{1/4}. \tag{3.46}$$

Proof. We evaluate

$$\frac{1}{2\pi}\int_C e^{-\mathrm{i}yx} M_N(\mathrm{i}y)\,\mathrm{d}y, \tag{3.47}$$

where C is the rectangle with vertices $-R$, R, $R+\mathrm{i}+\epsilon\mathrm{i}$, $-R+\mathrm{i}+\epsilon\mathrm{i}$, for ϵ a fixed real number subject to $0<\epsilon<1$ (see figure 3.2), and let $R\to\infty$. Note that the contour encloses only the simple pole at $y=\mathrm{i}$.

The asymptotics for $G(\cdot)$ show that the integral on the sides of the contour vanish as $R\to\infty$, which means

$$p(x) = \mathrm{i}\operatorname*{Res}_{y=\mathrm{i}}\left\{e^{-\mathrm{i}yx} M_N(\mathrm{i}y)\right\} + E, \tag{3.48}$$

where

$$E = \frac{e^{x+\epsilon x}}{2\pi}\int_{-\infty}^{\infty} e^{-\mathrm{i}tx} M_N(\mathrm{i}t-1-\epsilon)\,\mathrm{d}t. \tag{3.49}$$

The pole at $y = \mathrm{i}$ is simple, so

$$\mathrm{i}\operatorname*{Res}_{y=\mathrm{i}}\left\{e^{-\mathrm{i}yx}M_N(\mathrm{i}y)\right\} = e^x G^2\left(\tfrac{1}{2}\right)\frac{G(N)G(N+1)}{G^2\left(N+\frac{1}{2}\right)}\mathrm{i}\operatorname*{Res}_{y=\mathrm{i}}\left\{\frac{1}{G(1+\mathrm{i}y)}\right\} \tag{3.50}$$

$$= e^x G^2\left(\tfrac{1}{2}\right)\frac{G(N)G(N+1)}{G^2\left(N+\frac{1}{2}\right)} \tag{3.51}$$

$$\sim e^x \exp\left(3\zeta'(-1) + \tfrac{1}{12}\log 2 - \tfrac{1}{2}\log\pi\right) N^{1/4}, \tag{3.52}$$

and

$$|E| \leq \frac{e^{x+\epsilon x}}{2\pi}\int_{-\infty}^{\infty}|M_N(\mathrm{i}t-1-\epsilon)|\mathrm{d}t \tag{3.53}$$

$$\sim \frac{e^{x+\epsilon x}}{\sqrt{\pi}}\left|\frac{G^2\left(\frac{1}{2}-\frac{1}{2}\epsilon\right)}{G(-\epsilon)}\right| N^{1/4+\epsilon/2+\epsilon^2/4}(\log N)^{-1/2}. \tag{3.54}$$

(The constant term is well defined for $0 < \epsilon < 1$).

Thus $|E| \ll e^x N^{1/4}$ when

$$e^{x\epsilon}N^{\epsilon/2+\epsilon^2/4}(\log N)^{-1/2} \ll 1. \tag{3.55}$$

Thus the error term can be made subdominant if

$$\limsup_{N\to\infty}\frac{x}{\log N} < -\frac{1}{2} \tag{3.56}$$

by choosing

$$0 < \epsilon < \min\left\{-2 - 4\limsup_{N\to\infty}\frac{x}{\log N}\ ,\ 1\right\}, \tag{3.57}$$

which completes the proof of the theorem. □

For $x < 0$, it is possible to extend the above argument to include all poles, by integrating over the rectangle with vertices $-R$, R, $R + \mathrm{i}R$, $-R + \mathrm{i}R$, and letting $R \to \infty$ in order to show that

Theorem 3.10. *For $x < 0$,*

$$p(x) = \sum_{n=1}^{\infty} e^{(2n-1)x}\operatorname*{Res}_{s=0}\left\{e^{-sx}M_N(s-2n+1)\right\}. \tag{3.58}$$

The problem with this evaluation of $p(x)$ is that if one wishes to treat this as an asymptotic (i.e. truncatable) sum then $x \ll -\log N$ would be required. The reason for this is explained in the following theorem:

Theorem 3.11. *For $n = 1, 2, 3, \ldots$ and fixed, $N \gg 1$, $x < 0$ with $|x| \gg 1$,*

$$\operatorname*{Res}_{s=0} \left\{ e^{-sx} M_N(s - 2n + 1) \right\} = (-1)^{n-1} \frac{G^2\left(\frac{3}{2} - n\right)}{G(2n)} \frac{\left(x + (n - \frac{1}{2}) \log N\right)^{2n-2}}{(2n-2)!} N^{(n-1/2)^2} \times \times \left\{ 1 + \mathcal{O}\left(x^{-1}\right) + \mathcal{O}\left((\log N)^{-1}\right) \right\}. \quad (3.59)$$

Proof. The residue is the s^{-1} term in the Laurent expansion of

$$e^{-sx} M_N(s - 2n + 1) = \frac{G(N+1)G(s+N-2n+2)G^2\left(\frac{1}{2}s - n + \frac{3}{2}\right)}{G(s+1)G^2\left(\frac{1}{2}s + N - n + \frac{3}{2}\right)\exp(sx)} \prod_{j=0}^{2n-2} \Gamma(-j+s) \quad (3.60)$$

(which comes from (1.77) and the recurrence relation for the G–function). Now, for $j = 0, 1, 2, \ldots$, $\Gamma(-j+s)$ has a simple pole at $s = 0$ with residue $(-1)^j/j!$. Therefore,

$$\prod_{j=0}^{2n-2} \Gamma(-j+s) = \sum_{k=-2n+1}^{\infty} \alpha_k s^k \quad (3.61)$$

with

$$\alpha_{-2n+1} = (-1)^{n-1} \prod_{j=0}^{2n-2} \frac{1}{j!} \quad (3.62)$$

$$= (-1)^{n-1} \frac{1}{G(2n)}. \quad (3.63)$$

Now,

$$F(s) := \frac{G(N+1)G(s+N-2n+2)G^2\left(\frac{1}{2}s - n + \frac{3}{2}\right)}{G(s+1)G^2\left(\frac{1}{2}s + N - n + \frac{3}{2}\right)\exp(sx)} \quad (3.64)$$

is analytic at $s = 0$, and we require the Taylor expansion around that point. Actually, it is easier to get the Taylor expansion of $\log F(s)$:

$$\log F(s) = \sum_{k=0}^{\infty} a_k s^k \quad (3.65)$$

where

$$a_0 = \log \frac{G(N+1)G(N-2n+2)G^2\left(\frac{3}{2} - n\right)}{G^2\left(N - n + \frac{3}{2}\right)} \quad (3.66)$$

$$= \left(n - \tfrac{1}{2}\right)^2 \log N + 2 \log G\left(\tfrac{3}{2} - n\right) + \mathcal{O}\left(\frac{1}{N}\right), \quad (3.67)$$

$$a_1 = -x + \Phi(N-2n+2) + \Phi\left(\tfrac{3}{2}-n\right) - \Phi(1) - \Phi\left(N-n+\tfrac{3}{2}\right) \tag{3.68}$$

$$= -x - \left(n-\tfrac{1}{2}\right)\log N + \mathcal{O}(1), \tag{3.69}$$

and for $k \geq 2$,

$$a_k = \tfrac{1}{k!}\Phi^{(k-1)}(N-2n+2) + \tfrac{1}{2^{k-1}k!}\Phi^{(k-1)}\left(\tfrac{3}{2}-n\right) - \tfrac{1}{k!}\Phi^{(k-1)}(1) - \tfrac{1}{2^{k-1}k!}\Phi^{(k-1)}\left(N-n+\tfrac{3}{2}\right) \tag{3.70}$$

where $\Phi^{(k)}(z) := \frac{\mathrm{d}^{k+1}}{\mathrm{d}z^{k+1}}\log G(z)$. (The asymptotics of $\Phi^{(k)}(z)$ are described in appendix A). We have $a_2 = \frac{1}{4}\log N + \mathcal{O}(1)$ and $a_k = \mathcal{O}(1)$ for $k \geq 3$. So that a_1 is truly large, we require that $x/\log N$ is bounded away from $-(n-\frac{1}{2})$ as $N \to \infty$.

Therefore,

$$F(s) = e^{a_0}\exp\left(\sum_{k=1}^{\infty} a_k s^k\right) \tag{3.71}$$

$$= e^{a_0}\sum_{j=0}^{\infty} b_j s^j, \tag{3.72}$$

where $b_0 = 1$ and

$$b_j = a_j + \frac{1}{2!}\sum_{k_1+k_2=j} a_{k_1}a_{k_2} + \frac{1}{3!}\sum_{k_1+k_2+k_2=j} a_{k_1}a_{k_2}a_{k_3} + \cdots + \frac{1}{j!}a_1^j \tag{3.73}$$

$$= \frac{1}{j!}a_1^j + \mathcal{O}\left(a_1^{j-2}a_2\right). \tag{3.74}$$

And so,

$$\operatorname*{Res}_{s=0}\left\{e^{-sx}M_N(s-(2n-1))\right\} = \exp(a_0)\sum_{k=1}^{2n-1}\alpha_{-k}b_{k-1} \tag{3.75}$$

$$= \exp(a_0)\left\{\alpha_{-2n+1}b_{2n-2} + \mathcal{O}\left(b_{2n-3}\right)\right\} \tag{3.76}$$

which gives the required result and error terms. □

Remark. Note that for each n,

$$e^{(2n+1)x}\operatorname*{Res}_{s=0}\left\{e^{-sx}M_N(s-(2n+1))\right\} \ll e^{(2n-1)x}\operatorname*{Res}_{s=0}\left\{e^{-sx}M_N(s-(2n-1))\right\} \tag{3.77}$$

only if $\limsup_{N\to\infty}\frac{x}{\log N} < -n$, which explains the remark made after theorem 3.10.

It is tedious, but not hard, to obtain much better error terms for the large N expansions of the residue. For example, take $n = 2$ then

$$\operatorname*{Res}_{s=0}\left\{e^{-sx}M_N(s-3)\right\} = e^{a_0}\left(\alpha_{-3}b_2 + \alpha_{-2}b_1 + \alpha_{-1}b_0\right) \tag{3.78}$$

where $\alpha_{-3} = -1/2$, $\alpha_{-2} = 3\gamma/2 - 5/4$ and $\alpha_{-1} = -\pi^2/8 - 9\gamma^2/4 - 17/8 + 15\gamma/4$, and $b_0 = 1$, $b_1 = a_1$ and $b_2 = a_2 + \frac{1}{2}a_1^2$, where

$$a_1 = -x - \tfrac{3}{2}\log N + \tfrac{3}{2}\gamma - \tfrac{3}{2} + 3\log 2 + \mathcal{O}\left(\frac{1}{N}\right) \tag{3.79}$$

$$a_2 = \tfrac{1}{4}\log N - \tfrac{3}{4} + \tfrac{1}{4}\gamma - \tfrac{1}{2}\log 2 - \tfrac{3}{16}\pi^2 + \mathcal{O}\left(\frac{1}{N}\right) \tag{3.80}$$

Therefore, the residue of $e^{-sx}M_N(s-3)$ at $s = 0$ is

$$G^2(-\tfrac{1}{2})\left(-\tfrac{1}{4}x^2 + \left(-\tfrac{3}{4}\log N + \tfrac{1}{2} - \tfrac{3}{4}\gamma + \tfrac{3}{2}\log 2\right)x + A\right)N^{9/4} + \mathcal{O}\left(xN^{5/4}\right) + \mathcal{O}\left((\log N)N^{5/4}\right) \tag{3.81}$$

where

$$A = -\tfrac{9}{16}(\log N)^2 + \left(\tfrac{5}{8} + \tfrac{9}{4}\log 2 - \tfrac{9}{8}\gamma\right)\log N + \tfrac{5}{8}\gamma - \tfrac{5}{4}\log 2 - \tfrac{1}{32}\pi^2 - \tfrac{9}{4}(\log 2)^2 - \tfrac{9}{16}\gamma^2 - \tfrac{7}{16} + \tfrac{9}{4}\gamma\log 2 \tag{3.82}$$

One can write

$$p(x) = e^x \operatorname*{Res}_{s=0}\left\{e^{-sx}M_N(s-1)\right\} + e^{3x}\operatorname*{Res}_{s=0}\left\{e^{-sx}M_N(s-3)\right\} + E_2 \tag{3.83}$$

with

$$E_2 = \frac{e^{3x+\epsilon x}}{2\pi}\int_{-\infty}^{\infty} e^{-\mathrm{i}tx}M_N(\mathrm{i}t - 3 - \epsilon)\,\mathrm{d}t, \tag{3.84}$$

for $0 < \epsilon < 2$. An estimate similar to that used in theorem 3.9 gives that E_2 is subdominant to the first two terms for $\limsup_{N\to\infty}\frac{x}{\log N} < -3/2$.

The final result in this section is that for small N one can use the method of residues to calculate $p(x)$ explicitly.

Theorem 3.12. *For $N = 1$, and $x < \log 2$,*

$$p(x) = \frac{2}{\pi}\frac{e^x}{\sqrt{4 - e^{2x}}} \tag{3.85}$$

Proof. For $n = 1, 2, \ldots,$

$$\operatorname*{Res}_{s=0}\left\{e^{-sx}\frac{\Gamma(2-2n+s)}{\Gamma^2\left(\frac{3}{2} - n + \frac{1}{2}s\right)}\right\} = \frac{1}{\Gamma^2\left(\frac{3}{2} - n\right)(2n-2)!} \tag{3.86}$$

$$= \frac{1}{\pi}\frac{(2n-2)!}{2^{4n-4}(n-1)!^2} \tag{3.87}$$

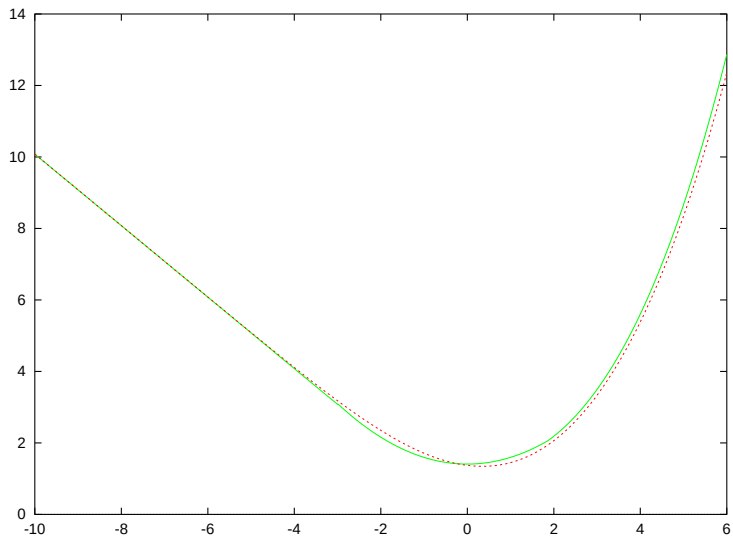

Figure 3.3: Graph of the negative log of the probability density of $\mathfrak{Re}\log Z_U(0)$ (from N. Snaith) (red) against its leading order asymptotics (green), for $N = 42$.

and so, by theorem 3.10, for $x < 0$,

$$p(x) = \frac{1}{\pi}\sum_{n=0}^{\infty} e^{(2n+1)x}\frac{(2n)!}{2^{4n}n!^2} \tag{3.88}$$

$$= \frac{2}{\pi}\frac{e^x}{\sqrt{4-e^{2x}}}. \tag{3.89}$$

But actually (3.89) is valid for all $x < \log 2$. (Note that $\mathfrak{Re}\log Z_U(0) \leq \log 2$ for $N = 1$). To see this, observe

$$p(x) = \frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}\left\{\log\left|1-e^{\mathrm{i}\theta_1}\right| \leq x\right\} \tag{3.90}$$

$$= \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbb{1}_{\{\log|1-e^{\mathrm{i}\theta_1}|\leq x\}}\,\mathrm{d}\theta_1 \tag{3.91}$$

$$= \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{\pi}\arccos\left(1-\tfrac{1}{2}e^{2x}\right) \tag{3.92}$$

$$= \frac{2}{\pi}\frac{e^x}{\sqrt{4-e^{2x}}}. \tag{3.93}$$

So maybe (by analytic continuation) theorem 3.10 gives $p(x)$ for all $x < N\log 2$, for all N. □

3.5.2 Refined large deviation estimates to the right

In the previous subsection

$$p(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\mathrm{i}yx}M_N(\mathrm{i}y)\,\mathrm{d}y \tag{3.94}$$

was found to be dominated by the pole at $y = \mathrm{i}$ when $x < -\frac{1}{2}\log N$. Here, the saddle point method will be used to estimate $p(x)$ for $|x| \ll \log N$ and for $\log N \ll x \ll N^{1/3}$.

As is usual in the saddle-point method, the contour of integration will be deformed to pass through the point where the Taylor expansion of $-\mathrm{i}yx + \log M_N(\mathrm{i}y)$ is stationary. Making the ansatz that such a saddle-point will occur for $|y_s| \ll 1$, then using Appendix C on the asymptotics of $M_N(\mathrm{i}y)$ and noting the error terms are differentiable,

$$y_s = \frac{-2\mathrm{i}x}{\log N + 1 + \gamma} + \mathcal{O}\left(\frac{x^2}{(\log N)^2}\right) \tag{3.95}$$

which will be valid so long as $|x| \ll \log N$. Writing $y = y_s + \epsilon$ and Taylor expanding around y_s,

$$-\mathrm{i}yx + \log M_N(\mathrm{i}y) = \frac{-x^2}{\log N + 1 + \gamma} - \frac{1}{4}(\log N + 1 + \gamma)\epsilon^2 + \mathcal{O}(\epsilon^3) \tag{3.96}$$

where the constant in the $\mathcal{O}(\epsilon^3)$ term can be made independent of N, and thus, integrating over $(-\epsilon_\infty, \epsilon_\infty)$ where $1/\sqrt{\log N} \ll \epsilon_\infty \ll 1$ we have

$$p(x) \sim \frac{1}{\sqrt{\pi(\log N + 1 + \gamma)}} \exp\left(\frac{-x^2}{\log N + 1 + \gamma}\right) \tag{3.97}$$

which is the density function of a normal random variable with mean zero and variance $\frac{1}{2}(\log N + 1 + \gamma)$.

If we make the ansatz that the saddle point, y_s, occurs for $1 \ll y_s \ll \sqrt[3]{N}$, then (again using Appendix C) we need to solve

$$\mathrm{i}x + \tfrac{1}{2}y_s\left(\log N - \log(\mathrm{i}y_s) - \log 2 + 1\right) + \mathcal{O}\left(\frac{y_s^2}{N}\right) + \mathcal{O}\left(\frac{1}{y_s^2}\right) = 0 \tag{3.98}$$

the solution being

$$y_s = \frac{2\mathrm{i}x}{W} + \mathcal{O}\left(\frac{x^2}{NW^2}\right) + \mathcal{O}\left(\frac{W^2}{x^2}\right) \tag{3.99}$$

where $W = W_{-1}\left(\frac{-4x}{eN}\right)$ is the -1 branch of the Lambert W–function (see appendix B). Our ansatz on y_s is valid so long as $\log N \ll x \ll \sqrt[3]{N}$.

Writing

$$y = \frac{2\mathrm{i}x}{W} + \epsilon \tag{3.100}$$

we find that $-\mathrm{i}yx + \log M_N(\mathrm{i}y)$ equals

$$\frac{2x^2}{W} - \mathrm{i}\epsilon x + \frac{1}{4}\left(\frac{2x}{-W} + \mathrm{i}\epsilon\right)^2 \left[\log N - \log\left(\frac{2x}{-W} + \mathrm{i}\epsilon\right) - \log 2 + \tfrac{3}{2}\right] + \tfrac{1}{6}\log 2 + \zeta'(-1) - \tfrac{1}{12}\log\left(\frac{2x}{-W} + \mathrm{i}\epsilon\right) + \mathcal{O}\left(\frac{x^3}{N}\right) + \mathcal{O}\left(\frac{1}{x}\right) \tag{3.101}$$

and expanding this in powers of ϵ, it equals

$$\left[\frac{x^2}{W} + \frac{x^2}{2W^2} + \tfrac{1}{12}\log(-W) + \tfrac{1}{12}\log 2 - \tfrac{1}{12}\log x + \zeta'(-1)\right] + \frac{\mathrm{i}W}{24x}\epsilon + \epsilon^2\left(\tfrac{1}{4}W + \tfrac{1}{4} - \frac{1}{96}\frac{W^2}{x^2}\right) + \mathcal{O}(\epsilon^3) + \mathcal{O}\left(\frac{x^3}{N}\right) + \mathcal{O}\left(\frac{1}{x}\right) \tag{3.102}$$

where we have used $\log(-W) = \log 4x - \log N - 1$. Integrating this over $(-\epsilon_\infty, \epsilon_\infty)$ with respect to ϵ, for $1/\sqrt{\log N} \ll \epsilon_\infty \ll 1$ the saddle-point method gives that $p(x)$ is asymptotic to

$$\frac{1}{\sqrt{\pi}\sqrt{-W}} \exp\left(\frac{x^2}{W} + \frac{x^2}{2W^2} + \tfrac{1}{12}\log(-W) - \tfrac{1}{12}\log x + \tfrac{1}{12}\log 2 + \zeta'(-1)\right) \tag{3.103}$$

Together, this proves

Theorem 3.13. *For $|x| \ll \log N$,*

$$p(x) \sim \frac{1}{\sqrt{\pi(\log N + 1 + \gamma)}} \exp\left(\frac{-x^2}{\log N + 1 + \gamma}\right) \tag{3.104}$$

while for $\log N \ll x \ll N^{1/3}$,

$$p(x) \sim \frac{1}{\sqrt{\pi}} \exp\left(\frac{-x^2}{-W} + \frac{x^2}{2W^2} - \tfrac{5}{12}\log(-W) - \tfrac{1}{12}\log x + \tfrac{1}{12}\log 2 + \zeta'(-1)\right) \tag{3.105}$$

where we have written W for $W_{-1}\left(-\frac{4x}{eN}\right)$.

3.6 Large deviations for the zeta function

3.6.1 Maximum size of $|\zeta(1/2 + \mathrm{i}t)|$

In this subsection we study how large $|\zeta(1/2 + \mathrm{i}t)|$ can get on $0 \le t \le T$.

Under the Riemann Hypothesis, it is known that (see theorem 14.14 of [90], for example) as $t \to \infty$

$$\log\left|\zeta\left(\tfrac{1}{2} + \mathrm{i}t\right)\right| = \mathcal{O}\left(\frac{\log t}{\log\log t}\right) \tag{3.106}$$

(This is stronger that the Lindelöf hypothesis, which conjectures that for any $\epsilon > 0$, $|\zeta(1/2 + \mathrm{i}t)| \leq t^{\epsilon}$ as $t \to \infty$).

Balasubramanian and Ramachandra [3] have shown that, independent of RH, there exists t_0 such that for $t > t_0$

$$\left|\zeta\left(\tfrac{1}{2} + \mathrm{i}t\right)\right| \geq \exp\left(\frac{3}{4}\sqrt{\frac{\log t}{\log\log t}}\right) \tag{3.107}$$

Montgomery [77] has conjectured that the right-hand side of (3.107) is about the correct rate of growth.

Jutila [65] has shown that, independent of RH, for $T \geq 2$ and $1 \leq V \leq \log T$

$$\frac{1}{T}\operatorname{meas}\left\{0 \leq t \leq T : \left|\zeta\left(\tfrac{1}{2} + \mathrm{i}t\right)\right| \geq V\right\} \leq A\exp\left(-\frac{(\log V)^2}{\log\log T}\left(1 + \mathcal{O}\left(\frac{\log V}{\log\log T}\right)\right)\right) \tag{3.108}$$

for A a certain constant. This is in agreement with the upper bound in the LDP for $\mathfrak{Re}\log Z_U(0)$ (theorem 3.6), for scalings between $\sqrt{\log N}$ and $\log N$ (that is with $\exp(\sqrt{\log\log T}) \ll V \leq \log T$). Selberg's result (1.71) covers $1 \leq V << \exp(\sqrt{\log\log T})$ in a more refined manner.

We will use the method of large moments and the Keating-Snaith conjecture (conjecture 1.4) in an attempt to calculate the maximum value $\log|\zeta(1/2 + \mathrm{i}t)|$ obtains for t between 0 and T. The first part of this proof is essentially the same as that in Conrey and Gonek's paper [30].

Since

$$\max_{t\in[0,T]} \log\left|\zeta\left(\tfrac{1}{2} + \mathrm{i}t\right)\right| = \log \max_{t\in[0,T]} \left|\zeta\left(\tfrac{1}{2} + \mathrm{i}t\right)\right|, \tag{3.109}$$

and for all k,

$$\log \max_{t\in[0,T]} \left|\zeta\left(\tfrac{1}{2} + \mathrm{i}t\right)\right| \geq \log\left(\frac{1}{T}\int_0^T \left|\zeta\left(\tfrac{1}{2} + \mathrm{i}t\right)\right|^{2k} \mathrm{d}t\right)^{\frac{1}{2k}} \tag{3.110}$$

it makes sense to maximise the right-hand side with respect to k. Assuming conjecture 1.4 is valid for $k = k(T) \to \infty$ (and this (false) assumption will be discussed later), we will maximise

$$\log\left(\frac{G^2(1+k)}{G(1+2k)}a(k)N^{k^2}\right)^{1/2k} = \frac{1}{2k}\left(2\log G(1+k) - \log G(1+2k) + \log a(k) + k^2\log N\right) \tag{3.111}$$

with respect to k, where we have written N for $\log \frac{T}{2\pi}$ for simplicity.

It has been shown [30] that

$$\log a(k) \sim -k^2 \log(2e^{\gamma} \log k) + o(k^2) \text{ for } k \to \infty \tag{3.112}$$

and from appendix A we have that as $k \to \infty$,

$$2 \log G(1+k) - \log G(1+2k) = k^2 \left(-\log k + \tfrac{3}{2} - 2\log 2 + o(1)\right) \tag{3.113}$$

Therefore $\log |\zeta\left(\frac{1}{2} + \mathrm{i}t\right)|$ must get larger than the maximum over k of

$$-\tfrac{1}{2}k \left(\log\log k + \log k - \log N + \gamma - \tfrac{3}{2} + 3\log 2 + o(1)\right) \tag{3.114}$$

Differentiating this, we wish to solve

$$-\tfrac{1}{2}(\log\log k + \log k - \log N + \gamma - \tfrac{3}{2} + 3\log 2) - \tfrac{1}{2}k\left(\frac{1}{\log k}\frac{1}{k} + \frac{1}{k}\right) = 0 \tag{3.115}$$

We may ignore the $(\log k)^{-1}$ term, since by assumption we have $k \gg 1$, thus we wish to solve

$$\log\log k + \log k - \log N + \gamma - \tfrac{1}{2} + 3\log 2 =: \log\left(\frac{k\log k}{BN}\right) \tag{3.116}$$

$$= 0 \tag{3.117}$$

where we have defined B by $-\log B = \gamma - \frac{1}{2} + 3\log 2$. The solution is

$$k = e^{W(BN)} = \frac{BN}{W(BN)} \tag{3.118}$$

where $W(x)$ is the principal branch of the Lambert W–function. (Note that, for sufficiently large N, $k \gg 1$ as predicted).

Substituting this back into (3.114), we find that $\log |\zeta\left(\frac{1}{2} + \mathrm{i}t\right)|$ must get larger than

$$-\frac{BN}{2W(BN)}\Big(\log W(BN) + W(BN) - \log N + \gamma - \frac{3}{2} + 3\log 2\Big) = \frac{BN}{2W(BN)} \tag{3.119}$$

where we have used the fact that $\log W(BN) = \log B + \log N - W(BN)$.

Since $B = \exp(\frac{1}{2} - \gamma - 3\log 2)$, $W(BN) = \log N + \mathcal{O}(\log\log N)$ and $N = \log \frac{T}{2\pi}$ we have, to dominant order,

$$\max_{t\in[0,T]} \log |\zeta\left(\tfrac{1}{2} + \mathrm{i}t\right)| \geq \exp(\tfrac{1}{2} - \gamma - 4\log 2)\frac{\log T}{\log\log T} \tag{3.120}$$

which suggests (3.106) is the correct rate of growth. The premultiplying constant is $\exp(\frac{1}{2} - \gamma - 4\log 2) \approx 0.0579$. (The constant in [30] is out by a factor of $1/e$).

However, it follows from lemma 3.14 that for any real continuous function $f(\cdot)$,

$$\sup_{x\in[0,T]} |f(x)| = \lim_{k\to\infty} \left(\frac{1}{T} \int_0^T |f(x)|^k \, \mathrm{d}x \right)^{1/k}. \tag{3.121}$$

Therefore, the true maximum of $\log|\zeta(1/2+\mathrm{i}t)|$ for $0 \le t \le T$ is the limit as $k \to \infty$ of (3.110). Substituting in the Keating-Snaith conjecture, this limit is found to be $-\infty$. This is because conjecture 1.4 only gives the leading order asymptotics for large T, and takes no account of large k error terms. We therefore cannot conclude that the Keating-Snaith conjecture implies (3.106) is the true rate of growth, since the fact that there existed a maximum of (3.114) means we were working in a regime where the errors which we ignored (since we don't know them) are important.

Lemma 3.14. *For any real, continuous function $f(\cdot)$,*

$$\log \int_0^T |f(t)|^k \mathrm{d}t \tag{3.122}$$

is a convex function of k.

Proof. Hölders inequality gives for $0 < \theta < 1$,

$$\int_0^T |f(t)|^{\theta\lambda_1} |f(t)|^{(1-\theta)\lambda_2} \mathrm{d}t \le \left(\int_0^T |f(t)|^{\theta\lambda_1/\theta} \mathrm{d}t \right)^{\theta} \left(\int_0^T |f(t)|^{(1-\theta)\lambda_1/(1-\theta)} \mathrm{d}t \right)^{1-\theta} \tag{3.123}$$

and so taking logarithms of both sides, we have

$$\log \int_0^T |f(t)|^{\theta\lambda_1+(1-\theta)\lambda_2} \mathrm{d}t \le \theta \log \int_0^T |f(t)|^{\lambda_1} \mathrm{d}t + (1-\theta) \log \int_0^T |f(t)|^{\lambda_2} \mathrm{d}t \tag{3.124}$$

which is the definition of convexity. □

Now, (3.110) is true for all k, so if one could find the largest value of k for which Keating-Snaith is valid, then this would still give a lower bound for the maximum size of the zeta function on the critical line. Lemma 3.14 says that

$$\log \left(\frac{1}{T} \int_0^T \left|\zeta\left(\tfrac{1}{2} + \mathrm{i}t\right)\right|^{2k} \, \mathrm{d}t \right) \tag{3.125}$$

is a convex function, and we can use this fact to obtain a not-necessarily-not-valid bound for zeta by finding the point where

$$\log \left(\frac{G^2(1+k)}{G(1+2k)} a(k) N^{k^2} \right) \tag{3.126}$$

ceases to be a convex function. This will be the largest value of k for which the Keating-Snaith conjecture *might* be valid (we are not saying it *is* valid at that point, only that it is not necessarily not valid).

Convexity will cease when

$$0 > \frac{\mathrm{d}^2}{\mathrm{d}k^2} \log \left(\frac{G^2(1+k)}{G(1+2k)} a(k) N^{k^2} \right) \tag{3.127}$$

$$= -2 \left(\log\log k + \log k - \log N + 3\log 2 + \gamma \right) + o(1) \tag{3.128}$$

using the same methods as above. Writing $C = \exp(-3\log 2 - \gamma)$ and solving for the smallest k such that convexity ceases, we find that

$$k = \frac{CN}{W(CN)} \tag{3.129}$$

Assuming the Keating-Snaith conjecture is valid at this point, since (3.110) is valid for all k, using this value of k in (3.114), exactly as we did before, we find that

$$\max_{t\in[0,T]} \log \left|\zeta(\tfrac{1}{2}+\mathrm{i}t)\right| \geq \exp(\log 3 - 5\log 2 - \gamma) \frac{\log T}{\log\log T} \tag{3.130}$$

which still suggests (3.106) is the correct rate of growth. The premultiplying constant here is $\exp(\log 3 - 5\log 2 - \gamma) \approx 0.0526$.

3.6.2 Large left deviations for $\log|\zeta(1/2+\mathrm{i}t)|$

Define, for $x > 0$,

$$P(T,x) = \frac{1}{T} \operatorname{meas} \left\{ t : 0 \leq t \leq T \ , \ \log \left|\zeta\left(\tfrac{1}{2}+\mathrm{i}t\right)\right| \leq -x \right\}, \tag{3.131}$$

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

Recall that

$$\mathbb{P}\left\{\log|Z_U(0)| \leq -x\right\} = \int_{-\infty}^{-x} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\mathrm{i}yz} M_N(\mathrm{i}y) \, \mathrm{d}y \, \mathrm{d}z \tag{3.132}$$

and, as shown in §3.5.1, for $x \gg \log N$ the Fourier integral is dominated by the simple pole of $M_N(\mathrm{i}y)$ at $y = \mathrm{i}$.

Recall also conjecture 1.4 which says that for fixed y, as $T \to \infty$,

$$\frac{1}{T} \int_0^T \exp\left(\mathrm{i}y \log \left|\zeta\left(\tfrac{1}{2}+\mathrm{i}t\right)\right|\right) \mathrm{d}t \sim \frac{G^2\left(1+\frac{1}{2}\mathrm{i}y\right)}{G(1+\mathrm{i}y)} \left(\log \frac{T}{2\pi}\right)^{-y^2/4} a\left(\tfrac{1}{2}\mathrm{i}y\right), \tag{3.133}$$

which will play the part of $M_N(\mathrm{i}y)$ in (3.132), with $N = \log \frac{T}{2\pi}$. This suggests that for large T,

$$P(T,x) \sim \int_{-\infty}^{-x} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\mathrm{i}yz} \frac{G^2\left(1+\frac{1}{2}\mathrm{i}y\right)}{G(1+\mathrm{i}y)} \left(\log \frac{T}{2\pi}\right)^{-y^2/4} a\left(\tfrac{1}{2}\mathrm{i}y\right) \, \mathrm{d}y \, \mathrm{d}z \tag{3.134}$$

Note that conjecture 1.4 is for fixed y, whereas the Fourier inversion (needed to obtain the probability density) involves an integral over all of $\mathbb{R}$. However, we believe that, for $x > 0$, the Fourier inversion of conjecture 1.4 is valid, since for $x \ll \log\log T$ the Fourier integral is dominated by the saddle point near $y = 0$, and if $x \geq (\frac{1}{2}+\epsilon)\log\log T$ it is dominated by the simple pole at $y = \mathrm{i}$. That is, if $x \geq (\frac{1}{2}+\epsilon)\log\log T$, it seems highly plausible that

$$P(T,x) = e^{-x} G^2\left(\tfrac{1}{2}\right) \left(\log \frac{T}{2\pi}\right)^{1/4} a\left(-\tfrac{1}{2}\right) \operatorname*{Res}_{y=\mathrm{i}} \left\{ \frac{1}{G(1+\mathrm{i}y)} \right\} + \mathcal{O}_T\left(e^{-3x+\epsilon}\right), \tag{3.135}$$

and so, writing out $G^2(1/2)$ explicitly, we have:

Conjecture 3.15. *As $T \to \infty$, if $x \geq (\frac{1}{2}+\epsilon)\log\log T$*

$$P(T,x) \sim e^{-x} \exp\left(3\zeta'(-1) + \tfrac{1}{12}\log 2 - \tfrac{1}{2}\log\pi\right) a\left(-\tfrac{1}{2}\right) \left(\log \frac{T}{2\pi}\right)^{1/4}. \tag{3.136}$$

Comparing this conjecture with theorem 3.9, we see they agree apart from the $a(-1/2)$ factor. Using figure 1.1, numerically we have for T close to the 10^{20}th zero,

$$-\log P(T,10) + \log \mathbb{P}\{\log|Z_U(0)| \leq -10\} \approx 0.087, \tag{3.137}$$

whereas $\log a(-1/2) = -0.085\ldots$, which is in good agreement.

That x is restricted to being greater than $\frac{1}{2}\log\log T$ is important, since, for similar reasons to those given in section 3.3 for $\mathfrak{Re}\log Z_U(0)$, we expect a phase transition to occur there:

Conjecture 3.16.

$$\lim_{T\to\infty} \frac{\log P(T, y\log\log T)}{\log\log T} = \begin{cases} -y^2 & \textit{for } 0 < y < \frac{1}{2} \\ \frac{1}{4} - y & \textit{for } y > \frac{1}{2} \end{cases} \tag{3.138}$$

Remark. Letting $x \to \infty$ in conjecture 3.15 one moves well away from the phase-transition point, and obtains

$$\lim_{x\to\infty} e^x P(T,x) \sim \exp\left(3\zeta'(-1) + \tfrac{1}{12}\log 2 - \tfrac{1}{2}\log\pi\right) a\left(-\tfrac{1}{2}\right) \left(\log \frac{T}{2\pi}\right)^{1/4} \tag{3.139}$$

as $T \to \infty$. The conjecture in this "safe-form" has previously been published in [59]. If $p(x)$ is the probability density of $|\zeta(1/2 + \mathrm{i}t)|$, then (3.139) is equivalent to the conjecture $p(0) \sim G^2(1/2)a(-1/2)\left(\log \frac{T}{2\pi}\right)^{1/4}$, which is equation (108) of [69].

3.7 Inside the circle

The sequence of spectral measures

$$S_N = \frac{1}{N}\sum_{n=1}^{N} \delta_{\theta_n} \tag{3.140}$$

satisfies the LDP in $M_1(\mathbb{T})$ (the set of all probability measures defined on $\mathbb{T}$) with speed N^2 and good convex rate function given by the logarithmic energy functional

$$\Sigma(\nu) = -\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \log|e^{\mathrm{i}\theta_1} - e^{\mathrm{i}\theta_2}|\nu(\mathrm{d}\theta_1)\nu(\mathrm{d}\theta_2). \tag{3.141}$$

For a proof of this fact, see [58]. (The analogue of this LDP for the GUE was obtained in [8]).

Varadhan's lemma (see, for example, [34]) enables us to calculate the logarithmic moment generating function (and hence the LDP) for functions of S_N as follows:

Theorem 3.17. *For any continuous $F : M_1(\mathbb{T}) \longrightarrow \mathbb{R}$ satisfying the condition*

$$\limsup_{N\to\infty} \frac{1}{N^2} \log \mathbb{E}\, e^{\lambda N^2 F(S_N)} < \infty \tag{3.142}$$

for some $\lambda > 1$, then

$$\lim_{N\to\infty} \frac{1}{N^2} \log \mathbb{E}\, e^{N^2 F(S_N)} = \sup_{\nu \in M_1(\mathbb{T})} \{F(\nu) - \Sigma(\nu)\}. \tag{3.143}$$

Remark. The supremising ν can be interpreted as the most likely distribution of eigenangles given that such a large deviation has occurred.

Now, we can write $s\mathfrak{Re}\log Z_U(0)/N + t\mathfrak{Im}\log Z_U(0)/N = F(S_N; s, t)$, where

$$F(\nu; s, t) := \int_{-\pi}^{\pi} \left(s\mathfrak{Re}\log\left(1 - e^{\mathrm{i}\theta}\right) + t\mathfrak{Im}\log\left(1 - e^{\mathrm{i}\theta}\right)\right)\nu(\mathrm{d}\theta), \tag{3.144}$$

but F is not continuous in the weak topology (the real part of the log has a singularity and the imaginary part has a jump discontinuity for all measures whose support includes $\theta = 0$), and so Varadhan's lemma does not apply.

Consider instead, for $\epsilon > 0$, the continuous function

$$F_\epsilon(\nu; s, t) := \int_{-\pi}^{\pi} \left(s\mathfrak{Re} \log \left(1 - e^{-\epsilon} e^{\mathrm{i}\theta}\right) + t\mathfrak{Im} \log \left(1 - e^{-\epsilon} e^{\mathrm{i}\theta}\right)\right) \nu(\mathrm{d}\theta). \tag{3.145}$$

Then $[s\mathfrak{Re} \log Z_\epsilon(U) + t\mathfrak{Im} \log Z_\epsilon(U)]/N = F_\epsilon(S_N; s, t)$, where

$$Z_\epsilon(U) = \prod_{n=1}^{N} \left(1 - e^{-\epsilon} e^{\mathrm{i}\theta_n}\right), \tag{3.146}$$

(so that $Z_U(0) = \lim_{\epsilon\to 0} Z_\epsilon(U)$, and we are effectively evaluating the log of the characteristic polynomial inside the unit circle). Applying Varadhan's lemma, we obtain

$$\Lambda_\epsilon(s, t) := \lim_{N\to\infty} \frac{1}{N^2} \log \mathbb{E}\, e^{Ns\mathfrak{Re} \log Z_\epsilon(U) + Nt\mathfrak{Im} \log Z_\epsilon(U)} \tag{3.147}$$

$$= \sup_{\nu\in M_1(\mathbb{T})} \{F_\epsilon(\nu; s, t) - \Sigma(\nu)\}. \tag{3.148}$$

Theorem 3.18. *In the restricted range*

$$\sqrt{(2 + s)^2 + t^2} \le e^\epsilon + e^{-\epsilon}(1 + s), \tag{3.149}$$

$$\sup_{\nu\in M_1(\mathbb{T})} \{F_\epsilon(\nu; s, t) - \Sigma(\nu)\} \tag{3.150}$$

is maximised by the measure having probability density

$$p(\theta) = \frac{e^{2\epsilon} + 1 - 2e^\epsilon \cos\theta - se^\epsilon \cos\theta + s - te^\epsilon \sin\theta}{2\pi \left|e^\epsilon - e^{\mathrm{i}\theta}\right|^2}, \tag{3.151}$$

the maximum being

$$\Lambda_\epsilon(s, t) = \tfrac{1}{4}\left(s^2 + t^2\right) \log \left(\frac{1}{1 - e^{-2\epsilon}}\right). \tag{3.152}$$

Proof. Writing $\nu(\mathrm{d}\theta) = p(\theta)\mathrm{d}\theta$, the following optimisation problem must be solved

$$\sup_{p(\cdot)} \Big\{ \int_{-\pi}^{\pi} \left[s\mathfrak{Re} \log(1 - e^{-\epsilon} e^{\mathrm{i}\theta}) + t\mathfrak{Im} \log(1 - e^{-\epsilon} e^{\mathrm{i}\theta})\right] p(\theta)\, \mathrm{d}\theta + \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \log \left|e^{\mathrm{i}\theta_1} - e^{\mathrm{i}\theta_2}\right| p(\theta_1)p(\theta_2)\, \mathrm{d}\theta_1\, \mathrm{d}\theta_2 \Big\} \tag{3.153}$$

Now,

$$s\mathfrak{Re} \log(1 - e^{-\epsilon} e^{\mathrm{i}\theta}) + t\mathfrak{Im} \log(1 - e^{-\epsilon} e^{\mathrm{i}\theta}) = \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \left[\frac{-s}{2|k|} e^{-|k|\epsilon} e^{\mathrm{i}k\theta} + \frac{\mathrm{i}t}{2k} e^{-|k|\epsilon} e^{\mathrm{i}k\theta} \right], \tag{3.154}$$

$$\log \left|e^{\mathrm{i}\theta_1} - e^{\mathrm{i}\theta_2}\right| = \sum_{\substack{k=\infty \\ k\neq 0}}^{\infty} \frac{-1}{2|k|} e^{-\mathrm{i}k\theta_1} e^{\mathrm{i}k\theta_2} \tag{3.155}$$

and (since $p(\theta)$ is 2π-periodic)

$$p(\theta) = \sum_{k=-\infty}^{\infty} \hat{p}_k e^{\mathrm{i}k\theta} \tag{3.156}$$

(with $\hat{p}_{-k} = \hat{p}_k^*$ to make $p(\cdot)$ real). Therefore we must solve

$$\sup_{p_k} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \left[\left(\frac{-\pi s}{|k|} + \frac{\mathrm{i}\pi t}{k} \right) \hat{p}_{-k} e^{-|k|\epsilon} + \frac{-2\pi^2}{|k|} \hat{p}_k \hat{p}_{-k} \right]$$

$$= \sup_{x_k, y_k} 4\pi^2 \sum_{k=1}^{\infty} \frac{-1}{k} \left(\frac{s}{2\pi} x_k e^{-k\epsilon} - \frac{t}{2\pi} y_k e^{-k\epsilon} + x_k^2 + y_k^2 \right), \tag{3.157}$$

where $\hat{p}_k = x_k + \mathrm{i}y_k$ (with $x_{-k} = x_k$ and $y_{-k} = -y_k$).

The supremum occurs when

$$x_k = \frac{-s}{4\pi} e^{-k\epsilon}, \tag{3.158}$$

$$y_k = \frac{t}{4\pi} e^{-k\epsilon} \tag{3.159}$$

for $k = 1, 2, \ldots$. Substituting these values into (3.157) and using Varadhan's lemma, we find that

$$\Lambda_\epsilon(s,t) = 4\pi^2 \sum_{k=1}^{\infty} \frac{-1}{k} \left(\frac{-s^2}{8\pi^2} e^{-2k\epsilon} - \frac{t^2}{8\pi^2} e^{-2k\epsilon} + \frac{s^2+t^2}{16\pi^2} e^{-2k\epsilon} \right) \tag{3.160}$$

$$= \tfrac{1}{4}(s^2+t^2) \log\left(\frac{1}{1-e^{-2\epsilon}} \right). \tag{3.161}$$

This is valid so long as the maximising $p(\theta)$ is a probability density. For normalization $p_0 = \frac{1}{2\pi}$, and so all that is required for $p(\theta)$ to be a probability density is that

$$p(\theta) = \frac{1}{2\pi} + 2\mathfrak{Re} \sum_{k=1}^{\infty} \left(\frac{-s}{4\pi} e^{-k\epsilon} + \frac{\mathrm{i}t}{4\pi} e^{-k\epsilon} \right) e^{\mathrm{i}k\theta} \tag{3.162}$$

$$= \frac{1}{2\pi} \left(1 - \mathfrak{Re} \left\{ \frac{(s-\mathrm{i}t)e^{-\epsilon} e^{\mathrm{i}\theta}}{1 - e^{-\epsilon} e^{\mathrm{i}\theta}} \right\} \right) \tag{3.163}$$

$$= \frac{1 + e^{-2\epsilon} - 2e^{-\epsilon}\cos\theta - se^{-\epsilon}\cos\theta + se^{-2\epsilon} - te^{-\epsilon}\sin\theta}{2\pi \left|1 - e^{-\epsilon}e^{\mathrm{i}\theta}\right|^2} \tag{3.164}$$

is a non-negative function.

Now, it is clear that if the numerator of (3.164) is greater than or equal to zero at its minimum point, then $p(\theta)$ is non-negative. The minima and maxima of its numerator occur when

$$e^{-\epsilon}(2+s)\sin\theta - te^{-\epsilon}\cos\theta = 0, \tag{3.165}$$

that is, when

$$\theta = \arctan\left(\frac{t}{2+s}\right). \tag{3.166}$$

So for what values of s and t is

$$\begin{aligned} 1 + (1+s)e^{-2\epsilon} - (2+s)e^{-\epsilon}\cos\arctan\left(\frac{t}{2+s}\right) - te^{-\epsilon}\sin\arctan\left(\frac{t}{2+s}\right) \\ = 1 + (1+s)e^{-2\epsilon} - e^{-\epsilon}\frac{(2+s)^2}{\sqrt{(2+s)^2+t^2}} - e^{-\epsilon}\frac{t^2}{\sqrt{(2+s)^2+t^2}} \end{aligned} \tag{3.167}$$

greater than zero? The answer is when

$$\sqrt{(2+s)^2+t^2} \leq e^{\epsilon} + e^{-\epsilon}(1+s) \tag{3.168}$$

which completes the proof of the theorem. □

The problem of extending s and t beyond the range given comes from the requirement of finding the maximum over the set of all *non-negative* functions; only within the range given does the maximiser lie away from the boundary of this set.

Although $Z_\epsilon(U) \to Z_U(0)$ as $\epsilon \to 0$, it does not necessarily follow that $\Lambda_\epsilon(s,t) \to \Lambda(s,t)$. However, formally taking the limit, we do obtain $\Lambda(s,0) = \infty$ for $t = 0$ and $-2 \leq s < 0$, which is indeed what is found in theorem 3.7 for this range of s at scaling N and speed N^2.

Applying theorem 1.6 to theorem 3.18 we are able to partially calculate the rate function for the large deviations for $\log Z_\epsilon(U)$, which occurs at speed N^2:

$$\lim_{N\to\infty} \frac{1}{N^2}\log\mathbb{P}\left\{\frac{1}{N}\log|Z_\epsilon(U)| \in B\right\} = -\inf_{(x,y)\in B} I(x,y) \tag{3.169}$$

where

$$I(x,y) = \sup_{s,t\in\mathbb{R}}\left\{xs + yt - \tfrac{1}{4}(s^2+t^2)\log\left(\frac{1}{1-e^{-2\epsilon}}\right)\right\} \tag{3.170}$$

$$= \frac{x^2+y^2}{\log\left(e^{2\epsilon}/\left(e^{2\epsilon}-1\right)\right)}. \tag{3.171}$$

This is valid so long as the point where the supremum of (3.170) occurs, which is

$$s = \frac{2x}{-\log\left(1-e^{-2\epsilon}\right)}, \tag{3.172}$$

$$t = \frac{2y}{-\log\left(1-e^{-2\epsilon}\right)}, \tag{3.173}$$

satisfies the conditions of theorem 3.18. Indeed, by looking at $\log|Z_\epsilon(U)|$, one can easily see that

$$\log\left|1-e^{-\epsilon}\right| \leq \frac{1}{N}\log|Z_\epsilon(U)| \leq \log\left|1+e^{-\epsilon}\right|, \tag{3.174}$$

so we conclude that the maximising probability density must change its nature near the extreme end-points, leading to a finite maximum / minimum values for $\log|Z_\epsilon(U)|/N$.

Finally, we remark that

$$\mathfrak{Re}\log\left(1-e^{-\epsilon}e^{\mathrm{i}\theta_n}\right) = \sum_{\substack{k=-\infty\\k\neq 0}}^{\infty} \frac{-e^{-|k|\epsilon}}{2|k|}e^{\mathrm{i}k\theta_n}, \tag{3.175}$$

and

$$\mathfrak{Im}\log\left(1-e^{-\epsilon}e^{\mathrm{i}\theta_n}\right) = \sum_{\substack{k=-\infty\\k\neq 0}}^{\infty} \frac{\mathrm{i}e^{-|k|\epsilon}}{2k}e^{\mathrm{i}k\theta_n}, \tag{3.176}$$

so Szegő's theorem (see section 2.3) implies that

$$\lim_{N\to\infty}\mathbb{E}\left\{|Z_\epsilon(U)|^{2s}\exp(2t\mathfrak{Im}\log Z_\epsilon(U))\right\} = \left(\frac{1}{1-e^{-2\epsilon}}\right)^{s^2+t^2} \tag{3.177}$$

which, by theorem 1.5, means $\log Z_\epsilon(U) \Longrightarrow X+\mathrm{i}Y$ where X and Y are iid normal, with mean 0 and variance $-\frac{1}{2}\log\left(1-e^{-2\epsilon}\right)$. Note the lack of $\sqrt{\log N}$ normalization, as required in the case $\epsilon = 0$.

There is an analogous lack of normalization in the zeta function. Bohr and Jessen [16, 17] showed that for fixed $\sigma > 1/2$, there exists a continuous function $F(z,\sigma)$ such that

$$\frac{1}{2T}\operatorname{meas}\{t\in[-T,T]\ :\ \log\zeta(\sigma+\mathrm{i}t)\in R\} \to \iint_R F(x+\mathrm{i}y,\sigma)\,\mathrm{d}x\,\mathrm{d}y \tag{3.178}$$

for any rectangle $R\subseteq\mathbb{C}$ with sides parallel to the real and imaginary axes. Compare the lack of $\sqrt{\log\log T}$ normalization, which is required in Selberg's result, (1.71), on the critical line. Moments off the critical line are also known, under the assumption of RH (see [90]): For fixed $\sigma > 1/2$,

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T |\zeta(\sigma+\mathrm{i}t)|^{2k}\,\mathrm{d}t = \sum_{n=1}^{\infty}\frac{d_k^2(n)}{n^{2\sigma}} \tag{3.179}$$

where $d_k(n)$ is the coefficient of n^{-s} in the expansion of $\zeta(s)^k$ for $\sigma > 1$. It would be interesting to compare $|Z_\epsilon(U)|^{2k}$ when $\epsilon = \delta/N$ for fixed δ as $N\to\infty$ with the $2k^{\text{th}}$ moment of $|\zeta(\sigma+\mathrm{i}t)|$ when $\sigma = 1/2+\delta/\log T$, as $T\to\infty$ (for some results in the zeta case, see [46]).

3.8 Moments of $\mathfrak{Re}\log Z_U(0)$

The large left deviations in $\mathfrak{Re}\log Z_U(0)$ have a significant impact in the moments of $\mathfrak{Re}\log Z_U(0)$. First, we will calculate an exact formula for each moment, and in doing so provide yet another proof of the central limit theorem of $\mathfrak{Re}\log Z_U(0)/\sigma$, where σ is given by (1.65). Then we will demonstrate the effect of the far left deviations, discussed in section 3.5.1, on the high moments, by asymptotically evaluating their defining integral.

3.8.1 Exact moments

The moment generating function of $\mathfrak{Re}\log Z_U(0)$ is $M_N(s)$, (1.77). Hence the kth moment is

$$M_k = \mathbb{E}\left((\mathfrak{Re}\log Z_U(0))^k\right) = \frac{\mathrm{d}^k}{\mathrm{d}s^k} M_N(s)\bigg|_{s=0} \tag{3.180}$$

And, for $k \geq 1$, the kth cumulant is

$$Q_k = \frac{\mathrm{d}^k}{\mathrm{d}s^k}\{\log M_N(s)\}\bigg|_{s=0} \tag{3.181}$$

$$= \left(1 - 2^{-(k-1)}\right)\sum_{j=1}^{N} \Psi^{(k-1)}(j) \tag{3.182}$$

where $\Psi^{(n)}(\cdot)$ is the nth polygamma function, described in appendix A. (Note that the zeroth cumulant is zero for any random variable).

Now, by their definitions,

$$M_N(s) = \sum_{k=0}^{\infty} \frac{M_k}{k!} s^k \tag{3.183}$$

$$\log M_N(s) = \sum_{r=1}^{\infty} \frac{Q_r}{r!} s^r \tag{3.184}$$

and so,

$$\sum_{k=0}^{\infty} \frac{M_k}{k!} s^k = \exp\left(\sum_{r=1}^{\infty} \frac{Q_r}{r!} s^r\right) \tag{3.185}$$

$$= \prod_{r=1}^{\infty} \exp\left(\frac{Q_r}{r!} s^r\right) \tag{3.186}$$

$$= \prod_{r=1}^{\infty} \sum_{m=0}^{\infty} \frac{1}{m!}\left(\frac{Q_r}{r!} s^r\right)^m \tag{3.187}$$

Thus, picking out the terms containing s^k, we have:

$$M_k = \sum \left(\frac{Q_1}{1!}\right)^{n_1} \left(\frac{Q_2}{2!}\right)^{n_2} \cdots \left(\frac{Q_k}{k!}\right)^{n_k} \frac{k!}{n_1!\, n_2! \cdots n_k!} \tag{3.188}$$

where the sum runs over all non-negative values of n_j $(j = 1, 2, \cdots, k)$ such that $\sum_{j=1}^{k} j n_j = k$ (the number of terms is the partition number of k).

This method is algebraically messy. There is an iterative method, however, which is far easier to implement on a computer:

Lemma 3.19.

$$M_k = \sum_{j=1}^{k} \binom{k-1}{j-1} M_{k-j}\, Q_j \tag{3.189}$$

with $M_0 = 1$.

(Note that the Q_j can be computed relatively easily from (3.182), and stored in a look-up table.)

Proof.

$$\log M_N(s) = \sum_{r=1}^{\infty} \frac{Q_r}{r!} s^r \tag{3.190}$$

so that

$$M_k = \frac{\mathrm{d}^k}{\mathrm{d}s^k} \exp(\log M_N(s))\bigg|_{s=0} \tag{3.191}$$

$$= \frac{\mathrm{d}^{k-1}}{\mathrm{d}s^{k-1}} \left\{ \frac{\mathrm{d}}{\mathrm{d}s} \{\log M_N(s)\} \exp(\log M_N(s)) \right\}\bigg|_{s=0} \tag{3.192}$$

$$= \frac{\mathrm{d}^{k-1}}{\mathrm{d}s^{k-1}} \left\{ \frac{\mathrm{d}}{\mathrm{d}s} \{\log M_N(s)\} M_N(s) \right\}\bigg|_{s=0} \tag{3.193}$$

By Leibnitz's rule,

$$M_k = \sum_{m=0}^{k-1} \binom{k-1}{m} \frac{\mathrm{d}^{m+1}}{\mathrm{d}s^{m+1}} \{\log M_N(s)\} \frac{\mathrm{d}^{k-1-m}}{\mathrm{d}s^{k-1-m}} \{M_N(s)\}\bigg|_{s=0} \tag{3.194}$$

$$= \sum_{m=0}^{k-1} \binom{k-1}{m} Q_{m+1} M_{k-1-m} \tag{3.195}$$

$$= \sum_{j=1}^{k} \binom{k-1}{j-1} M_{k-j}\, Q_j \tag{3.196}$$

□

3.8.2 Comparison with Gaussian moments

The kth moment of a normal random variable with mean zero and variance σ^2 is

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} x^k e^{-x^2/2\sigma^2}\mathrm{d}x = \begin{cases} \frac{k!}{(k/2)!\,2^{k/2}}\sigma^k & \text{if } k \text{ even} \\ 0 & \text{if } k \text{ odd} \end{cases} \tag{3.197}$$

Since the variance of $\mathfrak{Re}\log Z_U(0)$ is Q_2, in order to compare with the Gaussian, we need to compare

$$\frac{M_k}{Q_2^{k/2}} \text{ with } \begin{cases} \frac{k!}{(k/2)!\,2^{k/2}} & \text{if } k \text{ even} \\ 0 & \text{if } k \text{ odd} \end{cases} \tag{3.198}$$

k fixed, $N \to \infty$ limit

Since $Q_2 \sim \frac{1}{2}\log N$, and $Q_j = \mathcal{O}(1)$, for any $j \geq 3$, we need only consider the term in (3.188) with the largest exponent of Q_2. (Note that $Q_1 = 0$).

For k odd, this term is $\left(\frac{Q_2}{2}\right)^{(k-3)/2}\left(\frac{Q_3}{6}\right)\frac{k!}{((k-3)/2)!}$.

For k even, this term is $Q_2^{k/2}\frac{k!}{2^{k/2}(k/2)!}$.

Thus

$$\frac{M_k}{Q_2^{k/2}} \to \begin{cases} \frac{k!}{(k/2)!\,2^{k/2}} & \text{if } k \text{ even} \\ 0 & \text{if } k \text{ odd} \end{cases} \tag{3.199}$$

and so, for k fixed, N tending to infinity, the moments tend to those of the Gaussian.

This is, of course, another proof of the fact $\frac{\mathfrak{Re}\log Z_U(0)}{\sqrt{Q_2}}$ converges in distribution to a standard normal random variable, as found in [2, 69]. (Note that $Q_2 \sim \sigma^2$, with σ given by (1.65)).

$k \to \infty$ and $N \to \infty$

In order to obtain the correct limit for varying k and N, we would need to consider all the elements of the sum. This is very hard, so we only consider the subleading terms in Q_2, thus providing an upper bound for how large k can be with respect to N for the moments to remain Gaussian.

For k odd

At what value of k as a function of N is the $Q_2^{(k-3)/2}Q_3$ term comparable with

$Q_2^{k/2}$? This happens when

$$Q_2^{3/2} \approx \frac{Q_3}{6} \frac{k!}{((k-3)/2)!} 2^{-(k-3)/2}. \tag{3.200}$$

Working to leading order:

$$\frac{3}{2} \log Q_2 = k(\log k - 1) - \frac{k-3}{2} \left(\log \left(\frac{k-3}{2} \right) - 1 \right) - \frac{k-3}{2} \log 2 + \mathcal{O}(\log k) \tag{3.201}$$

$$= \frac{k}{2}(\log k - 1 - \log 2) + \mathcal{O}(\log k). \tag{3.202}$$

Since $Q_2 \sim \log N$, this shows that when $k \geq \log \log N$ and odd,

$$\frac{M_k}{Q_2^{k/2}} \not\to 0 \tag{3.203}$$

For k even

The subleading term is $Q_2^{(k-4)/2} Q_4$. This will be comparable to $Q_2^{k/2}$ when

$$Q_2^2 \approx \frac{Q_4}{4!} \frac{k!}{((k-4)/2)!} 2^{-(k-4)/2} \tag{3.204}$$

or

$$2 \log Q_2 \approx \frac{k}{2}(\log k - 1 - \log 2) + \mathcal{O}(\log k). \tag{3.205}$$

Thus, in a similar fashion to the above, the term contributes when

$$k \geq \log \log N \tag{3.206}$$

This term will dominate the Gaussian term when

$$(\log N)^{k/2-2} \frac{k!}{(k/2-2)!} 2^{-(k/2-2)} \gg \frac{k!}{2^{k/2}(k/2)!} (\log N)^{k/2} \tag{3.207}$$

which will happen for $k \gg \log N$.

Summary

- For k fixed, the kth moment tends to exactly the Gaussian, as $N \to \infty$.
- For $k = \mathcal{O}(\log \log N)$ the other terms are making a non-Gaussian contribution. (This is probably subdominant for even k).
- For $k \gg \log N$ the moments are no longer Gaussian.

3.8.3 Asymptotics for large moments

Writing $p(x)$ for the probability density of $\mathfrak{Re}\log Z_U(0)$, then the kth moment is

$$M_k = \int_{-\infty}^{\infty} x^k p(x)\mathrm{d}x. \tag{3.208}$$

Theorem 3.20. *For N fixed, and $k \gg 1$,*

$$M_k \sim (-1)^k \Gamma(k+1) G^2\left(\tfrac{1}{2}\right) \frac{G(N)\,G(N+1)}{G^2\left(N+\frac{1}{2}\right)}. \tag{3.209}$$

Proof. From §3.3 is is known that, for $x \gg \log N$, $p(-x) \gg p(x)$, which implies that if we hold N fixed and let $k \gg 1$, then

$$M_k \sim \int_{-\infty}^{0} x^k p(x)\mathrm{d}x \tag{3.210}$$

with an absolute error (which we will, from henceforth, ignore) estimated to be much much less than

$$\frac{1}{2}\frac{\Gamma(k+1)}{\Gamma(k/2+1)2^{k/2}}\left(\tfrac{1}{2}\log N\right)^k \tag{3.211}$$

(this comes from calculating $\frac{1}{\sqrt{\pi \log N}}\int_0^\infty x^k e^{-x^2/\log N}\mathrm{d}x >> \int_0^\infty x^k p(x)\mathrm{d}x$).

Recall that during the proof of theorem 3.9 we showed that for $x < 0$, $|x| \gg 1$ with N fixed,

$$p(x) \sim \exp(x) G^2(\tfrac{1}{2}) \frac{G(N)\,G(N+1)}{G^2(N+\frac{1}{2})} \tag{3.212}$$

and so, inserting this $p(x)$ into (3.210) (which is justified because the integral is dominated around the point $k = -x$, which is large and negative when $k \gg 1$) and comparing this with (A.8), we see that

$$M_k \sim (-1)^k \Gamma(k+1) G^2(\tfrac{1}{2}) \frac{G(N)\,G(N+1)}{G^2(N+\frac{1}{2})}, \tag{3.213}$$

as required. □

Remark. By the recurrence relation for the G–function, this can be written as

$$M_k \sim \frac{(-1)^k}{\pi}\Gamma(k+1)\Gamma(N)\prod_{j=1}^{N-1}\frac{\Gamma(j)^2}{\Gamma\left(j+\frac{1}{2}\right)^2}. \tag{3.214}$$

Remark. Allowing N to be large then if $k > \frac{1}{2}\log N$ (c.f. theorem 3.9), one has

$$M_k \sim (-1)^k \Gamma(k+1) G^2\left(\tfrac{1}{2}\right)\, N^{1/4}. \tag{3.215}$$

Chapter 4

The Derivative of the Riemann Zeta Function

The purpose of this chapter is to study the discrete moments of $\zeta'(s)$,

$$J_k(T) = \frac{1}{N(T)} \sum_{0<\gamma_n\leq T} \left|\zeta'\left(\tfrac{1}{2}+\mathrm{i}\gamma_n\right)\right|^{2k}, \tag{4.1}$$

where $N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi e} + \mathcal{O}(\log T)$ is the number of Riemann zeros up to height T.

$J_k(T)$ is clearly defined for all $k \geq 0$, and, on the additional assumption that all the zeros are simple, for all $k < 0$. It has previously been studied by Gonek [45, 46, 47] and Hejhal [56], and is discussed in §2.11 of Odlyzko [78] and §14 of Titchmarsh [90]. Many of the results in this chapter have previously been published in [59].

We consider

$$\mathbb{E}\left\{\frac{1}{N}\sum_{n=1}^{N}\left|Z_U'(\theta_n)\right|^{2k}\right\} \tag{4.2}$$

where $\mathbb{E}$ denotes expectation taken over the CUE, in the hope that it gives information about the universal part of $J_k(T)$, in a similar manner to conjecture 1.4. This is indeed the case; conjecture 4.3 is very similar in form to the Keating-Snaith conjecture.

We show that as $N \to \infty$, $\log|Z_U'(\theta_n)|$ asymptotically has mean $\log N$ and variance $\frac{1}{2}\log N$, and, when properly scaled, tends to a standard normal distribution.

Also, we study the asymptotics of the tails of this distribution, when scaled by a factor much greater than $\sqrt{\log N}$.

4.1 The discrete moments

4.1.1 The random matrix moments

Theorem 4.1. *For k bounded with $\mathfrak{Re}(k) > -3/2$,*

$$\mathbb{E}_N\left\{\frac{1}{N}\sum_{n=1}^{N}\left|Z_U'(\theta_n)\right|^{2k}\right\} = \frac{G^2(k+2)}{G(2k+3)}\frac{G(N+2k+2)G(N)}{N\,G^2(N+k+1)} \tag{4.3}$$

$$\sim \frac{G^2(k+2)}{G(2k+3)}N^{k(k+2)} \text{ as } N\to\infty. \tag{4.4}$$

Proof. Differentiating $Z_U(\theta)$ with respect to θ, we get

$$Z_U'(\theta) = \mathrm{i}\sum_{j=1}^{N} e^{\mathrm{i}(\theta_j-\theta)}\prod_{\substack{m=1\\ m\neq j}}^{N}\left(1-e^{\mathrm{i}(\theta_m-\theta)}\right), \tag{4.5}$$

and so

$$\left|Z_U'(\theta_n)\right| = \prod_{\substack{m=1\\ m\neq n}}^{N}\left|1-e^{\mathrm{i}(\theta_m-\theta_n)}\right|. \tag{4.6}$$

Due to the rotational invariance of Haar measure,

$$\mathbb{E}\left\{\frac{1}{N}\sum_{n=1}^{N}\left|Z_U'(\theta_n)\right|^{2k}\right\} = \mathbb{E}\left\{\left|Z_U'(\theta_N)\right|^{2k}\right\} \tag{4.7}$$

$$= \mathbb{E}\left\{\prod_{m=1}^{N-1}\left|1-e^{\mathrm{i}(\theta_m-\theta_N)}\right|^{2k}\right\} \tag{4.8}$$

This is in the form where lemma 1.9 applies, and so

$$\mathbb{E}_N\left\{\left|Z_U'(\theta_N)\right|^{2k}\right\} = \frac{1}{N}\,\mathbb{E}_{(N-1)}\left\{\left|Z_{\widetilde{U}}(0)\right|^{2k+2}\right\} \tag{4.9}$$

$$= \frac{G^2(k+2)}{G(2k+3)}\frac{G(N+2k+2)G(N)}{NG^2(N+k+1)} \tag{4.10}$$

which is valid for $\mathfrak{Re}(k) > -3/2$. Here $\widetilde{U}$ denotes the $(N-1)\times(N-1)$ unitary matrix with eigenangles $\theta_1,\ldots,\theta_{N-1}$. The evaluation of $\mathbb{E}_{(N-1)}\left\{\left|Z_{\widetilde{U}}(0)\right|^{2k+2}\right\}$ is given by (1.77).

Assuming k to be bounded, then as $N\to\infty$, the asymptotics for G, (A.3), imply

$$\frac{G(N+2k+2)G(N)}{NG^2(N+k+1)} = N^{k(k+2)}\left(1+\mathcal{O}\left(\frac{1}{N}\right)\right). \tag{4.11}$$

This proves theorem 4.1. □

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

$$\frac{G^2(k+2)}{G(2k+3)} = \prod_{j=0}^{k} \frac{j!}{(k+1+j)!}. \tag{4.12}$$

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

$$\frac{G^2(k+2)}{G(2k+3)} = \frac{\exp\left(3\zeta'(-1) + \log\pi - \frac{11}{12}\log 2 + k\log\pi - 3k\log 2 - 2k^2\log 2\right)}{G\left(k+\frac{3}{2}\right)G\left(k+\frac{5}{2}\right)}, \tag{4.13}$$

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

The existence of a pole at $k = -3/2$ in (4.10) means that $\mathbb{E}_N\left\{|Z_U'(\theta_N)|^{2k}\right\}$ diverges (for any $N \geq 2$) when $\mathfrak{Re}(k) \leq -3/2$. Its analytic continuation into this region is given by (4.10).

4.1.2 A heuristic analysis of $J_k(T)$

Recall (3.131), that for $x > 0$,

$$P(T,x) = \frac{1}{T}\,\text{meas}\left\{t : 0 \leq t \leq T\ ,\ \log\left|\zeta\left(\tfrac{1}{2}+\text{i}t\right)\right| \leq -x\right\}, \tag{4.14}$$

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

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

$$\left|\zeta\left(\tfrac{1}{2}+\text{i}\gamma_n+\text{i}\epsilon\right)\right| = |\epsilon|\left|\zeta'\left(\tfrac{1}{2}+\text{i}\gamma_n\right)\right| + \mathcal{O}_T\left(\epsilon^2\right), \tag{4.15}$$

which equals e^{-x} when

$$|\epsilon| = \frac{e^{-x}}{\left|\zeta'\left(\frac{1}{2}+\text{i}\gamma_n\right)\right|} + \mathcal{O}_T\left(e^{-2x}\right) \tag{4.16}$$

and so the length of each region is $2|\epsilon| + \mathcal{O}_T(\epsilon^2)$.

Thus,

$$\lim_{x\to\infty} e^x P(T,x) = \frac{2}{T}\sum_{0<\gamma_n\leq T}\left|\zeta'\left(\tfrac{1}{2}+\text{i}\gamma_n\right)\right|^{-1}. \tag{4.17}$$

A different evaluation of $P(T,x)$ comes from conjecture 3.15, which says

$$\lim_{x\to\infty} e^x P(T,x) \sim \left(\log\frac{T}{2\pi}\right)^{1/4} \exp\left(3\zeta'(-1) + \tfrac{1}{12}\log 2 - \tfrac{1}{2}\log\pi\right) a\left(-\tfrac{1}{2}\right). \tag{4.18}$$

Combining (4.17) and (4.18), we obtain

Conjecture 4.2.

$$J_{-1/2}(T) \sim \exp\left(3\zeta'(-1) + \tfrac{1}{12}\log 2 + \tfrac{1}{2}\log\pi\right) a\left(-\tfrac{1}{2}\right)\left(\log\frac{T}{2\pi}\right)^{-3/4}. \tag{4.19}$$

Note that for $k=-1/2$, the random matrix moment, (4.4), is asymptotic to

$$\exp\left(3\zeta'(-1) + \tfrac{1}{12}\log 2 + \tfrac{1}{2}\log\pi\right) N^{-3/4} \tag{4.20}$$

as $N\to\infty$. Since $a(k)$ is exactly the zeta-function-specific term in conjecture 1.4, and $N = \log\frac{T}{2\pi}$, this in turn leads us to

Conjecture 4.3. *For bounded k such that $\mathfrak{Re}(k) > -3/2$,*

$$J_k(T) \sim \frac{G^2(k+2)}{G(2k+3)} a(k)\left(\log\frac{T}{2\pi}\right)^{k(k+2)} \tag{4.21}$$

as $T\to\infty$.

Comparison with known results

If the tails of the distribution (4.42) are sufficiently small, one might expect [46, 56]

$$J_k(T) \asymp (\log T)^{k(k+2)}. \tag{4.22}$$

We show in §4.2.2 that the singularity at $k=-3/2$ in (4.19) comes from a large left tail of the distribution of $\log|Z'(\theta_1)|$.

Under RH Gonek [45] has proved that $J_1(T) \sim \frac{1}{12}(\log T)^3$. Under the additional assumption that all the non-trivial zeros are simple, he has conjectured in [46] that $J_{-1}(T) \sim \frac{6}{\pi^2}(\log T)^{-1}$.

Also, extending a theorem due to Conrey, Ghosh and Gonek (recorded in this thesis as theorem 5.4) beyond its range of (proven) applicability, conjecture 5.5.1 states that $J_2(T) \sim \frac{1}{1440\pi^2}\left(\log\frac{T}{2\pi}\right)^8$.

We observe that conjecture 4.3 agrees with all these results.

4.1.3 Discussion on the 'pole' at $k = -3/2$

Due to the divergence of the random matrix average, conjecture 4.3 is restricted to $2k > -3$. In this section, we argue that this restriction is necessary.

For k negative, but $|k|$ 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 [47], in a talk at the Mathematical Sciences Research Institute (MSRI), Berkeley, in June 1999, defined

$$\Theta = \inf\left\{\theta : |\zeta'\left(\tfrac{1}{2} + \mathrm{i}\gamma_n\right)|^{-1} = \mathcal{O}\left(|\gamma_n|^{\theta}\right)\ \forall\, n\right\}. \tag{4.23}$$

He observed that RH implies $\Theta \geq 0$, and that $\Theta \leq 1$ if the averaged Mertens hypothesis holds, that is if

$$\int_1^X \frac{1}{x^2}\left(\sum_{n\leq x}\mu(n)\right)^2 \mathrm{d}x = \mathcal{O}(\log X), \tag{4.24}$$

where $\mu(n)$ is the Möbius function, that is

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^k & \text{if } n \text{ is the product of } k \text{ distinct prime factors} \\ 0 & \text{otherwise} \end{cases} \tag{4.25}$$

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

$$|\zeta'\left(\tfrac{1}{2} + \mathrm{i}\gamma_n\right)|^{-1} > |\gamma_n|^{\Theta-\epsilon}. \tag{4.26}$$

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

$$J_k(T) > \frac{1}{N(T)}|\zeta'\left(\tfrac{1}{2} + \mathrm{i}\gamma\right)|^{2k} \tag{4.27}$$

$$> \frac{2\pi}{T\log T}T^{-2k(\Theta-\epsilon)}. \tag{4.28}$$

If $\Theta > 0$, then

$$\frac{2\pi}{T\log T}T^{-2k(\Theta-\epsilon)} \gg (\log T)^{k(k+2)} \tag{4.29}$$

when

$$2k < -\frac{1}{\Theta}, \tag{4.30}$$

implying that the conjectured scaling (4.22) is too small for $2k < -\frac{1}{\Theta}$. It follows from theorem 4.1 that the analogue of (4.22) for $Z'_U(\theta_1)$ fails for $2k \leq -3$, which implies, via conjecture 4.3, that $\Theta = 1/3$. This is precisely the value conjectured by Gonek [47], and is in line with the fact that Montgomery's pair correlation conjecture, (1.22), suggests that $\Theta \geq 1/3$.

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

$$J_k(T) = \Omega\left(T^{2|k|\Theta-1-\epsilon}\right). \tag{4.31}$$

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

$$J_k(T) = \mathcal{O}\left(T^{2|k|\Theta+\epsilon}\right) \tag{4.32}$$

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

Remark. If all the zeros are simple, then for $k \leq -\frac{3}{2}$, $J_k(T)$ 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 might still predict the universal behaviour.

4.2 The distribution of $\log|Z'(U, \theta_1)|$

4.2.1 The central limit theorem

Theorem 4.4. *For $a < b$,*

$$\lim_{N\to\infty} \mathbb{P}\left\{ \frac{\log\left|\frac{Z'_U(\theta_1)}{N\exp(\gamma-1)}\right|}{\sqrt{\frac{1}{2}(\log N + 3 + \gamma - \pi^2/2)}} \in (a, b) \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, \mathrm{d}x. \tag{4.33}$$

Proof. By theorem 4.1,

$$F_N(\lambda) := \mathbb{E}\left|Z'_U(\theta_1)\right|^\lambda \tag{4.34}$$

$$= \frac{G^2\left(2 + \frac{1}{2}\lambda\right)}{G(3+\lambda)} \frac{G(N+2+\lambda)G(N)}{G^2\left(N + 1 + \frac{1}{2}\lambda\right) N}. \tag{4.35}$$

The cumulants can now be calculated as

$$C_n = \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}\left\{\log F_N(\lambda)\right\}\Big|_{\lambda=0} \tag{4.36}$$

$$= \frac{1}{2^{n-1}}\Phi^{(n-1)}(2) - \Phi^{(n-1)}(3) + \Phi^{(n-1)}(N+2) - \frac{1}{2^{n-1}}\Phi^{(n-1)}(N+1) \tag{4.37}$$

where $\Phi^{(n)}(x) = \frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}}\log G(x)$. Expanding these for large N, the mean and the variance of $\log|Z'_U(\theta_1)|$ are

$$C_1 \sim \log N + \gamma - 1, \tag{4.38}$$

$$C_2 \sim \tfrac{1}{2}(\log N + \gamma + 3 - \pi^2/2), \tag{4.39}$$

with the higher cumulants being $C_n = \mathcal{O}(1)$.

Therefore, setting $\lambda = x/\sqrt{C_2}$,

$$\lim_{N\to\infty}\mathbb{E}\exp\left(\frac{x\left(\log|Z'_U(\theta_1)| - C_1\right)}{\sqrt{\frac{1}{2}(\log N + \gamma + 3 - \pi^2/2)}}\right) = e^{x^2/2} \tag{4.40}$$

which is the moment generating function of a standard normal random variable, and by theorem 1.5 this is sufficient to prove the theorem. □

Theorem 4.4 can be compared with the limiting distribution of $\log|\zeta'(\gamma_n)|$.

Theorem 4.5. (Hejhal [56]). *If one assumes RH and the existence of an α such that*

$$\limsup_{T\to\infty}\frac{1}{N(2T)-N(T)}\left|\left\{n\,:\,T\le\gamma_n\le 2T, 0\le\gamma_{n+1}-\gamma_n\le\frac{c}{\log T}\right\}\right| \le Mc^{\alpha} \tag{4.41}$$

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

$$\lim_{T\to\infty}\frac{1}{N(2T)-N(T)}\left|\left\{n\,:\,T\le\gamma_n\le 2T, \frac{\log\left|\frac{\zeta'(1/2+\mathrm{i}\gamma_n)}{\frac{1}{2\pi}\log\frac{\gamma_n}{2\pi}}\right|}{\sqrt{\frac{1}{2}\log\log T}}\in(a,b)\right\}\right| = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}\,\mathrm{d}x. \tag{4.42}$$

Recalling (1.72), that $N = \log\frac{T}{2\pi}$, theorem 4.4 is in line with Hejhal's distribution theorem, (4.42). (Note that the $\mathcal{O}(1)$ differences in the mean and variance are subdominant in the large N, large T limit).

Odlyzko [78] found numerically that, around the 10^{20}th zero, $\log|\zeta'|$ had mean 3.35 and variance 1.14. Compare this to the leading order asymptotic prediction in (4.42) of 1.91 and 1.89, and the above random matrix theory prediction of 3.33 and 1.21 respectively.

4.2.2 Asymptotics for deviations to the left

By the Fourier inversion theorem, the probability density function of $\log\left|\frac{Z_U'(\theta_1)}{\exp(C_1)}\right|$ exists and is given by

$$p(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\mathrm{i}y(t+C_1)}F_N(\mathrm{i}y)\,\mathrm{d}y \tag{4.43}$$

(so that for any measurable set A, $\mathbb{P}\left\{\log|Z_U'(\theta_1)| - C_1 \in A\right\} = \int_A p(t)\,\mathrm{d}t$).

Theorem 4.4 implies that for $|t| << \sqrt{\log N}$,

$$p(t) \sim \frac{1}{\sqrt{\pi\log N}}\exp\left(\frac{-t^2}{\log N}\right). \tag{4.44}$$

This section calculates $p(t)$ for $t < -(3/2+\epsilon)\log N$, in a very similar manner to that used in section 3.5.1.

Theorem 4.6. *If* $\limsup_{N\to\infty}\frac{t}{\log N} < -3/2$,

$$p(t) \sim e^{3t}N^{9/4}e^{3\gamma-3}G^2(1/2) \tag{4.45}$$

Proof. Integrating

$$\frac{1}{2\pi}\oint_C e^{-\mathrm{i}y(t+C_1)}F_N(\mathrm{i}y)\,\mathrm{d}y \tag{4.46}$$

over the rectangle C 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

$$p(t) = \mathrm{i}\operatorname*{Res}_{y=3\mathrm{i}}\left\{e^{-\mathrm{i}y(t+C_1)}F_N(\mathrm{i}y)\right\} + E, \tag{4.47}$$

where

$$E = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{(-\mathrm{i}y+3+\epsilon)(t+C_1)}F_N(\mathrm{i}y-3-\epsilon)\,\mathrm{d}y. \tag{4.48}$$

Asymptotic analysis shows that

$$|E| \le \frac{1}{2\pi}e^{(3+\epsilon)(t+C_1)}\int_{-\infty}^{\infty}|F_N(\mathrm{i}y-3-\epsilon)|\,\mathrm{d}y \tag{4.49}$$

$$\sim \frac{1}{\sqrt{\pi}}\left|\frac{G^2\left(\frac{1}{2}-\frac{1}{2}\epsilon\right)}{G(-\epsilon)}\right| e^{(3+\epsilon)(\gamma-1)}e^{(3+\epsilon)t}N^{9/4+3\epsilon/2+\epsilon^2/4}(\log N)^{-1/2}, \tag{4.50}$$

and that

$$\mathrm{i}\operatorname*{Res}_{y=3\mathrm{i}}\left\{e^{-\mathrm{i}y(t+C_1)}F_N(\mathrm{i}y)\right\} \sim e^{3t}N^{9/4}e^{3\gamma-3}G^2(1/2). \tag{4.51}$$

If $\limsup_{N\to\infty}\frac{t}{\log N} < -3/2$, then choosing

$$0 < \epsilon < \min\left\{-6 - 4\limsup_{N\to\infty}\frac{t}{\log N}\ ,\ 1\right\} \tag{4.52}$$

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

$$p(t) \sim e^{3t} N^{9/4} e^{3\gamma-3} G^2(\tfrac{1}{2}) \tag{4.53}$$

if $\limsup_{N\to\infty} \frac{t}{\log N} < -3/2$. □

Remark. Note that the divergence of $\int_{\mathcal{U}(N)} |Z'_U(\theta_1)|^{2k}\, \mathrm{d}\mu_N$ for $k \leq -3/2$ is due to the e^{3t} term in (4.45).

4.2.3 Large deviations

In this section we study the tails of the distribution of $\log|Z'_U(\theta_1)|$, beyond the scope of the central limit theorem, using large deviation theory. (In fact, we consider the random variable $(\log|Z'_U(\theta_1)| - C_1)$, since this has zero mean). Theorem 4.6 is a much more refined large deviation result, but it only applies for left deviations at scalings $A \leq \log N$.

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

$$R_N^A = \frac{\log\left|\frac{Z'_U(\theta_1)}{\exp(C_1)}\right|}{A(N)}, \tag{4.54}$$

where $A(N)$ is a given function, much greater than $\sqrt{C_2}$, where C_1 and C_2 are given by (4.38) and (4.39) respectively.

Denote the logarithmic moment generating function of R_N^A by

$$\Lambda_N(\lambda) = \log \mathbb{E}\, e^{\lambda R_N^A} \tag{4.55}$$

$$= \begin{cases} \log F\left(\frac{\lambda}{A(N)}, N\right) - \frac{\lambda}{A(N)} C_1 & \text{for } \frac{\lambda}{A(N)} > -3 \\ \infty & \text{for } \frac{\lambda}{A(N)} \leq -3 \end{cases} \tag{4.56}$$

(recall $F_N(\lambda)$ is given by (4.35)).

In order to apply theorem 1.7, we need to know the leading-order asymptotics of $\Lambda_N(B\lambda)$. Writing $\eta(N) = \frac{B(N)}{A(N)}\lambda$ for simplicity, then (4.38) and (A.3) imply that

as $N \to \infty$, for $\eta > -3$,

$$\begin{aligned}\log F_N(\eta) - \eta C_1 = \tfrac{1}{2}(N+\eta+1)^2 \log(N+\eta) + \tfrac{1}{2}(N-1)^2 \log N \\ - (N+\tfrac{1}{2}\eta)^2 \log(N+\tfrac{1}{2}\eta) - \tfrac{1}{12}\log(N+\eta) - \tfrac{13}{12}\log N \\ + \tfrac{1}{6}\log(N+\tfrac{1}{2}\eta) - \eta \log N - \gamma\eta - \tfrac{3}{8}\eta^2 - \log G(3+\eta) \\ + 2\log G\left(2+\tfrac{1}{2}\eta\right) + \mathcal{O}\left(\frac{1}{N}\right), \quad (4.57)\end{aligned}$$

uniform in η for $\eta > -3+\epsilon$.

Theorem 4.7. *For $\sqrt{\log N} \ll A \ll N$, R_N^A satisfies a large deviation principle at speed*

$$B = \frac{A^2}{-W_{-1}(-A/N)}. \quad (4.58)$$

For $\sqrt{\log N} \ll A \ll \log N$ the rate function is $I(x) = x^2$, for $A = \log N$ the rate function is

$$I(x) = \begin{cases} x^2 & \text{for } x \geq -3/2 \\ 3|x| - 9/4 & \text{for } x \leq -3/2 \end{cases}, \quad (4.59)$$

and for $\log N \ll A \ll N$ the rate function is

$$I(x) = \begin{cases} x^2 & \text{for } x \geq 0 \\ 0 & \text{for } x \leq 0 \end{cases}. \quad (4.60)$$

As in section 3.3, define $\chi(N)$ to be so that $\frac{B(N)}{A(N)} = \frac{N}{\chi(N)}$ (with $\chi(N) \gg 1$ as $N \to \infty$), then

$$\Lambda_N(B\lambda) = \begin{cases} \frac{1}{4}\lambda^2 \frac{N^2 \log \chi}{\chi^2} + \mathcal{O}_\lambda\left(\frac{N^2}{\chi^2}\right) & \text{if } \lambda N/\chi > -3 \\ \infty & \text{if } \lambda N/\chi \leq -3 \end{cases} \quad (4.61)$$

A non-trivial limit exists for $\frac{1}{B}\Lambda_N(B\lambda)$ if $\frac{N^2 \log \chi}{B\chi^2} = 1$, which happens when

$$B = \frac{A^2}{-W_{-1}\left(-\frac{A}{N}\right)}. \quad (4.62)$$

With such a B, set $\delta := \liminf_{N\to\infty} \chi/N = \liminf_{N\to\infty} A/B$, and define

$$\Lambda(\lambda) := \lim_{N\to\infty} \frac{1}{B}\Lambda_N(B\lambda) \quad (4.63)$$

$$= \begin{cases} \frac{1}{4}\lambda^2 & \text{if } \lambda > -3\delta \\ \infty & \text{otherwise} \end{cases} \quad (4.64)$$

- If $\sqrt{\log N} \ll A \ll \log N$, then $\delta = \infty$ and $\Lambda(\lambda) = \frac{1}{4}\lambda^2$ for all λ. Theorem 1.7 applies in this case, and $I(x) = \sup_{\lambda\in\mathbb{R}}\left\{\lambda x - \frac{1}{4}\lambda^2\right\} = x^2$.

- If $A = \log N$ then $B = \log N$ and $\chi = N$, and so $\delta = 1$. Hence

$$\Lambda(\lambda) = \begin{cases} \frac{1}{4}\lambda^2 & \text{if } \lambda > -3 \\ \infty & \text{if } \lambda \leq -3 \end{cases} \tag{4.65}$$

  so

$$I(x) = \begin{cases} x^2 & \text{if } x \geq -\frac{3}{2} \\ -3x - \frac{9}{4} & \text{if } x \leq -\frac{3}{2} \end{cases} \tag{4.66}$$

  Remark. Note that in this case theorem 1.7 only gives the upper bound, (1.100), on the probabilities for $x < -3/2$. However, we know from theorem 4.6 that for $x < -3/2$,

$$\mathbb{P}\left\{R_N^{\log N} \leq x\right\} \sim \tfrac{1}{3}e^{3\gamma-3}G^2(\tfrac{1}{2})N^{3x+9/4}. \tag{4.67}$$

- If $\log N \ll A \ll N$, then $\delta = 0$ and so

$$\Lambda(\lambda) = \begin{cases} \frac{1}{4}\lambda^2 & \text{if } \lambda \geq 0 \\ \infty & \text{if } \lambda < 0 \end{cases} \tag{4.68}$$

  If $x > 0$, theorem 1.7 implies that $I(x) = x^2$, and if $x < 0$ then theorem 4.6 implies that $I(x) = 0$.

This completes the proof of the theorem. □

Remark. If $A \gg \log N$, setting the speed $B = A$ makes $\delta = 1$, and thus

$$\lim_{N\to\infty} \frac{1}{B}\Lambda_N(B\lambda) = \begin{cases} 0 & \text{for } \lambda > -3 \\ \infty & \text{for } \lambda \leq -3 \end{cases} \tag{4.69}$$

the convex dual of this is

$$\Lambda^*(x) = \begin{cases} \infty & \text{for } x > 0 \\ -3x & \text{for } x \leq 0 \end{cases} \tag{4.70}$$

and this can be seen to be the correct rate function at this speed and scaling (even though theorem 1.7 doesn't apply) by theorem 4.6, since for $A \gg \log N$ and $x < 0$ it says that,

$$\mathbb{P}\left\{R_N^A \leq x\right\} \sim \tfrac{1}{3}e^{3\gamma-3}G^2(\tfrac{1}{2})N^{9/4}\exp(3Ax). \tag{4.71}$$

Theorem 4.8. *At scaling $A = N$ and speed $B = N^2$,*

$$I(x) = \begin{cases} I_c(x) & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases} \tag{4.72}$$

where

$$I_c(x) = \sup_{\lambda > 0} \Big\{ \lambda x - \tfrac{1}{2}(1+\lambda)^2 \log(1+\lambda) + (1+\tfrac{1}{2}\lambda)^2 \log(1+\tfrac{1}{2}\lambda) + \tfrac{1}{4}\lambda^2 \log 2\lambda \Big\}. \tag{4.73}$$

Proof. When $A = N$ and $B = N^2$, $\eta = N$ in (4.57), and so

$$\Lambda(\lambda) = \lim_{N\to\infty} \frac{1}{N^2} \Lambda_N(N^2\lambda) \tag{4.74}$$

$$= \tfrac{1}{2}(1+\lambda)^2 \log(1+\lambda) - (1+\tfrac{1}{2}\lambda)^2 \log(1+\tfrac{1}{2}\lambda) - \tfrac{1}{4}\lambda^2 \log 2\lambda \tag{4.75}$$

for $\lambda \geq 0$, and $\Lambda(\lambda) = \infty$ otherwise. Therefore, if $x \geq 0$, theorem 1.7 applies, and we have

$$I(x) = I_c(x). \tag{4.76}$$

For $x < 0$, theorem 4.6 implies that $I(x) = 0$ and this completes the proof. □

Remark. Assuming $x \geq 0$, the supremising λ in $I_c(x)$ occurs when

$$x = \tfrac{1}{2}\lambda \log\left(\frac{(\lambda+1)^2}{\lambda(\lambda+2)}\right) + \log\left(\frac{\lambda+1}{\lambda+2}\right) + \log 2. \tag{4.77}$$

Note that the right hand side is an increasing function of λ (for $\lambda > 0$) bounded between 0 and $\log 2$. This means, for $x \geq \log 2$ the supremum is ∞, which is in line with the fact that $\log |Z_U'(\theta_1)| \leq (N-1)\log 2$ for all $U \in \mathcal{U}(N)$.

Conclusion

For deviations to the right, we must take $x > 0$:

Scaling $A(N)$	Speed $B(N)$	Rate function $\Lambda^*(x)$, $x > 0$
$\sqrt{\log N} \ll A \ll N$	$\frac{A^2}{-W_{-1}(-\frac{A}{N})}$	x^2
$A = N$	N^2	$\begin{cases} I_c(x) & \text{if } 0 \le x \le \log 2 \\ \infty & \text{if } x \ge \log 2 \end{cases}$

Table 4.1: Deviations to the right

For deviations to the left, we must take $x < 0$:

Scaling $A(N)$	Speed $B(N)$	Rate function $\Lambda^*(x)$, $x < 0$
$\sqrt{\log N} \ll A \ll \log N$	$\frac{A^2}{\log N}$	x^2
$A = \log N$	$\log N$	$\begin{cases} x^2 & \text{if } -\frac{3}{2} \le x \le 0 \\ 3\lvert x\rvert - \frac{9}{4} & \text{if } x \le -\frac{3}{2} \end{cases}$
$A \gg \log N$	A	$3\lvert x\rvert$

Table 4.2: Deviations to the left

Remark. Note that the LDP for the deviations to the right is identical to that found for $\mathfrak{Re}\log Z_U(\theta)$ in §3.3. The LDP for deviations to the left is very similar, but the rate function there is linear for $x < -1/2$ rather than $x < -3/2$, though still at scaling $A = \log N$. That the rate function of $\log|Z'(\theta_1)|$ becomes linear further to the left than the rate function for $\log|Z(0)|$ is consistent with the observation that the (scaled) value distribution $\log|\zeta'(1/2 + \mathrm{i}\gamma_n)|$ is closer to the standard normal curve than $\log|\zeta(1/2 + \mathrm{i}t)|$ in Odlyzko's numerical data (see page 55 of [78]).

4.3 Other unitary ensembles

The other unitary ensembles — the COE ($\beta = 1$) and the CSE ($\beta = 4$) — 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 are [74]

$$\mathrm{d}\mu_N^\beta = \frac{((\beta/2)!)^N}{(N\beta/2)!(2\pi)^N} \prod_{1\le j<k\le N} \left|e^{\mathrm{i}\theta_j} - e^{\mathrm{i}\theta_k}\right|^\beta \prod_{n=1}^N \mathrm{d}\theta_n. \tag{4.78}$$

Using Selberg's integral, lemma 1.8, Keating and Snaith [69] found that

$$\int_{\mathcal{U}(N)} |Z_U(\theta)|^s \,\mathrm{d}\mu_N^\beta = \prod_{j=0}^{N-1} \frac{\Gamma(1+j\beta/2)\Gamma(1+s+j\beta/2)}{(\Gamma(1+s/2+j\beta/2))^2} \tag{4.79}$$

$$= M_N(\beta, s). \tag{4.80}$$

As in §4.1.1, we find that

$$\int_{\mathcal{U}(N)} \frac{1}{N}\sum_{n=1}^N |Z_U'(\theta_n)|^s \,\mathrm{d}\mu_N^\beta = \int_{\mathcal{U}(N)} |Z_U'(\theta_1)|^s \,\mathrm{d}\mu_N^\beta \tag{4.81}$$

$$= (\beta/2)!\frac{((N-1)\beta/2)!}{(N\beta/2)!} M_{N-1}(\beta, s+\beta). \tag{4.82}$$

Calculating the cumulants,

$$C_1^\beta = \sum_{j=0}^{N-2} \Psi(1+\beta+j\beta/2) - \Psi(1+\beta/2+j\beta/2) \tag{4.83}$$

$$= \log N + \mathcal{O}(1) \tag{4.84}$$

$$C_2^\beta = \sum_{j=0}^{N-2} \Psi^{(1)}(1+\beta+j\beta/2) - \tfrac{1}{2}\Psi^{(1)}(1+\beta/2+j\beta/2) \tag{4.85}$$

$$= \frac{1}{\beta}\log N + \mathcal{O}(1) \tag{4.86}$$

$$C_n^\beta = \sum_{j=0}^{N-2} \Psi^{(n-1)}(1+\beta+j\beta/2) - 2^{-(n-1)}\Psi^{(n-1)}(1+\beta/2+j\beta/2) \tag{4.87}$$

$$= \mathcal{O}(1) \text{ for } n \ge 3 \tag{4.88}$$

which shows that

$$\lim_{N\to\infty} \mathbb{P}_\beta \left\{ \frac{\log|Z_U'(\theta_1)| - C_1^\beta}{\sqrt{C_2^\beta}} \in (a,b) \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} \,\mathrm{d}x. \tag{4.89}$$

Chapter 5

A Unified Approach for Moments of $|Z_U(\theta)|$ and $|Z'_U\theta_1)|$

In this chapter we will use RMT to study the large N asymptotics of

$$\mathbb{E}_N\left\{\frac{1}{N}\sum_{n=1}^{N}\left|Z_U\left(\theta_n+\frac{2\pi\alpha}{N}\right)\right|^{2k}\right\}=\mathbb{E}_N\left\{\left|Z_U\left(\theta_1+\frac{2\pi\alpha}{N}\right)\right|^{2k}\right\} \tag{5.1}$$

and use those to make a conjecture on the asymptotic behaviour of

$$\frac{1}{N(T)}\sum_{0<\gamma_n\leq T}\left|\zeta\left(\tfrac{1}{2}+\mathrm{i}\left(\gamma_n+\alpha/L\right)\right)\right|^{2k} \tag{5.2}$$

where $L=\frac{1}{2\pi}\log\frac{T}{2\pi}$ is the mean density of zeros at height T. In particular we will show that one can recover both a variant of the Keating-Snaith conjecture (conjecture 1.4) by letting $\alpha\to\infty$, and also conjecture 4.3 (the discrete derivatives conjecture) by letting $\alpha\to 0$. So this chapter can be thought of as unifying these two conjectures. These results can also be used to calculate a lower bound for how large a gap can exist between consecutive scaled zeros of the zeta function.

5.1 The random matrix calculation

$$\mathbb{E}_N\left\{\left|Z_U\left(\theta_N+\frac{y}{N}\right)\right|^{2k}\right\}=\mathbb{E}_N\left\{\left|1-e^{-\mathrm{i}y/N}\right|^{2k}\prod_{n=1}^{N-1}\left|1-e^{\mathrm{i}(\theta_n-\theta_N-y/N)}\right|^{2k}\right\} \tag{5.3}$$

which is in a form where lemma 1.9 applies, and so

$$\mathbb{E}_N\left\{\left|Z_U\left(\theta_N+\frac{y}{N}\right)\right|^{2k}\right\}=\frac{1}{N}\left|2\sin\left(\frac{y}{2N}\right)\right|^{2k}\mathbb{E}_{N-1}\left\{\left|Z_{\widetilde{U}}(0)\right|^2\left|Z_{\widetilde{U}}\left(\frac{y}{N}\right)\right|^{2k}\right\}, \tag{5.4}$$

where $Z_{\widetilde{U}}$ is the characteristic polynomial of an $(N-1)\times(N-1)$ unitary matrix. The first theorem deals with the right-hand side of this equation.

Theorem 5.1. *For integer k,*

$$\mathbb{E}_{N-1}\left\{\left|Z_{\widetilde{U}}(0)\right|^2\left|Z_{\widetilde{U}}(\beta)\right|^{2k}\right\}=\frac{G^2(k+2)}{G(2k+3)}\frac{G(N+2k+2)G(N)}{G^2(N+k+1)}\times$$
$$\times(x_1y_2-x_2y_1)e^{-\mathrm{i}(N-1)\beta} \tag{5.5}$$

where

$$x_1=\sum_{n=0}^{N-1}\frac{(2k)!}{(2k+n)!}\binom{k+n-1}{n}\frac{(N-1)!}{(N-n-1)!}\left(e^{\mathrm{i}\beta}-1\right)^n \tag{5.6}$$

$$x_2=\sum_{n=1}^{N}\frac{(2k)!}{(2k+n)!}\binom{k+n-1}{n}n\frac{(N-1)!}{(N-n)!}\left(e^{\mathrm{i}\beta}-1\right)^n \tag{5.7}$$

$$y_1=\sum_{n=0}^{N-2}\frac{(2k+1)!}{(2k+1+n)!}\binom{k+n}{n}\frac{(N-1)!}{(N-n-2)!}\left(e^{\mathrm{i}\beta}-1\right)^n \tag{5.8}$$

$$y_2=\sum_{n=0}^{N-1}\frac{(2k+1)!}{(2k+1+n)!}\binom{k+n}{n}(n+1)\frac{(N-1)!}{(N-n-1)!}\left(e^{\mathrm{i}\beta}-1\right)^n \tag{5.9}$$

Proof By rotation invariance of Haar measure and a little trivial rearrangement,

$$\mathbb{E}_{N-1}\left\{\left|Z_{\widetilde{U}}(0)\right|^2\left|Z_{\widetilde{U}}(\beta)\right|^{2k}\right\}=\mathbb{E}_{N-1}\left\{\left|Z_{\widetilde{U}}(0)\right|^{2k}\left|Z_{\widetilde{U}}(\beta)\right|^2\right\} \tag{5.10}$$

$$=e^{-\mathrm{i}(N-1)\beta}\,\mathbb{E}_{N-1}\left\{Z_{\widetilde{U}}(0)^{k+1}Z^*_{\widetilde{U}}(0)^{k-1}Z^*_{\widetilde{U}}(\beta)^2\right\}. \tag{5.11}$$

The right-hand side of this is, in the notation of (6.19), $e^{-\mathrm{i}(N-1)\beta}T_{k+1,2}$ (with N replaced by $N-1$, due to the average being over $\mathcal{U}(N-1)$ rather than $\mathcal{U}(N)$). Lemma 6.3 (which is more general than we need here) evaluates $T_{k+1,2}$. □

Remark Putting $\beta=0$ into theorem 5.1 we see that

$$x_1y_2-x_2y_1=1, \tag{5.12}$$

and thus,

$$\mathbb{E}_{N-1}\left\{\left|Z_{\widetilde{U}}(0)\right|^{2k+2}\right\}=\frac{G^2(k+2)}{G(2k+2)}\frac{G(N+2k+2)G(N)}{G^2(N+k+1)} \tag{5.13}$$

$$=M_{N-1}(2k+2), \tag{5.14}$$

as it should.

Remark. Although not obvious from the definitions of x_1, x_2, y_1, y_2 it must be that $(x_1y_2 - x_2y_1)e^{-\mathrm{i}(N-1)\beta}$ is a real, even function of β, since the left-hand side of (5.5) is.

We are now in a position to state and prove our main theorem of this chapter:

Theorem 5.2. *For integer k, and $|y| \ll N$,*

$$\mathbb{E}_N\left\{\left|Z_U\left(\theta_1 + \frac{y}{N}\right)\right|^{2k}\right\} \sim \frac{G^2(k+1)}{G(2k+1)}F_k(y)N^{k^2} \tag{5.15}$$

where

$$F_k(y) = \frac{A(y)\cos y + B(y)\sin y + C(y)}{y^{2k}} \tag{5.16}$$

where $A(y), B(y)$ and $C(y)$ are all polynomials in y with integer coefficients depending on k only, and are given by

$$A(y) = \sum_{n=0}^{k-1} 2(-1)^n b_{2n} y^{2n} \tag{5.17}$$

$$B(y) = \sum_{n=1}^{k-1} 2(-1)^{n-1} b_{2n-1} y^{2n-1} \tag{5.18}$$

$$C(y) = \sum_{n=0}^{k} (-1)^n c_{2n} y^{2n} \tag{5.19}$$

with

$$b_r = \sum_{n=0}^{k} \frac{(2k-n-1)!}{k!}\frac{(2k+n-r-1)!}{(k-1)!}\binom{k}{n}\binom{k}{r+1-n}(2n-r-1) \tag{5.20}$$

and

$$c_{2r} = \sum_{n=1}^{\min\{r,k-1\}} \frac{2b_{2n-1}}{(2r-2n+1)!} - \sum_{n=0}^{\min\{r,k-1\}} \frac{2b_{2n}}{(2r-2n)!} \tag{5.21}$$

Proof. Set

$$f(y) := \sum_{n=0}^{\infty} \frac{(k-1)!}{(2k+n)!}\binom{k+n-1}{n} y^n. \tag{5.22}$$

Replace β with y/N in the definitions of x_1, x_2, y_1 and y_2 from theorem 5.1. Since by assumption $|y| \ll N$ as $N \to \infty$, we have

$$\frac{(N-1)!}{(N-n-1)!}\left(e^{\mathrm{i}y/N} - 1\right)^n \sim \mathrm{i}^n y^n, \tag{5.23}$$

and so

$$x_1 \sim \sum_{n=0}^{\infty} \frac{(2k)!}{(2k+n)!} \binom{k+n-1}{n} \mathrm{i}^n y^n \tag{5.24}$$

$$= \frac{(2k)!}{(k-1)!} f(\mathrm{i}y), \tag{5.25}$$

and similarly,

$$N x_2 \sim \sum_{n=1}^{\infty} \frac{(2k)!}{(2k+n)!} \binom{k+n-1}{n} n \mathrm{i}^n y^n \tag{5.26}$$

$$= \frac{(2k)!}{(k-1)!} \mathrm{i}y f'(\mathrm{i}y), \tag{5.27}$$

$$\frac{1}{N} y_1 \sim \sum_{n=0}^{\infty} \frac{(2k+1)!}{(2k+1+n)!} \binom{k+n}{n} \mathrm{i}^n y^n \tag{5.28}$$

$$= \frac{(2k+1)!}{k!} f'(\mathrm{i}y), \tag{5.29}$$

and

$$y_2 \sim \sum_{n=1}^{\infty} \frac{(2k+1)!}{(2k+1+n)!} \binom{k+n}{n} (n+1) \mathrm{i}^n y^n \tag{5.30}$$

$$= \frac{(2k+1)!}{k!} (\mathrm{i}y f''(\mathrm{i}y) + f'(\mathrm{i}y)). \tag{5.31}$$

Thus, defining

$$F_k(y) = k e^{-\mathrm{i}y} \left\{ \mathrm{i}y^{2k+1} f(\mathrm{i}y) f''(\mathrm{i}y) + y^{2k} f(\mathrm{i}y) f'(\mathrm{i}y) - \mathrm{i}y^{2k+1} f'(\mathrm{i}y) f'(\mathrm{i}y) \right\} \tag{5.32}$$

we have

$$e^{-\mathrm{i}y} (x_1 y_2 - x_2 y_1) \sim \frac{1}{y^{2k}} \frac{(2k)!(2k+1)!}{k!k!} F_k(y). \tag{5.33}$$

(The reason $F_k(y)$ is defined with the (apparently unnecessary) factors of y^{2k} will be obvious later).

Also, from the asymptotics for the G–function, one can show that

$$\frac{G(N+2k+2)G(N)}{G^2(N+k+1)} \sim N^{(k+1)^2} \tag{5.34}$$

so, on combining (5.4) with theorem 5.1, and inserting the above asymptotics, we obtain

$$\mathbb{E}_N \left\{ \left| Z_U \left(\theta_1 + \frac{y}{N} \right) \right|^{2k} \right\} \sim y^{2k} N^{-2k-1} \, \mathbb{E}_{N-1} \left\{ \left| Z_{\widetilde{U}}(0) \right|^{2k} \left| Z_{\widetilde{U}} \left(\frac{y}{N} \right) \right|^2 \right\} \tag{5.35}$$

$$\sim N^{k^2} \frac{G^2(k+2)}{G(2k+3)} y^{2k} e^{-\mathrm{i}y} (x_1 y_2 - x_2 y_1) \tag{5.36}$$

$$\sim \frac{G^2(k+1)}{G(2k+1)} F_k(y) N^{k^2}. \tag{5.37}$$

This is valid as $N \to \infty$, uniformly in y for $|y| \ll N$.

Since the LHS of this equation is real, then $F_k(y)$ must be real for real y (we will use this fact to save us some work later on).

So, to prove theorem 5.2, all that remains is to show that

$$F_k(y) = \frac{A(y)\cos(y) + B(y)\sin(y) + C(y)}{y^{2k}}. \tag{5.38}$$

To calculate $F_k(y)$ observe that from (5.22),

$$\begin{aligned}
f(y) &= \sum_{n=0}^{\infty} \frac{(k-1)!}{(2k+n)!} \binom{k+n-1}{n} y^n && (5.39)\\
&= \sum_{n=0}^{\infty} \frac{1}{(2k+n)!} (k+n-1)\cdots(n+1) y^n && (5.40)\\
&= \frac{\mathrm{d}^{k-1}}{\mathrm{d}y^{k-1}} \left(\sum_{n=0}^{\infty} \frac{y^{k+n-1}}{(2k+n)!} \right) && (5.41)\\
&= \frac{\mathrm{d}^{k-1}}{\mathrm{d}y^{k-1}} \left(\sum_{n=0}^{\infty} \frac{y^{n}}{(k+1+n)!} \right) && (5.42)
\end{aligned}$$

which is true since the $y^0, \ldots, y^{k-2}$ terms vanish on differentiation. The series can be evaluated exactly as

$$\sum_{n=0}^{\infty} \frac{y^n}{(k+1+n)!} = \frac{1}{y^{k+1}} \left[e^y - \sum_{n=0}^{k} \frac{y^n}{n!} \right] \tag{5.43}$$

and so

$$\begin{aligned}
f(y) &= \frac{\mathrm{d}^{k-1}}{\mathrm{d}y^{k-1}} \left(\frac{1}{y^{k+1}} \left[e^y - \sum_{n=0}^{k} \frac{y^n}{n!} \right] \right) && (5.44)\\
&= \sum_{m=0}^{k-1} \binom{k-1}{m} (-1)^m \frac{(k+m)!}{k!} e^y y^{-k-m-1} + (-1)^k \sum_{n=0}^{k} \frac{(2k-n-1)!}{(k-n)!} \frac{y^{n-2k}}{n!}. && (5.45)
\end{aligned}$$

Inserting this form of $f(\mathrm{i}y)$ into (5.32), $y^{2k}F_k(y)$ equals

$$
\begin{aligned}
e^{\mathrm{i}y}\Bigg\{&\sum_{m=0}^{k-1}\sum_{n=0}^{k+1}(-1)^{m+n}\binom{k-1}{m}\binom{k+1}{n}\frac{(k+m)!}{(k-1)!}\frac{(k+n)!}{k!}(\mathrm{i}y)^{2k-m-n-1}+\\
&+\sum_{m=0}^{k-1}\sum_{n=0}^{k}(-1)^{m+n}\binom{k-1}{m}\binom{k}{n}\frac{(k+m)!}{(k-1)!}\frac{(k+n)!}{k!}(\mathrm{i}y)^{2k-m-n-2}+\\
&-\sum_{m=0}^{k}\sum_{n=0}^{k}(-1)^{m+n}\binom{k}{m}\binom{k}{n}\frac{(k+m)!}{(k-1)!}\frac{(k+n)!}{k!}(\mathrm{i}y)^{2k-m-n-1}\Bigg\}\\
+e^{-\mathrm{i}y}\Bigg\{&k\sum_{m=0}^{k}\sum_{n=0}^{k}\Big[\frac{(2k-m-1)!}{(k-m)!}\frac{(2k-n+1)!}{(k-n)!}\frac{(\mathrm{i}y)^{m+n-1}}{m!n!}+\\
&-\frac{(2k-m-1)!}{(k-m)!}\frac{(2k-n)!}{(k-n)!}\frac{(\mathrm{i}y)^{m+n-1}}{m!n!}+\\
&-\frac{(2k-m)!}{(k-m)!}\frac{(2k-n)!}{(k-n)!}\frac{(\mathrm{i}y)^{m+n-1}}{m!n!}\Big]\Bigg\}\\
+\Bigg\{&\sum_{m=0}^{k-1}\sum_{n=0}^{k}(-1)^{k+m}\binom{k-1}{m}\frac{(k+m)!}{(k-1)!}\frac{(2k-n+1)!}{(k-n)!}\frac{(\mathrm{i}y)^{k-m+n-2}}{n!}+\\
&+\sum_{n=0}^{k}\sum_{m=0}^{k+1}(-1)^{k+m}\frac{(2k-n-1)!}{(k-n)!}\binom{k+1}{m}\frac{(k+m)!}{(k-1)!}\frac{(\mathrm{i}y)^{k-m+n}}{n!}\\
&+\sum_{m=0}^{k-1}\sum_{n=0}^{k}(-1)^{k+m+1}\binom{k-1}{m}\frac{(k+m)!}{(k-1)!}\frac{(2k-n)!}{(k-n)!}\frac{(\mathrm{i}y)^{k-m+n-2}}{n!}+\\
&+\sum_{n=0}^{k}\sum_{m=0}^{k}(-1)^{k+m}\frac{(2k-n-1)!}{(k-n)!}\binom{k}{m}\frac{(k+m)!}{(k-1)!}\frac{(\mathrm{i}y)^{k-m+n-1}}{n!}\\
&+2\sum_{m=0}^{k}\sum_{n=0}^{k}(-1)^{k+m}\binom{k}{m}\frac{(k+m)!}{(k-1)!}\frac{(2k-n)!}{(k-n)!}\frac{(\mathrm{i}y)^{k-m+n-1}}{n!}\Bigg\}
\end{aligned}
\tag{5.46}
$$

For notational (and space!) convenience, we write this as

$$
y^{2k}F_k(y)=e^{\mathrm{i}y}\sum_{r=-1}^{2k-1}a_r(\mathrm{i}y)^r+e^{-\mathrm{i}y}\sum_{r=-1}^{2k-1}b_r(\mathrm{i}y)^r+\sum_{r=-1}^{2k}c_r(\mathrm{i}y)^r \tag{5.47}
$$

The requirement that $F_k(y)$ is real for real y allows us to deduce many things about the coefficients a_r, b_r and c_r. Namely:

$$
c_{2n-1}=0 \text{ for } n=0,1,\ldots,k \tag{5.48}
$$

$$
a_{2n}=b_{2n} \text{ for } n=0,1,\ldots,k-1 \tag{5.49}
$$

$$
a_{2n-1}=-b_{2n-1} \text{ for } n=0,1,\ldots,k \tag{5.50}
$$

(later on we will also show that $b_{-1} = b_{2k-1} = 0$). Thus we have

$$y^{2k}F_k(y) = 2\cos(y)\sum_{n=0}^{k-1}(-1)^n b_{2n}y^{2n} + 2\sin(y)\sum_{n=1}^{k-1}(-1)^{n-1}b_{2n-1}y^{2n-1} + \\ + \sum_{n=0}^{k}(-1)^n c_{2n}y^{2n}. \quad (5.51)$$

From (5.33) and (5.6–5.9) we know that

$$F_k(y) = \frac{k!k!}{(2k)!(2k+1)!}y^{2k} + \mathcal{O}\left(y^{2k+1}\right) \qquad \text{as } y \to 0 \quad (5.52)$$

(that is, the first $4k$ terms of the Taylor expansion of (5.51) vanish) so we can find c_r in terms of b_r by equating coefficients as follows:

$$\sum_{r=0}^{k}(-1)^r c_{2r}y^{2r} = \left(\sum_{m=0}^{\infty}\frac{(-1)^m y^{2m}}{(2m)!}\right)\left(\sum_{n=0}^{k-1}2(-1)^{n-1}b_{2n}y^{2n}\right) + \\ + \left(\sum_{m=0}^{\infty}\frac{(-1)^m y^{2m+1}}{(2m+1)!}\right)\left(\sum_{n=1}^{k-1}2(-1)^{n}b_{2n-1}y^{2n-1}\right) + \mathcal{O}\left(y^{4k}\right) \quad (5.53)$$

and so, for $r = 0, \ldots, k-1$,

$$c_{2r} = \sum_{n=1}^{r}\frac{2b_{2n-1}}{(2r-2n+1)!} - \sum_{n=0}^{r}\frac{2b_{2n}}{(2r-2n)!} \quad (5.54)$$

and

$$c_{2k} = \sum_{n=1}^{k-1}\frac{2b_{2n-1}}{(2r-2n+1)!} - \sum_{n=0}^{k-1}\frac{2b_{2n}}{(2r-2n)!}. \quad (5.55)$$

Thus the problem has been reduced to calculating b_r. Note that from (5.46),

$$\sum_{r=-1}^{2k-1} b_r(\mathrm{i}y)^r = k\sum_{m=0}^{k}\sum_{n=0}^{k}\frac{(2k-m-1)!}{(k-m)!}\frac{(2k-n)!}{(k-n)!}\frac{(\mathrm{i}y)^{m+n-1}}{m!n!}(m-n) \quad (5.56)$$

so

$$b_r = \begin{cases} k\sum_{a=0}^{r+1}\frac{(2k-a-1)!}{(k-a)!}\frac{(2k+a-r-1)!}{(k+a-r-1)!}\frac{(2a-r-1)}{a!(r+1-a)!} & \text{for } -1 \le r \le k-1 \\ k\sum_{a=r+1-k}^{k}\frac{(2k-a-1)!}{(k-a)!}\frac{(2k+a-r-1)!}{(k+a-r-1)!}\frac{(2a-r-1)}{a!(r+1-a)!} & \text{for } k \le r \le 2k-1 \end{cases} \quad (5.57)$$

which can be combined as follows

$$b_r = k\sum_{a=0}^{k}\frac{(2k-a-1)!}{(k-a)!}\frac{(2k+a-r-1)!}{(k+a-r-1)!}\frac{(2a-r-1)}{a!(r+1-a)!} \quad (5.58)$$

$$= \sum_{a=0}^{k}(2a-r-1)\binom{2k-a-1}{k-1}\binom{2k+a-r-1}{k}\frac{k!k!}{a!(r+1-a)!} \quad (5.59)$$

$$= \sum_{a=0}^{k}\frac{(2k-a-1)!}{k!}\frac{(2k+a-r-1)!}{(k-1)!}\binom{k}{a}\binom{k}{r+1-a}(2a-r-1) \quad (5.60)$$

for $-1 \leq r \leq 2k-1$ (this is okay because $\binom{k}{r+1-a}$ kills off the terms we don't want). By direct calculation, we see that $b_{-1} = b_{2k-1} = 0$.

Observing that the b_r are integers completes the proof of theorem 5.2. □

Remark. The above large-N asymptotics are uniform in y for $|y| \ll N$. This can be extended to $|y| \leq AN$ for $A < 2\pi$ an arbitrary constant as follows: Let β be a fixed constant subject to $0 < \beta < 2\pi$. By (5.4) and the results of Basor [6] (which has already been quoted in the proof of lemma 2.1),

$$\mathbb{E}_N \left\{ |Z_U(\theta_1 + \beta)|^{2k} \right\} = \frac{1}{N} \left| 2 \sin\left(\tfrac{1}{2}\beta\right) \right|^{2k} \mathbb{E}_{N-1} \left\{ \left| Z_{\widetilde{U}}(0) \right|^{2k} \left| Z_{\widetilde{U}}(\beta) \right|^2 \right\} \tag{5.61}$$

$$\sim \frac{G^2(k+1)}{G(2k+1)} N^{k^2}. \tag{5.62}$$

If one lets $y = N\beta$ then it is easily checked that

$$F_k(N\beta) = 1 + \mathcal{O}\left(\frac{1}{N}\right). \tag{5.63}$$

So theorem 5.2 gives the correct first order term, for $|y| \leq AN$ for $A < 2\pi$ a constant.

Remark. Note that b_r is a strictly positive integer for $r = 0, \ldots, 2k-2$, and that the c_{2r} are integers also.

Remark An alternative expression is

$$F_k(2x) = \frac{X(x)\sin^2(x) + Y(x)\sin(2x) + Z(x)}{x^{2k}} \tag{5.64}$$

with

$$X(x) = \frac{-A(2x)}{2^{2k-1}} \tag{5.65}$$

$$Y(x) = \frac{B(2x)}{2^{2k}} \tag{5.66}$$

$$Z(x) = \frac{A(2x) + C(2x)}{2^{2k}} \tag{5.67}$$

all polynomials with coefficients which are smaller than the coefficients of A, B, C. This means

$$\mathbb{E}_N \left\{ \left| Z_U\left(\theta_1 + \frac{2\pi\alpha}{N}\right) \right|^{2k} \right\} \sim \frac{G^2(k+1)}{G(2k+1)} \frac{X(\pi\alpha)\sin^2(\pi\alpha) + Y(\pi\alpha)\sin(2\pi\alpha) + Z(\pi\alpha)}{(\pi\alpha)^{2k}} N^{k^2} \tag{5.68}$$

and this is appears to be the most useful and compact way of expressing this result.

k	$X(x)$	$Y(x)$	$Z(x)$
1	-1	0	x^2
2	$2x^2-3$	$3x$	x^4-3x^2
3	$-3x^4+72x^2-45$	$-12x^3+45x$	$x^6-3x^4-45x^2$

Table 5.1: X, Y, Z for $k = 1, 2, 3$.

5.2 Conjecture about the zeta function

Conjecture 5.3. *For k an integer,*

$$\frac{1}{N(T)}\sum_{0<\gamma_n\leq T}\left|\zeta\left(\tfrac{1}{2}+\mathrm{i}\left(\gamma_n+\alpha/L\right)\right)\right|^{2k}\sim\frac{G^2(k+1)}{G(2k+1)}a(k)F_k(2\pi\alpha)\left(\log\frac{T}{2\pi}\right)^{k^2} \tag{5.69}$$

as $T\to\infty$, uniformly in α for $|\alpha|\leq AL$, where $A<1$ is a constant, $L=\frac{1}{2\pi}\log\frac{T}{2\pi}$ is the density of zeros of height T, and $a(k)$ is given by (1.85).

If this conjecture is true, then we are able to prove conjecture 4.3 from chapter 4 (though only for integer k) and a variant of the Keating-Snaith conjecture (conjecture 1.4), again only for integer k.

Corollary 5.3.1. *If RH and conjecture 5.3 are true, then*

$$\frac{1}{N(T)}\sum_{0<\gamma_n\leq T}\left|\zeta'\left(\tfrac{1}{2}+\mathrm{i}\gamma_n\right)\right|^{2k}\sim\frac{G^2(k+2)}{G(2k+3)}a(k)\left(\log\frac{T}{2\pi}\right)^{k(k+2)} \tag{5.70}$$

Proof. By the definition of differentiation,

$$\left|\zeta'\left(\tfrac{1}{2}+\mathrm{i}\gamma_n\right)\right|^{2k}=L^{2k}\lim_{a\to 0}\frac{\left|\zeta\left(\frac{1}{2}+\mathrm{i}\left(\gamma_n+\frac{\alpha}{L}\right)\right)\right|^{2k}}{\alpha^{2k}} \tag{5.71}$$

(we assume RH so that $\zeta(1/2+\mathrm{i}\gamma_n)=0$). From (5.52) we have

$$\lim_{\alpha\to 0}\frac{F(2\pi\alpha)}{\alpha^{2k}}=(2\pi)^{2k}\frac{k!k!}{(2k)!(2k+1)!} \tag{5.72}$$

so applying conjecture 5.3 and using uniformity to swap the $\alpha \to 0$ and $N \to \infty$ limits, we have

$$\frac{1}{N(T)} \sum_{0<\gamma_n \leq T} \left|\zeta'\left(\tfrac{1}{2} + \mathrm{i}\gamma_n\right)\right|^{2k} \sim \frac{G^2(k+1)}{G(2k+1)} a(k) L^{2k} (2\pi)^{2k} \frac{k!k!}{(2k)!(2k+1)!} \left(\log \frac{T}{2\pi}\right)^{k^2} \tag{5.73}$$

$$= \frac{G^2(k+2)}{G(2k+3)} a(k) \left(\log \frac{T}{2\pi}\right)^{k(k+2)} \tag{5.74}$$

as required. □

Corollary 5.3.2. *From conjecture 5.3 it follows that for $\beta > 0$ fixed,*

$$\frac{1}{N(T)} \sum_{0<\gamma_n \leq T} \left|\zeta\left(\tfrac{1}{2} + \mathrm{i}(\gamma_n + \beta)\right)\right|^{2k} \sim \frac{G^2(k+1)}{G(2k+1)} a(k) \left(\log \frac{T}{2\pi}\right)^{k^2} \tag{5.75}$$

Proof. Trivial, as

$$\lim_{\alpha \to \infty} F_k(2\pi\alpha) = 1 \tag{5.76}$$

and put $\alpha = L\beta$ in conjecture 5.3.

Remark. Note that the left-hand side of (5.75) behaves like

$$\frac{1}{T} \int_0^T \left|\zeta\left(\tfrac{1}{2} + \mathrm{i}t\right)\right|^{2k} \, \mathrm{d}t \tag{5.77}$$

so this can be thought of as a variant of the Keating-Snaith conjecture.

5.2.1 Comparison with the zeta function

Gonek [45] showed, amongst other things, that

$$\frac{1}{N(T)} \sum_{0<\gamma_n \leq T} \left|\zeta\left(\tfrac{1}{2} + \mathrm{i}\left(\gamma_n + \alpha/L\right)\right)\right|^2 \sim \left(1 - \left(\frac{\sin(\pi\alpha)}{\pi\alpha}\right)^2\right) \log \frac{T}{2\pi} \tag{5.78}$$

uniformly in α for $|\alpha| \leq L/2$, which is in perfect agreement with conjecture 5.3 when $k = 1$.

There is no proof of the conjecture for $k = 2$ (unlike conjecture 1.4 which is proven for $k = 1$ and 2). But there are theorems along the lines of conjecture 5.3 for $k = 2$:

Theorem 5.4. (Conrey, Ghosh and Gonek [29].) *Assume GRH and let* $A(s) = \sum_{n\le x} n^{-s}$ *where* $x = \left(\frac{T}{2\pi}\right)^{\eta}$ *for some* $\eta \in (0, \frac{1}{2})$. *Then,*

$$\frac{1}{N(T)} \sum_{0<\gamma_n\le T} \left|\zeta A\left(\tfrac{1}{2} + \mathrm{i}\left(\gamma_n + \alpha/L\right)\right)\right|^2 \sim \frac{6}{\pi^2} \sum_{j=0}^{\infty} \frac{(-1)^{j+1}(2\pi\alpha)^{2j+2}}{(2j+5)!} \times$$
$$\left(-\eta^2 + \tfrac{1}{3}(2j+5)\eta^3 - \frac{2j+5}{j+3}\eta^{2j+6} + \eta^{2j+7} + \eta^2(1-\eta)^{2j+5}\right)\left(\log\frac{T}{2\pi}\right)^4 \quad (5.79)$$

uniformly for bounded α.

(We have slightly changed notation from [29], to be consistent with our definition of $L = \frac{1}{2\pi}\log\frac{T}{2\pi}$).

Putting $\eta = 1$ in the above (which, as it stands, is not allowed under the conditions of the theorem) then $A(\frac{1}{2} + \mathrm{i}t) = \zeta(\frac{1}{2} + \mathrm{i}t) + \mathcal{O}(t^{-1/2})$, and we have

$$\frac{1}{N(T)} \sum_{0<\gamma_n\le T} \left|\zeta^2\left(\tfrac{1}{2} + \mathrm{i}\left(\gamma_n + \alpha/L\right)\right)\right|^2$$
$$\sim \frac{4}{\pi^2} \sum_{j=1}^{\infty} \frac{(-1)^{j+1}(2\pi\alpha)^{2j+2}}{(2j+6)!}(2j^2+5j)\left(\log\frac{T}{2\pi}\right)^4 \quad (5.80)$$

Note that

$$\frac{4}{\pi^2} \sum_{j=1}^{\infty} \frac{(-1)^{j+1}(2\pi\alpha)^{2j+2}}{(2j+6)!}(2j^2+5j) =$$
$$\frac{1}{12}a(2)\frac{(2\pi^2\alpha^2-3)\sin^2(\pi\alpha) + 3\pi\alpha\sin(2\pi\alpha) + (\pi\alpha)^4 - 3(\pi\alpha)^2}{(\pi\alpha)^4} \quad (5.81)$$

which is what is predicted in conjecture 5.3.

That is, from a purely number theoretical calculation involving no random matrix theory, we have

Conjecture 5.5. *Assuming that* $\eta = 1$ *is permissible in theorem 5.4 then*

$$\frac{1}{N(T)} \sum_{0<\gamma_n\le T} \left|\zeta\left(\tfrac{1}{2} + \mathrm{i}\left(\gamma_n + \alpha/L\right)\right)\right|^4 \sim \frac{1}{2\pi^2}F_2(2\pi\alpha)\left(\log\frac{T}{2\pi}\right)^4 \quad (5.82)$$

where

$$F_2(2x) = \frac{(2x^2-3)\sin^2(x) + 3x\sin(2x) + x^4 - 3x^2}{x^4} \quad (5.83)$$

So, following the proof of corollary 5.3.1, we may deduce

Corollary 5.5.1. *If RH and conjecture 5.5 are both true then*

$$\frac{1}{N(T)} \sum_{0<\gamma_n\le T} \left|\zeta'\left(\tfrac{1}{2} + \mathrm{i}\gamma_n\right)\right|^4 \sim \frac{1}{1440\pi^2}\left(\log\frac{T}{2\pi}\right)^8. \quad (5.84)$$

5.3 Large gaps between zeros of the zeta function

Using conjecture 5.3, large gaps between the non-trivial zeros can be studied.

Corollary 5.5.2. *Assuming the Riemann hypothesis, if conjecture 5.5 is true, then for*

$$\lambda := \limsup_{n\to\infty}(\gamma_{n+1}-\gamma_n)\frac{1}{2\pi}\log\frac{\gamma_n}{2\pi} \tag{5.85}$$

we have $\lambda > 2.7$.

Remark. Random matrix theory predicts that $\lambda = \infty$ since the spacing distribution, (1.59), does not have compact support.

Our proof will follow Mueller's [75, 45] proof for the case $k = 1$ (Gonek's result, (5.78)), where she showed $\lambda > 1.9$ under the assumption of the Riemann hypothesis. One of the main reasons Conrey, Ghosh and Gonek proved theorem 5.4 was to prove that $\lambda > 2.68$, this time under the assumption of GRH. The best unconditional result is $\lambda > \sqrt{11/2} = 2.345\ldots$, which is due to Hall [50].

Proof. Note that if $\beta > \lambda$ then

$$\sum_{0<\gamma_n\le T}\int_{\gamma_n-\beta/2L}^{\gamma_n+\beta/2L}|\zeta(\tfrac{1}{2}+\mathrm{i}t)|^4\,\mathrm{d}t > \int_0^T|\zeta(\tfrac{1}{2}+\mathrm{i}t)|^4\,\mathrm{d}t \tag{5.86}$$

Ingham's result [61] deals with the right-hand side, saying

$$\int_0^T|\zeta(\tfrac{1}{2}+\mathrm{i}t)|^4\,\mathrm{d}t \sim \frac{1}{2\pi^2}T\left(\log\frac{T}{2\pi}\right)^4. \tag{5.87}$$

Changing variables in the left-hand integral to $t = \gamma_n + \alpha/L$ gives

$$\sum_{0<\gamma_n\le T}\int_{\gamma_n-\beta/2L}^{\gamma_n+\beta/2L}|\zeta(\tfrac{1}{2}+\mathrm{i}t)|^4\,\mathrm{d}t = \int_{-\beta/2}^{\beta/2}\sum_{0<\gamma_n\le T}\left|\zeta(\tfrac{1}{2}+\mathrm{i}(\gamma_n+\alpha/L))\right|^4\mathrm{d}\alpha\frac{2\pi}{\log(T/2\pi)}. \tag{5.88}$$

Using conjecture 5.5, this is asymptotic to

$$\frac{1}{2\pi^2}\left(\log\frac{T}{2\pi}\right)^4\int_{-\beta/2}^{\beta/2}F_2(2\pi\alpha)\,\mathrm{d}\alpha\frac{2\pi N(T)}{\log(T/2\pi)}. \tag{5.89}$$

Combining these, we see that if $\beta > \lambda$ then

$$\int_{-\beta/2}^{\beta/2}F_2(2\pi\alpha)\,\mathrm{d}\alpha > 1. \tag{5.90}$$

However, using Maple it is possible to show that

$$\int_{-2.7/2}^{2.7/2}F_2(2\pi\alpha)\,\mathrm{d}\alpha < 0.994 \tag{5.91}$$

and so $\lambda > 2.7$. □

Chapter 6

Joint moments of the Riemann zeta function and its logarithmic derivative

6.1 Introduction

Recall the definition of Hardy's function[1] from (1.8):

$$\mathcal{Z}(t) = \sqrt{\chi(\tfrac{1}{2} - \mathrm{i}t)}\zeta(\tfrac{1}{2} + \mathrm{i}t). \tag{6.1}$$

$\mathcal{Z}(t)$ has the useful property that it is real for real t, and that $|\zeta(1/2 + \mathrm{i}t)| = |\mathcal{Z}(t)|$. However, $|\zeta'(1/2 + \mathrm{i}t)| \neq |\mathcal{Z}'(t)|$ unless $\zeta(1/2 + \mathrm{i}t) = 0$ or $\vartheta'(t) = 0$ (which, for real t, only occurs when $t \approx \pm 6.29$).

Observe that

$$Z_U^*(\theta) = \prod_{n=1}^{N} \left(1 - e^{-\mathrm{i}(\theta_n - \theta)}\right) \tag{6.2}$$

$$= \prod_{n=1}^{N} \left(-e^{\mathrm{i}\theta} e^{-\mathrm{i}\theta_n}\right) \left(1 - e^{\mathrm{i}(\theta_n - \theta)}\right) \tag{6.3}$$

$$= (-1)^N e^{\mathrm{i}N\theta} e^{-\mathrm{i}\sum_{n=1}^{N}\theta_n} Z_U(\theta), \tag{6.4}$$

[1] Hardy's function is often denoted $Z(t)$, but since we use $Z(\theta)$ for the characteristic polynomial, we have introduced the notation $\mathcal{Z}(t)$ to avoid confusion. Note that in [56], Hejhal uses $V(s) \equiv \mathcal{Z}(\frac{1}{2}\mathrm{i} - \mathrm{i}s)$, which is Hardy's function for $s = \frac{1}{2} + \mathrm{i}t$.

which implies that for real θ,

$$V_U(\theta) := e^{\mathrm{i}N(\theta+\pi)/2} e^{-\mathrm{i}\sum_{n=1}^{N}\theta_n/2} Z_U(\theta) \tag{6.5}$$

is a real function.

In this chapter we will use random matrix theory to estimate

$$\frac{1}{T}\int_0^T \left|\zeta(\tfrac{1}{2}+\mathrm{i}t)\right|^{2k} \left|\frac{\zeta'(\frac{1}{2}+\mathrm{i}t)}{\zeta(\frac{1}{2}+\mathrm{i}t)}\right|^{2h} \mathrm{d}t \tag{6.6}$$

and

$$\frac{1}{T}\int_0^T |\mathcal{Z}(t)|^{2k} \left|\frac{\mathcal{Z}'(t)}{\mathcal{Z}(t)}\right|^{2h} \mathrm{d}t, \tag{6.7}$$

by studying

$$\mathbb{E}\left\{|Z_U(0)|^{2k-2h}|Z_U'(0)|^{2h}\right\} \tag{6.8}$$

and

$$\mathbb{E}\left\{|V_U(0)|^{2k-2h}|V_U'(0)|^{2h}\right\}. \tag{6.9}$$

As has already been seen, $Z_U(\theta)$ is a good model for $\zeta(\frac{1}{2}+\mathrm{i}t)$. We shall argue that $Z_U'(\theta)$ is also a good model for $\frac{\mathrm{d}}{\mathrm{d}t}\zeta(\frac{1}{2}+\mathrm{i}t) = \mathrm{i}\zeta'(1/2+\mathrm{i}t)$ (since we consider $|\zeta'(1/2+\mathrm{i}t)|$, the factor of i is irrelevant[2]), but one needs $V_U'(\theta)$ in order to correctly model $\mathcal{Z}'(t)$.

6.1.1 Summary of relevant results on the zeta function

Hall [50] has conjectured that

$$\frac{1}{T}\int_0^T \left|\zeta(\tfrac{1}{2}+\mathrm{i}t)\right|^{2k-2h} \left|\zeta'(\tfrac{1}{2}+\mathrm{i}t)\right|^{2h} \mathrm{d}t \sim C(h,k)\left(\log\frac{T}{2\pi}\right)^{k^2+2h}. \tag{6.10}$$

For integer h and k this conjecture can be shown (see §6.3) to be equivalent to

$$\frac{1}{T}\int_0^T |\mathcal{Z}(t)|^{2k-2h} \left|\mathcal{Z}'(t)\right|^{2h} \mathrm{d}t \sim \widetilde{C}(h,k)\left(\log\frac{T}{2\pi}\right)^{k^2+2h}, \tag{6.11}$$

with the $\widetilde{C}(h,k)$ given in terms of $C(h,k)$ for $h = 0,\ldots,k$, and vice versa.

For $k = 1, 2$, the problem has been completely solved to leading order in T (which is what we want):

- For $k = 1$, Hardy and Littlewood [52] proved that $C(0,1) = 1$ and in [61] Ingham showed that $C(1,1) = \frac{1}{3}$.

[2] but see §6.3

- For $k = 2$, we have $C(0,2) = \frac{1}{2\pi^2}$ due to Ingham [61], and $C(1,2) = \frac{2}{15\pi^2}$ and $C(2,2) = \frac{61}{1680\pi^2}$ both due to Conrey [22].

The analogous results for $\mathcal{Z}(t)$ are easy to deduce from these, and for completeness we record them here as $\widetilde{C}(1,1) = \frac{1}{12}$, $\widetilde{C}(1,2) = \frac{1}{120\pi^2}$ and $\widetilde{C}(2,2) = \frac{1}{1120\pi^2}$.

Conrey and Ghosh [27] have proved $\widetilde{C}(1/2,1) = \frac{e^2-5}{4\pi}$, but $C(1/2,1)$ is currently unknown.

When $h = 0$ the Keating-Snaith conjecture (conjecture 1.4) explicitly calculates the constants as

$$C(0,k) = \widetilde{C}(0,k) = \frac{G^2(1+k)}{G(1+2k)} a(k). \tag{6.12}$$

Finally, when $k = h$, conjecture (6.10) is consistent with the work of Hejhal, [56], which suggests that

$$\frac{1}{T}\int_0^T \left|\zeta'(\tfrac{1}{2}+\mathrm{i}t)\right|^{2h} \mathrm{d}t \asymp \left(\log\frac{T}{2\pi}\right)^{h^2+2h}. \tag{6.13}$$

6.2 Random matrix results

Table 6.1 gives the results for $k = 1, 2$, calculated using the method described below.

k	h	$\mathbb{E}\left\{\lvert Z_U(0)\rvert^{2k-2h}\lvert Z_U'(0)\rvert^{2h}\right\}$	$\mathbb{E}\left\{\lvert V_U(0)\rvert^{2k-2h}\lvert V_U'(0)\rvert^{2h}\right\}$
1	0	N	N
1	1	$\frac{1}{3}N^3$	$\frac{1}{12}N^3$
2	0	$\frac{1}{12}N^4$	$\frac{1}{12}N^4$
2	1	$\frac{1}{45}N^6$	$\frac{1}{720}N^6$
2	2	$\frac{61}{10080}N^8$	$\frac{1}{6720}N^8$

Table 6.1: Table of results for $k = 1$ and 2.

The main result of this section is theorem 6.4, which calculates

$$F(h,k) := \lim_{N\to\infty} \frac{1}{N^{k^2+2h}} \mathbb{E}\left\{|Z_U(0)|^{2k-2h}|Z_U'(0)|^{2h}\right\} \tag{6.14}$$

and

$$\widetilde{F}(h,k) := \lim_{N\to\infty} \frac{1}{N^{k^2+2h}} \mathbb{E}\left\{|V_U(0)|^{2k-2h}|V_U'(0)|^{2h}\right\}. \tag{6.15}$$

From this we will conjecture that $C(h,k) = F(h,k)a(k)$ for the Riemann zeta function, and for Hardy's function $\widetilde{C}(h,k) = \widetilde{F}(h,k)a(k)$. That is,

Conjecture 6.1.

$$\frac{1}{T}\int_0^T \left|\zeta(\tfrac{1}{2}+\mathrm{i}t)\right|^{2k-2h}\left|\zeta'(\tfrac{1}{2}+\mathrm{i}t)\right|^{2h} \sim F(h,k)a(k)\left(\log\frac{T}{2\pi}\right)^{k^2+2h} \tag{6.16}$$

and

$$\frac{1}{T}\int_0^T |\mathcal{Z}(t)|^{2k-2h}\left|\mathcal{Z}'(t)\right|^{2h} \sim \widetilde{F}(h,k)a(k)\left(\log\frac{T}{2\pi}\right)^{k^2+2h} \tag{6.17}$$

where $a(k)$ is given by (1.85).

Note that $a(1) = 1$ and $a(2) = \frac{6}{\pi^2}$, so, using the results in table 6.1 for $k = 1$ and $k = 2$, we have perfect agreement with all the results in section 6.1.1, with the exception of Conrey and Ghosh's [27] result that

$$\frac{1}{T}\int_1^T |\mathcal{Z}(t)\mathcal{Z}'(t)|\,\mathrm{d}t \sim \frac{e^2-5}{4\pi}\log^2 T \tag{6.18}$$

which we are (currently) unable to reproduce in random matrix theory, since all our results rely on h being an integer. If it could be shown that $\widetilde{F}(1/2,1) = \frac{e^2-5}{4\pi}$, then this would be evidence that conjecture 6.1 holds for all h,k such that $h > -1/2$ and $k > h - 1/2$.

6.2.1 Various lemmas

In this subsection we will evaluate

$$T_{k,n} := \mathbb{E}_N\left\{Z_U(0)^k Z_U^*(0)^{k-n} Z_U^*(\beta)^n\right\} \tag{6.19}$$

for $n = 0,\ldots,k$. Lemma 6.2 calculates $T_{k,n}$ as a $k\times k$ determinant, essentially using Heine's identity and a very useful trick due to Basor and Forrester [7]. Lemma 6.3 is a continuation of lemma 6.2, evaluating the determinant by row and column

manipulation to show that $T_{k,n} = M_N(2k)$ times a certain $n \times n$ determinant, which is itself just a large polynomial in $e^{\mathrm{i}\beta}$. Finally, in corollary 6.3.1, this polynomial is written as an infinite series in β, and the coefficients are evaluated for large N.

Lemma 6.2. *For $n = 0, \ldots, k$,*

$$T_{k,n} = \det \begin{bmatrix} \binom{N+k}{k} & \cdots & \binom{N+2k-1}{k} \\ \vdots & & \vdots \\ \binom{N+k}{2k-n-1} & \cdots & \binom{N+2k-1}{2k-n-1} \\ z_{k-n+1,1} & \cdots & z_{k-n+1,k} \\ \vdots & & \vdots \\ z_{k,1} & \cdots & z_{k,k} \end{bmatrix} \tag{6.20}$$

with, for $i = k-n+1, \ldots, k$ and $j = 1, \ldots, k$,

$$z_{i,j} = \sum_{m=0}^{N+j-i} \binom{N+k+j-1}{k+i+m-1} \binom{m+i-k+n-1}{i-k+n-1} (e^{\mathrm{i}\beta} - 1)^m \tag{6.21}$$

Proof. The proof essentially follows Basor and Forrester's method, [7].

First of all, note that Heine's identity (lemma 1.10) gives that

$$\mathbb{E}_N \left\{ Z_U(0)^k Z_U^*(0)^{k-n} Z_U^*(\beta)^n \right\} = D_N[g] \tag{6.22}$$

where $D_N[g]$ is the Toeplitz determinant with symbol

$$g(\theta) = \left(1 - e^{\mathrm{i}\theta}\right)^k \left(1 - e^{-\mathrm{i}\theta}\right)^{k-n} \left(1 - e^{-\mathrm{i}\theta} e^{\mathrm{i}\beta}\right)^n \tag{6.23}$$

$$= (-1)^k e^{-\mathrm{i}k\theta} \left(e^{\mathrm{i}\theta} - 1\right)^{2k-n} \left(e^{\mathrm{i}\theta} - e^{\mathrm{i}\beta}\right)^n . \tag{6.24}$$

In order to calculate this, we will let $\alpha_1, \ldots, \alpha_{2k}$ be distinct modulo 2π, and consider $D_N[f]$, the Toeplitz determinant with symbol

$$f(\theta) = (-1)^k e^{-\mathrm{i}k\theta} \prod_{j=1}^{2k} \left(e^{\mathrm{i}\theta} - e^{\mathrm{i}\alpha_j}\right). \tag{6.25}$$

In [7], Basor and Forrester calculate Toeplitz determinants with such symbols.

They show that $D_N[f]$ equals a $2k$ by $2k$ determinant:

$$D_N[f] = \prod_{1\le m<n\le 2k} \frac{1}{e^{\mathrm{i}\alpha_n} - e^{\mathrm{i}\alpha_m}} \times \det \begin{bmatrix} 1 & e^{\mathrm{i}\alpha_1} & \dots & e^{(k-1)\mathrm{i}\alpha_1} & e^{(N+k)\mathrm{i}\alpha_1} & e^{(N+k+1)\mathrm{i}\alpha_1} & \dots & e^{(N+2k-1)\mathrm{i}\alpha_1} \\ 1 & e^{\mathrm{i}\alpha_2} & \dots & e^{(k-1)\mathrm{i}\alpha_2} & e^{(N+k)\mathrm{i}\alpha_2} & e^{(N+k+1)\mathrm{i}\alpha_2} & \dots & e^{(N+2k-1)\mathrm{i}\alpha_2} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & e^{\mathrm{i}\alpha_{2k}} & \dots & e^{(k-1)\mathrm{i}\alpha_{2k}} & e^{(N+k)\mathrm{i}\alpha_{2k}} & e^{(N+k+1)\mathrm{i}\alpha_{2k}} & \dots & e^{(N+2k-1)\mathrm{i}\alpha_{2k}} \end{bmatrix} \tag{6.26}$$

To obtain $\mathbb{E}\left\{Z_U(0)^k Z_U^*(0)^{k-n} Z_U^*(\beta)^n\right\}$ we wish to set $\alpha_1 = \alpha_2 = \dots = \alpha_{2k-n} = 0$ and $\alpha_{2k-n+1} = \alpha_{2k} = \beta$. We achieve this by means of L'Hopital's rule, exactly as in [7]. First set $\alpha_1 = 0$ in ρ_1 (for simplicity we will denote the ith row by ρ_i). Then subtract ρ_1 from ρ_2 and let $\alpha_2 \to 0$, noting that

$$\lim_{\alpha_2\to 0} \frac{e^{\mathrm{i}n\alpha_2} - 1}{e^{\mathrm{i}\alpha_2} - 1} = \binom{n}{1} \tag{6.27}$$

implies this limit is well defined in all elements in the new second row. Repeat this procedure for the third row, the fourth row, and so on, up to the $(2k-n)^{\text{th}}$ row. (That is, to ρ_i add

$$-\rho_1 - (e^{\mathrm{i}\alpha_i} - 1)\rho_2 - (e^{\mathrm{i}\alpha_i} - 1)^2\rho_3 - \dots - (e^{\mathrm{i}\alpha_i} - 1)^{i-2}\rho_{i-1} \tag{6.28}$$

and let $\alpha_i \to 0$, remembering there is a $1/(e^{\mathrm{i}\alpha_i} - 1)^{i-1}$ factor outside the determinant).

This done,

$$D_N[f] = \prod_{j=2k-n+1}^{2k} \frac{1}{\left(e^{\mathrm{i}\alpha_j} - 1\right)^{2k-n}} \times \prod_{2k-n+1\le m<n\le 2k} \frac{1}{(e^{\mathrm{i}\alpha_n} - e^{\mathrm{i}\alpha_m})} \times \det \begin{bmatrix} 1 & \dots & 1 & 1 & \dots & 1 \\ \binom{0}{1} & \dots & \binom{k-1}{1} & \binom{N+k}{1} & \dots & \binom{N+2k-1}{1} \\ \vdots & & \vdots & \vdots & & \vdots \\ \binom{0}{2k-n-1} & \dots & \binom{k-1}{2k-n-1} & \binom{N+k}{2k-n-1} & \dots & \binom{N+2k-1}{2k-n-1} \\ 1 & \dots & e^{(k-1)\mathrm{i}\alpha_{2k-n+1}} & e^{(N+k)\mathrm{i}\alpha_{2k-n+1}} & \dots & e^{(N+2k-1)\mathrm{i}\alpha_{2k-n+1}} \\ \vdots & & \vdots & \vdots & & \vdots \\ 1 & \dots & e^{(k-1)\mathrm{i}\alpha_{2k}} & e^{(N+k)\mathrm{i}\alpha_{2k}} & \dots & e^{(N+2k-1)\mathrm{i}\alpha_{2k}} \end{bmatrix} \tag{6.29}$$

To ρ_i for $i = 2k-n+1, \dots, 2k$ add

$$-\sum_{m=1}^{2k-n} (e^{\mathrm{i}\alpha_i} - 1)^{m-1} \rho_m \tag{6.30}$$

and note that the new ρ_i has zero in the first k elements, and that when the $(k+j)$th element is divided by $(e^{\mathrm{i}\alpha_i} - 1)^{2k-n}$, it can be written as

$$z_j^{(i)} := \frac{e^{(N+k+j-1)\mathrm{i}\alpha_i} - \sum_{m=0}^{2k-n-1} \binom{N+k+j-1}{m} (e^{\mathrm{i}\alpha_i} - 1)^m}{(e^{\mathrm{i}\alpha_i} - 1)^{2k-n}} \tag{6.31}$$

$$= \sum_{m=0}^{N-k+n+j-1} \binom{N+k+j-1}{2k-n+m} (e^{\mathrm{i}\alpha_i} - 1)^m \tag{6.32}$$

We now have

$$\mathbb{E}\left\{ Z_U(0)^k Z_U^*(0)^{k-n} Z_U^*(\alpha_{2k-n+1}) \cdots Z_U^*(\alpha_{2k}) \right\} = \prod_{2k-n+1 \le m < n \le 2k} \frac{1}{(e^{\mathrm{i}\alpha_n} - e^{\mathrm{i}\alpha_m})} \times$$

$$\det \begin{bmatrix} 1 & \cdots & 1 & 1 & \cdots & 1 \\ \binom{0}{1} & \cdots & \binom{k-1}{1} & \binom{N+k}{1} & \cdots & \binom{N+2k-1}{1} \\ \vdots & & \vdots & \vdots & & \vdots \\ \binom{0}{2k-n-1} & \cdots & \binom{k-1}{2k-n-1} & \binom{N+k}{2k-n-1} & \cdots & \binom{N+2k-1}{2k-n-1} \\ 0 & \cdots & 0 & z_1^{(2k-n+1)} & \cdots & z_k^{(2k-n+1)} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & z_1^{(2k)} & \cdots & z_k^{(2k)} \end{bmatrix} \tag{6.33}$$

The first k diagonal elements are 1, with zeros underneath, and so the determinant reduces to a k by k one:

$$\prod_{2k-n+1 \le m < n \le 2k} \frac{1}{(e^{\mathrm{i}\alpha_n} - e^{\mathrm{i}\alpha_m})} \det \begin{bmatrix} \binom{N+k}{k} & \cdots & \binom{N+2k-1}{k} \\ \vdots & & \vdots \\ \binom{N+k}{2k-n-1} & \cdots & \binom{N+2k-1}{2k-n-1} \\ z_1^{(2k-n+1)} & \cdots & z_k^{(2k-n+1)} \\ \vdots & & \vdots \\ z_1^{(2k)} & \cdots & z_k^{(2k)} \end{bmatrix} \tag{6.34}$$

Finally, set $\alpha_{2k-n+1} = \beta$ in ρ_{k-n+1} of the above determinant, then subtract this new ρ_{k-n+1} from ρ_{k-n+2} and let $\alpha_{2k-n+2} \to \beta$ etc. (i.e. proceed as before: to row i, for $i = k-n+2, \dots, k$, add

$$-\rho_{k-n+1} - (e^{\mathrm{i}\alpha_{k+i}} - e^{\mathrm{i}\beta})\rho_{k-n+2} - \cdots - (e^{\mathrm{i}\alpha_{k+i}} - e^{\mathrm{i}\beta})^{i-k+n-2} \rho_{i-1} \tag{6.35}$$

and let $\alpha_{k+i} \to \beta$ recalling there is a $1/(e^{\mathrm{i}\alpha_{k+i}} - e^{\mathrm{i}\beta})^{i-k+n-1}$ factor outside the determinant). This done, the jth element of row i, for $i = k-n+1, \ldots, k$, equals

$$\sum_{m=i-k+n-1}^{N+j-k+n-1} \binom{N+k+j-1}{2k-n+m} \binom{m}{i-k+n-1} (e^{\mathrm{i}\beta} - 1)^{m-i+k-n+1} \tag{6.36}$$

which, after a slight renumbering of the sum, is seen to equal $z_{i,j}$ defined in (6.21), and this completes the proof of the lemma. □

Remark. Although we have calculated $T_{k,n}$ only for $0 \le n \le k$, it is valid for $0 \le n \le 2k$, and, in fact, we have $T_{k,n} = e^{\mathrm{i}N(k-n)\beta} T_{k,2k-n}$.

It is possible to further simplify the determinant appearing in lemma 6.2.

Lemma 6.3.

$$T_{k,n} = M_N(2k) \det \{a_{i,j}\}_{i,j=1,\ldots,n} \tag{6.37}$$

where $M_N(2k)$ is given by (1.77) *and*

$$a_{i,j} = \sum_{m=0}^{N+j-i} \frac{(2k-n+i-1)!}{(2k-n+i-1+m)!} \binom{i+k-n-1+m}{m} \binom{i+m-1}{j-1} \times \\ \times \frac{N!}{(N+j-i-m)!} (e^{\mathrm{i}\beta} - 1)^m \tag{6.38}$$

Proof. We start with $T_{k,n}$ as given in lemma 6.2. From row i (for $i = 1, \ldots, k$) pull out a factor of $\frac{1}{(k+i-1)!}$, and from column j (for $j = 1, \ldots, k$) pull out a factor of $\frac{(N+k+j-1)!}{(N+j-1)!}$. So now we have,

$$T_{k,n} = \frac{G(k+1)}{G(2k+1)} \frac{G(N+2k+1)G(N+1)}{G^2(N+k+1)} \det \begin{bmatrix} 1 & \cdots & 1 \\ \frac{N!}{(N-1)!} & \cdots & \frac{(N+k-1)!}{(N+k-2)!} \\ \vdots & & \vdots \\ \frac{N!}{(N-k+n+1)!} & \cdots & \frac{(N+k-1)!}{(N+n)!} \\ & x_{i,j}^{(0)} & \end{bmatrix} \tag{6.39}$$

where for $i = k-n+1, \ldots, k$ and $j = 1, \ldots, k$, $x_{i,j}^{(0)}$ equals

$$\sum_{m=0}^{N+j-i} \frac{(N+j-1)!}{(N+j-i-m)!} \frac{(k+i-1)!}{(k+i-1+m)!} \binom{m+i-k+n-1}{i-k+n-1} (e^{\mathrm{i}\beta} - 1)^m \tag{6.40}$$

From column k subtract column $k-1$. From column $k-1$ subtract column $k-2$ etc. The top row becomes $1, 0, \ldots, 0$, and so the above determinant equals a $k-1$

by $k-1$ determinant:

$$\det \begin{bmatrix} 1 & \cdots & 1 \\ 2\frac{N!}{(N-1)!} & \cdots & 2\frac{(N+k-2)!}{(N+k-3)!} \\ \vdots & & \vdots \\ (k-n-1)\frac{N!}{(N-k+n+2)!} & \cdots & (k-n-1)\frac{(N+k-2)!}{(N+n)!} \\ & x_{i,j}^{(1)} & \end{bmatrix} \tag{6.41}$$

where for $i = k-n, \ldots, k-1$ and $j = 1, \ldots, k-1$,

$$x_{i,j}^{(1)} = \sum_{m=0}^{N+j-i} (i+m) \frac{(N+j-1)!}{(N+j-i-m)!} \frac{(k+i)!}{(k+i+m)!} \binom{m+i-k+n}{i-k+n} (e^{\mathrm{i}\beta}-1)^m \tag{6.42}$$

Factor out i from row i, then repeat: subtract column $k-2$ from column $k-1$ etc., the new top row is $1, 0, \ldots, 0$, and expanding about this row, we obtain

$$(k-1)! \det \begin{bmatrix} 1 & \cdots & 1 \\ 2\frac{N!}{(N-1)!} & \cdots & 2\frac{(N+k-3)!}{(N+k-4)!} \\ \vdots & & \vdots \\ (k-n-2)\frac{N!}{(N-k+n+3)!} & \cdots & (k-n-2)\frac{(N+k-3)!}{(N+n)!} \\ & x_{i,j}^{(2)} & \end{bmatrix} \tag{6.43}$$

where for $i = k-n-1, \ldots, k-2$ and $j = 1, \ldots, k-2$,

$$x_{i,j}^{(2)} = \sum_{m=0}^{N+j-i} (i+m) \frac{(i+1+m)}{(i+1)} \frac{(N+j-1)!}{(N+j-i-m)!} \frac{(k+i+1)!}{(k+i+1+m)!} \times \\ \times \binom{m+i+1-k+n}{i+1-k+n} (e^{\mathrm{i}\beta}-1)^m \tag{6.44}$$

If we repeat this process enough times, then the determinant becomes an $n \times n$ one:

$$\frac{G(k+1)}{G(n+1)} \det \begin{bmatrix} x_{1,1}^{(k-n)} & \cdots & x_{1,n}^{(k-n)} \\ \vdots & & \vdots \\ x_{n,1}^{(k-n)} & \cdots & x_{n,n}^{(k-n)} \end{bmatrix} \tag{6.45}$$

where for $i = 1, \ldots, n$ and $j = 1, \ldots, n$,

$$x_{i,j}^{(k-n)} = \sum_{m=0}^{N+j-i} \frac{(i+m)\cdots(i+k-n-1+m)}{(i)\cdots(i+k-n-1)} \frac{(N+j-1)!}{(N+j-i-m)!} \times \\ \times \frac{(2k-n+i-1)!}{(2k-n+i-1+m)!} \binom{m+i-1}{i-1} (e^{\mathrm{i}\beta}-1)^m \tag{6.46}$$

which simplifies down to

$$x_{i,j}^{(k-n)} = \sum_{m=0}^{N+j-i} \frac{(N+j-1)!}{(N+j-i-m)!} \binom{i+k-n-1+m}{m} \times \\ \times \frac{(2k-n+i-1)!}{(2k-n+i-1+m)!} (e^{\mathrm{i}\beta}-1)^m. \quad (6.47)$$

Thus, to complete the proof, it remains to show that

$$\frac{1}{G(n+1)} \det\left\{x_{i,j}^{(k-n)}\right\}_{1\le i,j\le n} = \det\left\{a_{i,j}\right\}_{i,j=1,\dots,n}. \quad (6.48)$$

To do this, for $j = n, n-1, \dots, 1$ define

$$\text{new col}_j = \sum_{l=0}^{j-1} (-1)^l \binom{j}{l} \text{old col}_{j-l} \quad (6.49)$$

(this does not change the value of the determinant). Note that

$$\sum_{l=0}^{j-1} (-1)^l \binom{j}{l} \frac{(N+j-l-1)!}{(N+j-l-i-m)!} = \frac{N!}{(N+j-i-m)!} \prod_{l=1}^{j-1} (i+m-l). \quad (6.50)$$

Multiplying new jth column by $1/(j-1)!$, for $j = 1, \dots, n$ (which will come from $1/G(n+1)$), we see that $x_{i,j} = a_{i,j}$ for $1 \le i, j \le n$, and this completes the proof of lemma 6.3. □

Corollary 6.3.1. *For fixed h and fixed k,*

$$T_{k,n} \sim \frac{G^2(k+1)}{G(2k+1)} N^{k^2} \det\left\{b_{i,j}\right\}_{1\le i,j\le n} + \mathcal{O}_N(\beta^{2h+1}) \quad (6.51)$$

where, for $i, j = 1, \dots, n$

$$b_{i,j} = \sum_{m=0}^{2h} \frac{(2k-n+i-1)!}{(2k-n+i-1+m)!} \binom{i+k-n-1+m}{m} \binom{i+m-1}{j-1} (\mathrm{i}N\beta)^m \quad (6.52)$$

and where $\sim$ means the leading order (in N) term in each *coefficient of β^n.*

Proof. First of all, recall that

$$M_N(2k) = \frac{G^2(k+1)}{G(2k+1)} N^{k^2} + \mathcal{O}\left(N^{k^2-1}\right) \quad (6.53)$$

and secondly note that

$$a_{i,j} \sim \sum_{m=0}^{2h} \frac{(2k-n+i-1)!}{(2k-n+i-1+m)!} \binom{i+k-n-1+m}{m} \binom{i+m-1}{j-1} \times \\ \times N^{m+i-j} \left((\mathrm{i}\beta)^m + \mathcal{O}(\beta^{m+1})\right) + \mathcal{O}_N(\beta^{2h+1}) \quad (6.54)$$

(to get this, first Taylor expand in terms of β and then find the leading order (in N) term in each coefficient of the Taylor expansion). To show that

$$\det\{a_{i,j}\}_{1\le i,j\le n} \sim \det\{b_{i,j}\}_{1\le i,j\le n} \tag{6.55}$$

multiply row i by $\frac{1}{N^i}$ and column j by N^j for each i and j. (This does not effect the value of the determinant, just the entries in the matrix). □

Remark. The definition of a Schur polynomial in $2k$ variables is

$$S_\lambda(\boldsymbol{x}) = \frac{\det\{x_j^{\lambda_i+2k-i}\}_{1\le i,j\le 2k}}{\det\{x_j^{2k-i}\}_{1\le i,j\le 2k}} \tag{6.56}$$

where $\lambda = (\lambda_1, \dots, \lambda_{2k})$ is a partition (which means that $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_{2k} \ge 0$) and $\boldsymbol{x} = (x_1, \dots, x_{2k})$. The bottom determinant is the usual Vandermonde determinant. Comparing this with (6.26) we see that

$$D_N[f] = S_{(N,\dots,N,0,\dots,0)}(e^{\mathrm{i}\alpha_1}, \dots, e^{\mathrm{i}\alpha_{2k}}) \tag{6.57}$$

where the partition is $\lambda = 0^k N^k$. Setting $\alpha_1 = \cdots = \alpha_{2k-n} = 0$ and $\alpha_{2k-n+1} = \cdots = \alpha_{2k} = \beta$ we have

$$T_{k,n} = S_\lambda(1, \dots, 1, e^{\mathrm{i}\beta}, \dots, e^{\mathrm{i}\beta}). \tag{6.58}$$

But evaluating the Schur polynomial as a polynomial in β appears to be just as difficult as evaluating the determinants, as done above.

6.2.2 Calculating $\widetilde{F}(h,k)$

Keeping the notation of corollary 6.3.1, we write

$$T_{k,n} \sim \frac{G^2(k+1)}{G(2k+1)} N^{k^2} \sum_{j=0}^{2h} \alpha_{k,n}^{(j)} (\mathrm{i}N\beta)^j + \mathcal{O}_N\left(\beta^{2h+1}\right) \tag{6.59}$$

Theorem 6.4. *For integer $h \ge 1$ and $k \ge h$ an integer*

$$\begin{aligned}\mathbb{E}\left\{V_U(0)^{2k}\left(\frac{V_U'(0)}{V_U(0)}\right)^{2h}\right\} &= \lim_{\beta\to 0}\frac{1}{\beta^{2h}}\sum_{n=0}^{2h}(-1)^n\binom{2h}{n}e^{-\mathrm{i}Nn\beta/2}T_{k,n} \\ &\sim \frac{G^2(k+1)}{G(2k+1)}\sum_{n=1}^{2h}\sum_{m=0}^{2h}(-1)^{h+m+n}\binom{2h}{n}\binom{2h}{m}\left(\tfrac{1}{2}n\right)^m \alpha_{k,n}^{(2h-m)} N^{k^2+2h}\end{aligned} \tag{6.60}$$

as $N \to \infty$.

Note that this is actually an analytic function of k for $\mathfrak{Re}(k) > h - 1/2$, but the restriction of h to be a positive integer is, at least in this method of proof, a genuine one.

Proof. By the definition of derivative we have

$$\mathbb{E}\left\{V_U(0)^{2k-2h}V_U'(0)^{2h}\right\} = \lim_{\beta\to 0}\frac{1}{\beta^{2h}}\mathbb{E}\left\{V_U(0)^{2k-2h}(V_U(\beta)-V_U(0))^{2h}\right\} \tag{6.61}$$

$$= \lim_{\beta\to 0}\frac{1}{\beta^{2h}}\sum_{n=0}^{2h}(-1)^n\binom{2h}{n}\mathbb{E}\left\{V_U(0)^{2k-n}V_U(\beta)^n\right\} \tag{6.62}$$

From the definitions of $V_U(\theta)$ and $Z_U(\theta)$, one can see that

$$V_U(0)^{2k-n}V_U(\beta)^n = e^{-\mathrm{i}Nn\beta/2}Z_U(0)^k Z_U^*(0)^{k-n}Z_U^*(\beta)^n \tag{6.63}$$

and so

$$\mathbb{E}\left\{V_U(0)^{2k-2h}V_U'(0)^{2h}\right\} = \lim_{\beta\to 0}\frac{1}{\beta^{2h}}\sum_{n=0}^{2h}(-1)^n\binom{2h}{n}e^{-\mathrm{i}Nn\beta/2}T_{k,n}$$

$$\sim \frac{G^2(k+1)}{G(2k+1)}\sum_{n=0}^{2h}\sum_{m=0}^{2h}(-1)^{h+m+n}\binom{2h}{n}\binom{2h}{m}\left(\tfrac{1}{2}n\right)^m\alpha_{k,n}^{(2h-m)}N^{k^2+2h} \tag{6.64}$$

(We have cheated slightly here, by writing down the $\beta\to 0$ limit as the coefficient of β^{2h} in the sum. There is correct so long as there are no terms of $\mathcal{O}(\beta^{2h-1})$ or less. But the fact that the derivative is defined means that no such terms can exist). The $n=0$ term does not exist for $h\neq 0$ due to the fact that $T_{k,0}$ has no β dependence.

□

Remark. $\left\{|Z_U(0)|^{2k-2h}|Z_U'(0)|^{2h}\right\}$ can be dealt with using exactly the same ideas, but the formulae are slightly messier and so are not included here. Such work is unnecessary, anyway, since the results in §6.3 show that $\mathbb{E}\left\{|Z_U(0)|^{2k-2h}|Z_U'(0)|^{2h}\right\}$ can be calculated from $\mathbb{E}\left\{V_U(0)^{2k-2h}V_U'(0)^{2h}\right\}$.

6.2.3 Worked example: $h = 1$

From theorem 6.4 we have

$$\mathbb{E}\left\{V_U(0)^{2k-2}V_U'(0)^2\right\} = \lim_{\beta\to 0}\frac{1}{\beta^2}\left(T_{k,0} - 2e^{-\mathrm{i}N\beta/2}T_{k,1} + e^{-\mathrm{i}N\beta}T_{k,2}\right) \tag{6.65}$$

From corollary 6.3.1 we have

$$T_{k,1} \sim \frac{G^2(k+1)}{G(2k+1)}N^{k^2}\sum_{m=0}^{2}\frac{(2k-1)!}{(2k-1+m)!}\binom{k+m-1}{m}(N\mathrm{i}\beta)^m + \mathcal{O}_N(\beta^3) \tag{6.66}$$

and

$$T_{k,2} \sim \frac{G^2(k+1)}{G(2k+1)} N^{k^2}(x_1 y_2 - x_2 y_1) + \mathcal{O}_N(\beta^3) \tag{6.67}$$

where

$$x_1 = \sum_{m=0}^{2} \frac{(2k-2)!}{(2k-2+m)!} \binom{k+m-2}{m} (N\mathrm{i}\beta)^m \tag{6.68}$$

$$x_2 = \sum_{m=1}^{2} \frac{(2k-2)!}{(2k-2+m)!} \binom{k+m-2}{m} m(N\mathrm{i}\beta)^m \tag{6.69}$$

$$y_1 = \sum_{m=0}^{2} \frac{(2k-1)!}{(2k-1+m)!} \binom{k+m-1}{m} (N\mathrm{i}\beta)^m \tag{6.70}$$

$$y_2 = \sum_{m=0}^{2} \frac{(2k-1)!}{(2k-1+m)!} \binom{k+m-1}{m} (m+1)(N\mathrm{i}\beta)^m \tag{6.71}$$

Thus

$$T_{k,2} \sim \frac{G^2(k+1)}{G(2k+1)} N^{k^2} \left(1 + N(\mathrm{i}\beta) + \frac{4k^2+k-2}{2(4k^2-1)} N^2(\mathrm{i}\beta)^2 + \mathcal{O}_N(\beta^3)\right) \tag{6.72}$$

And so, if we let $[\beta^n \ : \ \ldots]$ denote "the coefficient of β^n in $\ldots$", then for large N we have

$$\left[\beta^2 \ : \ e^{-\mathrm{i}N\beta} T_{k,2}\right] \sim \frac{G^2(k+1)}{G(2k+1)} \frac{1-k}{2(4k^2-1)} N^{k^2+2} \tag{6.73}$$

Similarly, for large N, (6.66) gives

$$T_{k,1} \sim \frac{G^2(k+1)}{G(2k+1)} N^{k^2} \left(1 + \tfrac{1}{2} N(\mathrm{i}\beta) + \frac{k+1}{4(2k+1)} N^2(\mathrm{i}\beta)^2 + \mathcal{O}_N(\beta^3)\right) \tag{6.74}$$

and so

$$\left[\beta^2 \ : \ e^{-\mathrm{i}N\beta/2} T_{k,1}\right] \sim \frac{G^2(k+1)}{G(2k+1)} \frac{-1}{8(2k+1)} N^{k^2+2} \tag{6.75}$$

Since $T_{k,0}$ has no β dependence, inserting (6.73) and (6.75) into (6.65), we get

$$\mathbb{E}\left\{V_U(0)^{2k-2} V_U'(0)^2\right\} = \left[\beta^2 \ : \ e^{-\mathrm{i}N\beta} T_{k,2} - 2e^{-\mathrm{i}N\beta/2} T_{k,1}\right] \tag{6.76}$$

$$\sim \frac{G^2(k+1)}{G(2k+1)} \frac{1}{4(4k^2-1)} N^{k^2+2} \tag{6.77}$$

To get the corresponding result for $|Z_U|^{2k-2}|Z_U'|^2$, we make use of the work in §6.3, which says

$$F(1,k) = \tfrac{1}{4}\widetilde{F}(0,k) + \widetilde{F}(1,k) \tag{6.78}$$

$$= \frac{1}{4}\frac{G^2(k+1)}{G(2k+1)} + \frac{G^2(k+1)}{G(2k+1)} \frac{1}{4(4k^2-1)} \tag{6.79}$$

$$= \frac{G^2(k+1)}{G(2k+1)} \frac{k^2}{4k^2-1} \tag{6.80}$$

The conjecture for the Riemann zeta function would be

$$\frac{1}{T}\int_0^T |\zeta(\tfrac{1}{2}+\mathrm{i}t)|^{2k-2}|\zeta'(\tfrac{1}{2}+\mathrm{i}t)|^2\,\mathrm{d}t \sim \frac{G^2(k+1)}{G(2k+1)}\frac{k^2}{4k^2-1}a(k)\left(\log\frac{T}{2\pi}\right)^{k^2+2} \tag{6.81}$$

which agrees with what is known (summarized in §6.1.1), since $a(1)=1$ and $a(2)=\frac{6}{\pi^2}$.

6.2.4 General formula: $h=2$ and $h=3$

Doing the analogous thing when $h=2$, we get

$$\widetilde{F}(2,k) = \frac{G^2(k+1)}{G(2k+1)}\frac{3}{16}\frac{1}{(4k^2-1)(4k^2-9)} \tag{6.82}$$

and

$$F(2,k) = \frac{G^2(k+1)}{G(2k+1)}\frac{8k^4-16k^2-3}{8(4k^2-1)(4k^2-9)} \tag{6.83}$$

Also, for $h=3$,

$$\widetilde{F}(3,k) = \frac{G^2(k+1)}{G(2k+1)}\frac{15}{64}\frac{1}{(4k^2-1)^2(4k^2-25)} \tag{6.84}$$

so

$$F(3,k) = \frac{G^2(k+1)}{G(2k+1)}\frac{32k^8-264k^6+396k^4+37k^2-45}{8(4k^2-1)^2(4k^2-9)(4k^2-25)} \tag{6.85}$$

These examples were calculated with the aid of Maple. It appears clear from these results that theorem 6.4 can be simplified, but for general integer h I am currently unable to put $\widetilde{F}(h,k)$ in to a form similar to the above.

6.3 Converting between the real and complex cases

6.3.1 Converting from $\widetilde{F}(h,k)$ to $F(h,k)$

By logarithmically differentiating the definition of $V_U(\theta)$ we have

$$\frac{V_U'(\theta)}{V_U(\theta)} = \tfrac{1}{2}\mathrm{i}N + \frac{Z_U'(\theta)}{Z_U(\theta)}. \tag{6.86}$$

But

$$\frac{V_U'(\theta)}{V_U(\theta)} = \frac{V_U^{*\prime}(\theta)}{V_U^*(\theta)} \tag{6.87}$$

$$= -\tfrac{1}{2}\mathrm{i}N + \frac{Z_U^{*\prime}(\theta)}{Z_U^*(\theta)}, \tag{6.88}$$

and so

$$\mathfrak{Im}\frac{Z_U'(\theta)}{Z_U(\theta)} = -\tfrac{1}{2}N \tag{6.89}$$

and

$$\frac{V_U'(\theta)}{V_U(\theta)} = \mathfrak{Re}\frac{Z_U'(\theta)}{Z_U(\theta)}. \tag{6.90}$$

Now,

$$|Z_U|^{2k}\left|\frac{Z_U'}{Z_U}\right|^{2h} = |Z_U|^{2k}\left(\left(\mathfrak{Re}\frac{Z_U'}{Z_U}\right)^2 + \left(\mathfrak{Im}\frac{Z_U'}{Z_U}\right)^2\right)^h \tag{6.91}$$

$$= \sum_{n=0}^{h}\binom{h}{n}|Z_U|^{2k}\left(\mathfrak{Re}\frac{Z_U'}{Z_U}\right)^{2n}\left(\mathfrak{Im}\frac{Z_U'}{Z_U}\right)^{2h-2n} \tag{6.92}$$

$$= \sum_{n=0}^{h}\binom{h}{n}|V_U|^{2k}\left(\frac{V_U'}{V_U}\right)^{2n}\frac{N^{2h-2n}}{2^{2h-2n}}, \tag{6.93}$$

so taking expectations, and letting $N \to \infty$, we obtain

$$F(h,k) = \sum_{n=0}^{h}\binom{h}{n}\frac{\widetilde{F}(n,k)}{2^{2h-2n}}. \tag{6.94}$$

By inverting this we get

$$\widetilde{F}(h,k) = \sum_{n=0}^{h}(-1)^n\binom{h}{n}\frac{F(h-n,k)}{2^{2n}}. \tag{6.95}$$

6.3.2 Converting from $\widetilde{C}(h,k)$ to $C(h,k)$

Recall the functional equation, $\zeta(1-s) = \chi(1-s)\zeta(s)$. Logarithmically differentiating this, we get

$$\frac{\chi'}{\chi}(\tfrac{1}{2}-\mathrm{i}t) = \frac{\zeta'}{\zeta}(\tfrac{1}{2}+\mathrm{i}t) + \frac{\zeta'}{\zeta}(\tfrac{1}{2}-\mathrm{i}t) \tag{6.96}$$

$$= 2\mathfrak{Re}\frac{\zeta'}{\zeta}(\tfrac{1}{2}+\mathrm{i}t). \tag{6.97}$$

From this follows two results:

Firstly,

$$\mathfrak{Re}\frac{\zeta'}{\zeta}(\tfrac{1}{2}+\mathrm{i}t) \sim -\tfrac{1}{2}\log\frac{T}{2\pi}, \tag{6.98}$$

and secondly, from the definition $\mathcal{Z}(t) = \sqrt{\chi(\frac{1}{2}-\mathrm{i}t)}\zeta(\frac{1}{2}+\mathrm{i}t)$, we have

$$\frac{\mathcal{Z}'}{\mathcal{Z}}(t) = -\tfrac{1}{2}\mathrm{i}\frac{\chi'}{\chi}(\tfrac{1}{2}-\mathrm{i}t) + \mathrm{i}\frac{\zeta'}{\zeta}(\tfrac{1}{2}+\mathrm{i}t) \tag{6.99}$$

$$= -\mathfrak{Im}\frac{\zeta'}{\zeta}(\tfrac{1}{2}+\mathrm{i}t). \tag{6.100}$$

Now

$$|\zeta|^{2k}\left|\frac{\zeta'}{\zeta}\right|^{2h} = |\zeta|^{2k}\left(\left(\mathfrak{Re}\frac{\zeta'}{\zeta}\right)^2 + \left(\mathfrak{Im}\frac{\zeta'}{\zeta}\right)^2\right)^h \tag{6.101}$$

$$= \sum_{n=0}^{h}\binom{h}{n}|\zeta|^{2k}\left(\mathfrak{Im}\frac{\zeta'}{\zeta}\right)^{2n}\left(\mathfrak{Re}\frac{\zeta'}{\zeta}\right)^{2h-2n} \tag{6.102}$$

so we have

$$\frac{1}{T}\int_0^T |\zeta(\tfrac{1}{2}+\mathrm{i}t)|^{2k}\left|\frac{\zeta'}{\zeta}(\tfrac{1}{2}+\mathrm{i}t)\right|^{2h}\mathrm{d}t = \sum_{n=0}^{h}\binom{h}{n}\frac{1}{T}\int_0^T |\mathcal{Z}(t)|^{2k}\left(-\frac{\mathcal{Z}'(t)}{\mathcal{Z}(t)}\right)^{2n}\left(\mathfrak{Re}\frac{\zeta'}{\zeta}(\tfrac{1}{2}+\mathrm{i}t)\right)^{2h-2n}\mathrm{d}t \tag{6.103}$$

so

$$\frac{1}{T}\int_0^T |\zeta(\tfrac{1}{2}+\mathrm{i}t)|^{2k}\left|\frac{\zeta'}{\zeta}(\tfrac{1}{2}+\mathrm{i}t)\right|^{2h}\mathrm{d}t \sim \sum_{n=0}^{h}\binom{h}{n}\widetilde{C}(n,k)\left(\log\frac{T}{2\pi}\right)^{k^2+2n}\frac{1}{2^{2h-2n}}\left(\log\frac{T}{2\pi}\right)^{2h-2n}. \tag{6.104}$$

Thus,

$$C(h,k) = \sum_{n=0}^{h}\binom{h}{n}\frac{\widetilde{C}(n,k)}{2^{2h-2n}}. \tag{6.105}$$

By inverting this:

$$\widetilde{C}(h,k) = \sum_{n=0}^{h}(-1)^n\binom{h}{n}\frac{C(h-n,k)}{2^{2n}}. \tag{6.106}$$

6.3.3 Explaining the difference

We have

$$\mathfrak{Im}\frac{Z_U'(\theta)}{Z_U(\theta)} = -\tfrac{1}{2}N \tag{6.107}$$

and

$$\mathfrak{Re}\frac{Z_U'(\theta)}{Z_U(\theta)} = \frac{V_U'(\theta)}{V_U(\theta)}. \tag{6.108}$$

But we have

$$\mathfrak{Re}\frac{\zeta'}{\zeta}(\tfrac{1}{2}+\mathrm{i}t) \sim -\tfrac{1}{2}\log\frac{T}{2\pi} \tag{6.109}$$

and

$$\mathfrak{Im}\frac{\zeta'}{\zeta}(\tfrac{1}{2}+\mathrm{i}t) = -\frac{\mathcal{Z}'}{\mathcal{Z}}(t). \tag{6.110}$$

The resolution of this difference is we should really be considering the function

$$\zeta_c(t) := \zeta(\tfrac{1}{2} + \mathrm{i}t), \tag{6.111}$$

since then

$$\zeta_c'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\zeta(\tfrac{1}{2} + \mathrm{i}t) \tag{6.112}$$

$$= \mathrm{i}\zeta'(\tfrac{1}{2} + \mathrm{i}t), \tag{6.113}$$

and so

$$\mathfrak{Im}\frac{\zeta_c'(t)}{\zeta(t)} \sim -\tfrac{1}{2}\log\frac{T}{2\pi} \tag{6.114}$$

and

$$\mathfrak{Re}\frac{\zeta_c'(t)}{\zeta_c(t)} = \frac{\mathcal{Z}'(t)}{\mathcal{Z}(t)} \tag{6.115}$$

which, since $N = \log\frac{T}{2\pi}$ from (1.72), is in perfect agreement with the random matrix results. Since we have only considered the modulus of zeta in this chapter, the missing factors of i are in fact irrelevant.

Appendix A

The Barnes G–function and the Euler Gamma function

The Barnes G–function is defined [5] for all z by

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

where $\gamma = 0.5772\ldots$ is Euler's constant, or alternatively via the integral

$$\log G(z+1) = \int_0^\infty \frac{e^{-t}}{t(1-e^{-t})^2}\left(1 - zt - \tfrac{1}{2}z^2t^2 - e^{-zt}\right) \mathrm{d}t - \tfrac{1}{2}z^2(1+\gamma) + \tfrac{1}{2}z \log\frac{2\pi}{e} \tag{A.2}$$

where $\mathfrak{Re}(z+1) > 0$.

The G-function has the following properties [5, 94]:

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

Complex conjugation: $G^*(z) = G(z^*)$.

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

$$\log G(z+1) \sim z^2\left(\tfrac{1}{2}\log z - \tfrac{3}{4}\right) + \tfrac{1}{2}z\log 2\pi - \tfrac{1}{12}\log z + \zeta'(-1) + \mathcal{O}\left(\frac{1}{z}\right) \tag{A.3}$$

Taylor expansion for $|z| < 1$,

$$\log G(z+1) = \tfrac{1}{2}(\log 2\pi - 1)z - \tfrac{1}{2}(1+\gamma)z^2 + \sum_{n=3}^{\infty}(-1)^{n-1}\zeta(n-1)\frac{z^n}{n} \tag{A.4}$$

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

$G(z+1)$ has zeros at $z = -n$ of order n, where $n = 1, 2, \ldots$.

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

$$\frac{\mathrm{d}^{n+1}}{\mathrm{d}z^{n+1}} \log G(z) = \Phi^{(n)}(z) \tag{A.5}$$

where

$$\Phi^{(0)}(z) = \tfrac{1}{2} \log 2\pi - z + \tfrac{1}{2} + (z-1)\Psi^{(0)}(z). \tag{A.6}$$

and

$$\Psi^{(n)}(z) = \frac{\mathrm{d}^{n+1}}{\mathrm{d}z^{n+1}} \log \Gamma(z). \tag{A.7}$$

Euler's Gamma function is discussed in [1]. It is defined for $\mathfrak{Re}(z) > 0$ by

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, \mathrm{d}t \tag{A.8}$$

For $|z| < \infty$ Hankel's contour integral is

$$\frac{1}{\Gamma(z)} = \frac{\mathrm{i}}{2\pi} \int_C (-t)^{-z} e^{-t} \, \mathrm{d}t \tag{A.9}$$

where C starts at $+\infty$ on the real axis, circles the origin once in the counterclockwise direction, and returns to the starting point. With the above, this gives a functional equation:

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)} \tag{A.10}$$

For $|z| < \infty$, Euler's infinite product is

$$\frac{1}{\Gamma(z)} = z e^{\gamma z} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-z/n} \tag{A.11}$$

The Gamma function has the following properties:

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

Stirling's asymptotic formula, valid for $|z| \to \infty$ with $|\arg(z)| < \pi$,

$$\log \Gamma(z) \sim (z - \tfrac{1}{2}) \log z - z + \tfrac{1}{2} \log(2\pi) + \frac{1}{12z} - \frac{1}{360z^3} + \cdots + \frac{B_{2m}}{2m(2m-1)z^{2m-1}} + \cdots \tag{A.12}$$

where the B_{2m} are the Bernoulli numbers, defined by

$$\sum_{n=0}^{\infty} B_n \frac{t^n}{n!} = \frac{t}{e^t - 1} \tag{A.13}$$

There is a Taylor expansion for $|z| < 2$:

$$\log \Gamma(z+1) = -\log(z+1) + z(1-\gamma) + \sum_{n=2}^{\infty} (-1)^n (\zeta(n) - 1) \frac{z^n}{n} \tag{A.14}$$

Special values: $\Gamma(1) = 1$ and $\Gamma(1/2) = \sqrt{\pi}$.

$\Gamma(z)$ has simple poles at $z = -n$, $n = 0, 1, 2, \ldots$ of residue $(-1)^n/n!$.

Appendix B

The Lambert W–function

The Lambert W-function (sometimes called the Omega function) is defined to be the solution of

$$W(x)e^{W(x)} = x \tag{B.1}$$

It has a branch point at $x = 0$, and is double real-valued for $-e^{-1} < x < 0$.

The unique branch that is analytic at the origin is called the principal branch. It is real in the domain $-e^{-1} < x < \infty$, with a range -1 to ∞. The second real branch is referred to as the -1 branch, denoted W_{-1}. It is real in the domain $-e^{-1} < x < 0$, with a range $-\infty$ to -1.

The equation

$$\log x = vx^{\beta} \tag{B.2}$$

has solution

$$x = \exp\left(\frac{-W(-\beta v)}{\beta}\right) \tag{B.3}$$

There are various asymptotic expansions of the W function:

- As $x \to \infty$,

$$W(x) \sim \log x - \log\log x + \frac{\log\log x}{\log x} \tag{B.4}$$

- As $x \to 0$ on the principal branch,

$$W(x) \sim x - x^2 + \tfrac{3}{2}x^3 \tag{B.5}$$

- As $x \to 0^-$ on the -1 branch,

$$W(x) \sim \log|x| - \log|\log|x|| + \frac{\log|\log|x||}{\log|x|} \tag{B.6}$$

Appendix C

Asymptotics of $\log M_N(x)$

$M_N(x)$ is given by (1.77). From the asymptotics for the G-function, (A.3), we have for $x > -1$

$$\begin{aligned}\log M_N(x) = 2\log G\left(1+\tfrac{1}{2}x\right) - \log G(1+x) - \tfrac{3}{8}x^2 + \tfrac{1}{2}N^2\log N \\ + \tfrac{1}{2}(N+x)^2\log(N+x) - \left(N+\tfrac{1}{2}x\right)^2\log\left(N+\tfrac{1}{2}x\right) \\ + \tfrac{1}{6}\log\left(N+\tfrac{1}{2}x\right) - \tfrac{1}{12}\log(N+x) - \tfrac{1}{12}\log N + \mathcal{O}\left(\frac{1}{N}\right)\end{aligned} \tag{C.1}$$

where the error term is independent of x.

This may be simplified if we assume that $x(N)$ is restricted to various regimes:

- If $|x| \ll 1$ then

$$\log M_N(x) = \tfrac{1}{4}x^2(\log N + 1 + \gamma) + \mathcal{O}\left(x^3\right) + \mathcal{O}\left(\frac{1}{N}\right) \tag{C.2}$$

- If $x = \mathcal{O}(1)$ and $x > -1$ then

$$\log M_N(x) = \tfrac{1}{4}x^2\log N + 2\log G\left(1+\tfrac{1}{2}x\right) - \log G(1+x) + \mathcal{O}\left(\frac{1}{N}\right) \tag{C.3}$$

- If $1 \ll x \ll \sqrt[3]{N}$ then

$$\begin{aligned}\log M_N(x) = \tfrac{1}{4}x^2\left(\log N - \log x - \log 2 + \tfrac{3}{2}\right) + \tfrac{1}{6}\log 2 - \tfrac{1}{12}\log x + \zeta'(-1) \\ + \mathcal{O}\left(\frac{x^3}{N}\right) + \mathcal{O}\left(\frac{1}{x}\right)\end{aligned} \tag{C.4}$$

- If $x = \lambda N$ with $\lambda = \mathcal{O}(1)$ and $\lambda > 0$ then

$$\begin{aligned}\log M_N(x) = N^2 \Big\{ \tfrac{1}{2}(1+\lambda)^2 \log(1+\lambda) - \big(1+\tfrac{1}{2}\lambda\big)^2 \log\big(1+\tfrac{1}{2}\lambda\big) \\ -\tfrac{1}{4}\lambda^2 \log(2\lambda)\Big\} - \tfrac{1}{12}\log N - \tfrac{1}{12}\log\lambda + \zeta'(-1) \\ + \tfrac{1}{6}\log(2+\lambda) - \tfrac{1}{12}\log(1+\lambda) + \mathcal{O}\left(\frac{1}{N}\right) \qquad \text{(C.5)}\end{aligned}$$

Appendix D

Asymptotics of $\log L_N(\mathrm{i}x)$

$L_N(\mathrm{i}x)$ is given by (1.80). From the asymptotics for the G-function, (A.3), if $x \in \mathbb{R}$ we have

$$\begin{aligned}\log L_N(\mathrm{i}x) &= \log G\left(1+\tfrac{1}{2}\mathrm{i}x\right) + \log G\left(1-\tfrac{1}{2}\mathrm{i}x\right) - \tfrac{3}{8}x^2 + N^2\log N \\ &\quad - \tfrac{1}{2}\left(N+\tfrac{1}{2}\mathrm{i}x\right)^2\log\left(N+\tfrac{1}{2}\mathrm{i}x\right) - \tfrac{1}{2}\left(N-\tfrac{1}{2}\mathrm{i}x\right)^2\log\left(N-\tfrac{1}{2}\mathrm{i}x\right) \\ &\quad - \tfrac{1}{6}\log N + \tfrac{1}{12}\log\left(N+\tfrac{1}{2}\mathrm{i}x\right) + \tfrac{1}{12}\log\left(N-\tfrac{1}{2}\mathrm{i}x\right) + \mathcal{O}\left(\frac{1}{N}\right) \qquad \text{(D.1)}\end{aligned}$$

Constraining $x(N)$ to lie in various regimes simplifies the above considerably:

- If $|x| \ll 1$ then

$$\log L_N(\mathrm{i}x) = \tfrac{1}{4}x^2(\log N + 1 + \gamma) + \mathcal{O}(x^4) + \mathcal{O}\left(\frac{1}{N}\right) \qquad \text{(D.2)}$$

- If $x = \mathcal{O}(1)$ then

$$\log L_N(\mathrm{i}x) = \log G\left(1+\tfrac{1}{2}\mathrm{i}x\right) + \log G\left(1-\tfrac{1}{2}\mathrm{i}x\right) + \tfrac{1}{4}x^2\log N + \mathcal{O}\left(\frac{1}{N}\right) \qquad \text{(D.3)}$$

- If $1 \ll |x| \ll \sqrt{N}$ then

$$\begin{aligned}\log L_N(\mathrm{i}x) &= \tfrac{1}{4}x^2\left(\log N - \log x + \log 2 + \tfrac{3}{2}\right) - \tfrac{1}{6}\log x + \tfrac{1}{6}\log 2 + 2\zeta'(-1) \\ &\quad + \mathcal{O}\left(\frac{x^4}{N^2}\right) + \mathcal{O}\left(\frac{1}{x^2}\right) \qquad \text{(D.4)}\end{aligned}$$

- If $x = \lambda N$ with $\lambda = \mathcal{O}(1)$ then

$$\begin{aligned}\log L_N(\mathrm{i}x) &= N^2\left\{\tfrac{1}{8}\lambda^2\log\left(1+4\lambda^{-2}\right) - \tfrac{1}{2}\log\left(1+\tfrac{1}{4}\lambda^2\right) + \lambda\tan^{-1}\tfrac{1}{2}\lambda\right\} \\ &\quad - \tfrac{1}{6}\log N + \tfrac{1}{12}\log\left(1+4\lambda^{-2}\right) + 2\zeta'(-1) + \mathcal{O}\left(\frac{1}{N}\right) \qquad \text{(D.5)}\end{aligned}$$

Bibliography

[1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, ninth printing 1972)

[2] T.H. Baker and P.J. Forrester, "Finite-N fluctuation formulas for random matrices", *J. Stat. Phys.* **88** (1997) 1371–1386

[3] R. Balasubramanian and K. Ramachandra, "On the frequency of Titchmarsh's phenomenon for $\zeta(s)$ III", *Proc. Indian Acad. Sci. A* **86** (1977) 341–351

[4] R. Balasubramanian and K. Ramachandra, "Proof of some conjectures on the mean-value of Titchmarsh series. I" *Hardy-Ramanujan J.* **13** (1990) 1–20

[5] E.W. Barnes, "The theory of the G–function", *Quart. J. Pure Appl. Math.* **31** (1899) 264–314

[6] E.L. Basor, "Asymptotic formulas for Toeplitz determinants", *Trans. Amer. Math. Soc.* **239** (1978) 33–65

[7] E.L. Basor and P.J. Forrester, "Formulas for the evaluation of Toeplitz determinants with rational generating functions", *Math. Nachr.* **170** (1994) 5–18

[8] G. Ben Arous and A. Guionnet, "Large deviations for Wigner's law and Voiculescu's non-commutative entropy", *Probab. Theor. Rel. Fields* **108** (1997) 517–542

[9] M.V. Berry, "Semiclassical formula for the number variance of the Riemann zeros", *Nonlinearity* **1** (1988) 399–407

[10] M.V. Berry and J.P. Keating, "The Riemann zeros and eigenvalue asymptotics", *SIAM Rev.* **41** (1999) 236–266

[11] P. Billingsley, *Convergence of Probability Measures*, (John Wiley & Sons, 1968)

[12] P. Billingsley, *Probability and Measure*, (John Wiley & Sons, 1979)

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

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

[15] E.B. Bogomolny and J.P. Keating, "Gutzwiller's trace formula and spectral statistics: beyond the diagonal approximation" *Phys. Rev. Lett.* **77** (1996) 1472–1475

[16] H. Bohr and B. Jessen, "Über die Werteverteilung der Riemannschen Zetafunktion", *Acta. Math.* **54** (1930) 1–35

[17] H. Bohr and B. Jessen, "Über die Werteverteilung der Riemannschen Zetafunktion", *Acta. Math.* **58** (1932) 1–55

[18] A. Böttcher and B. Silbermann, *Introduction to Large Truncated Toeplitz Matrices* (Springer-Verlag, 1999)

[19] E. Brézin and S. Hikami, "Characteristic polynomials of random matrices", *Comm. Math. Phys.* **214** (2000) 111–135

[20] J. Bucklew, *Large Deviation Techniques in Decision, Simulation, and Estimation* (Wiley Interscience, 1990)

[21] D. Bump, P. Diaconis and J.B. Keller, "Unitary correlations and the Fejér kernel", (preprint) (2000)

[22] J.B. Conrey, "The fourth moment of the derivative of the Riemann zeta function", *Quart. J. Math. Oxford (2)* **39** (1988) 21–36

[23] J.B. Conrey, "More than two-fiths of the zeros of the Riemann zeta-function are on the critical line", *J. Reine Angew. Math.* **399** (1989) 1–26

[24] J.B. Conrey, "A note on the fourth power moment of the Riemann zeta function" *Analytic Number Theory*, Progress in Mathematics **138** (1996) 225–230

[25] J.B. Conrey and D.W. Farmer, "Mean values of L–functions and symmetry", *Internat. Math. Res. Notices* **17** (2000) 883–908

[26] J.B. Conrey and A. Ghosh, "On mean values of the zeta-function", *Mathematika* **31** (1984) 159–161

[27] J.B. Conrey and A. Ghosh, "On mean values of the zeta-function, II", *Acta Arith.* **52** (1989) 367–371

[28] J.B. Conrey and A. Ghosh, "A conjecture for the sixth power moment of the Riemann zeta-function", *Internat. Math. Res. Notices* **15** (1998) 775–780

[29] J.B. Conrey, A. Ghosh and S.M. Gonek, "Large gaps between zeros of the zeta function", *Mathematika* **33** (1986) 212–238

[30] J.B. Conrey and S.M. Gonek, "High moments of the Riemann zeta function", *Duke Math. J.* **107** (2001) 577–604

[31] M. Coram and P. Diaconis. "New tests of the correspondence between unitary eigenvalues and the zeros of Riemann's zeta function.", preprint (2000)

[32] R.M. Corless *et. al.*, "On the Lambert W function", *Adv. Comput. Math.* **5** (1996) 329–359

[33] O. Costin and J.L. Lebowitz, "Gaussian fluctuation in random matrices", *Phys. Rev. Lett.* **75** (1995) 69–72

[34] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications, 2nd Ed.* (Springer-Verlag, 1998)

[35] P. Diaconis and S.N. Evans, "Linear functionals of eigenvalues of random matrices", *Trans. Amer. Math. Soc.* **353** (2001) 2615–2633

[36] P. Diaconis and M. Shahshahani, "On the eigenvalues of random matrices", *J. Appl. Probab.* **31A** (1994) 49–62

[37] F.J. Dyson, "Statistical theory of energy levels of complex systems, I, II and III", *J. Math. Phys.* **3** (1962) 140–156, 157–165, 166–175

[38] H.M. Edwards, *Riemann's Zeta Function* (Academic Press, 1974)

[39] L. Euler, "Variae observationes circa series infinitas", *Comm. Acd. Sci. Petropolitanae* **9** (1737) 222–236 (also in "Opera" (1), Vol 14, 216–244)

[40] P.J. Forrester and N.S. Witte, "Exact Wigner surmise type evaluation of the spacing distribution in the bulk of the scaled random matrix ensembles", *Lett. Math. Phys.* **53** (2000) 195–200

[41] A. Fujii, "On the distribution of the zeros of the Riemann zeta function in short intervals", *Proc. Japan Acad. A* **66** (1990) 75–79

[42] A. Fujii, "Explicit formulas and oscillations", *Emerging Applications of Number Theory*, D.A. Hejhal, J. Friedman, M.C. Gutzwiller, A.M. Odlyzko, eds. (Springer, 1999) 219–267

[43] D.M. Gangardt, "Second quantization approach to characteristic polynomials in RMT", *J. Phys. A* **34** (2001) 3553–3560

[44] D.A. Goldston, "On the function $S(T)$ in the theory of the Riemann zeta function", *J. Number Theory* **27** (1987) 149–177

[45] S.M. Gonek, "Mean values of the Riemann zeta function and its derivatives", *Invent. Math.* **75** (1984) 123–141

[46] S.M. Gonek, "On negative moments of the Riemann zeta-function", *Mathematika* **36** (1989) 71–88

[47] S.M. Gonek, "The second moment of the reciprocal of the Riemann zeta function and its derivative", lecture at the Mathematical Sciences Research Institute, Berkeley, (June 1999)

[48] F. Haake *et. al.*, "Secular determinants of random unitary matrices", *J. Phys. A* **29** (1996) 3641–3658

[49] J. Hadamard, "Sur la distribution des zeros de la fonction $\zeta(s)$ et ses consequences arithmétiques", *Bull. Soc. Math. Frances* **24** (1896) 199–220

[50] R.R. Hall, "A Wirtinger type inequality and the spacing of the zeros of the Riemann zeta function", preprint (2001)

[51] B.M. Hambly, P. Keevash, N. O'Connell and D. Stark, "The characteristic polynomial of a random permutation matrix", *Stochastic Process. Appl.* **90** (2000) 335–346

[52] G.H. Hardy and J.E. Littlewood, "Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes", *Acta Arith.* **41** (1918) 119–196

[53] D.R. Heath-Brown, "The fourth power moment of the Riemann zeta function", *Proc. Lond. Math. Soc. (3)* **38** (1979) 385–422

[54] D.R. Heath-Brown, "Fractional moments of the Riemann zeta function", *J. Lond. Math. Soc. (2)* **24** (1981) 65–78

[55] D.R. Heath-Brown and M.N. Huxley, "Exponential sums with a difference", *Proc. Lond. Math. Soc. (3)* **61** (1990) 227–250

[56] D.A. Hejhal, "On the distribution of $\log|\zeta'(1/2+\mathrm{i}t)|$", *Number Theory, Trace Formulas, and Discrete Groups*, K.E. Aubert, E. Bombieri and D.M. Goldfeld, eds., Proceedings of the 1987 Selberg Symposium, (Academic Press, 1989) 343–370

[57] D.A. Hejhal, "On the triple correlation of zeros of the zeta function", *Internat. Math. Res. Notices* **7** (1994) 293–302

[58] F. Hiai and D. Petz, "A large deviation theorem for the empirical eigenvalue distribution of random unitary matrices", *Ann. Inst. H. Poincaré Probab. Statist.* **36** (2000) 71–85

[59] C.P. Hughes, J.P. Keating and N. O'Connell, "Random matrix theory and the derivative of the Riemann zeta function", *Proc. R. Soc. Lond. A* **456** (2000) 2611–2627

[60] C.P. Hughes, J.P. Keating and N. O'Connell, "On the characteristic polynomial of a random unitary matrix", *Comm. Math. Phys.* **220** (2001) 429–451

[61] A.E. Ingham, "Mean-values theorems in the theory of the Riemann zeta-function", *Proc. Lond. Math. Soc.* **27** (1926) 273–300

[62] M. Jimbo, T. Miwa, Y. Môri and M. Sato, "Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent", *Physica* **1D** (1980) 80–158

[63] K. Johansson, "On Szegő's asymptotic formula for Toeplitz determinants and generalizations", *Bull. Sc. Math.* **112** (1988) 257–304

[64] K. Johansson, "On fluctuations of eigenvalues of random Hermitian matrices", *Duke J. Math.* **91** (1998) 151–204

[65] M. Jutila, "On the value distribution of the zeta-function on the critical line", *Bull. Lond. Math. Soc.* **15** (1983) 513-518

[66] N.M. Katz and P. Sarnak, *Random Matrices, Frobenius Eigenvalues, and Monodromy*, (AMS Colloquium Publications, 1999)

[67] J.P. Keating, "The Riemann zeta-function and quantum chaology", *Quantum Chaos*, G. Casati, I. Guarneri and U. Smilansky, eds. (North-Holland, 1993) 145–185

[68] J.P. Keating, "Periodic orbits, spectral statistics, and the Riemann zeros" *Supersymmetry and Trace Formulae: Chaos and Disorder*, I.V. Lerner, J.P. Keating and D.E. Khmelnitskii eds. (Kluwer Academic, 1999) 1–15

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

[70] J.P. Keating and N.C. Snaith, "Random matrix theory and L–functions at $s = 1/2$", *Comm. Math. Phys.* **214** (2000) 91–110

[71] A. Laurinčikas, *Limit Theorems for the Riemann Zeta Function* (Kluwer Academic Publishers, 1996)

[72] J. van der Lune, H.J.J. te Riele and D.T. Winter, "On the zeros of the Riemann zeta function in the critical strip IV", *Math. Comp.* **46** (1986) 667–681

[73] H. von Mangoldt, "Zur Verteilung der Nullstellen der Riemannschen Funktionen $\xi(t)$", *Math. Ann.* **60** (1905) 1–19

[74] M.L. Mehta, *Random Matrices* (Academic Press, 1991)

[75] J. Mueller, "On the difference between consecutive zeros of the Riemann zeta function", *J. Number Theory* **14** (1982) 327–331

[76] H.L. Montgomery, "The pair correlation of zeros of the zeta function", *Analytic Number Theory*, Proceedings of Symposia in Pure Mathematics **24** (1973) 181–193

[77] H.L. Montgomery, "Extreme values of the Riemann zeta function", *Comment. Math. Helv.* **52** (1977) 511–518

[78] A.M. Odlyzko, "The 10^{20}-th zero of the Riemann zeta function and 175 million of its neighbors", unpublished. Available from the author's website at `www.research.att.com/~amo/`

[79] L. Pastur, "Random matrices as paradigm", *Mathematical Physics 2000*, A. Fokas *et. al.* eds. (Imperial College Press, 2000) 216–265

[80] E. Rains, "High powers of random elements of compact Lie groups", *Probab. Theor. Rel. Fields* **107** (1997) 219–241

[81] G.F.B. Riemann "Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse" *Monatsberichte der Berliner Akademie* (1859) 671–680. An English translation of the paper is available in the appendix of [38] and also at `www.maths.tcd.ie/pub/HistMath/People/Riemann/Zeta/`

[82] Z. Rudnick and P. Sarnak "The n–level correlations of zeros of the zeta function", *C. R. Acad. Sci. Paris Sér. I Math.* **319** (1994) 1027–1032

[83] Z. Rudnick and P. Sarnak, "Zeros of principal L–functions and random matrix theory", *Duke Math. J.* **81** (1996) 269–322

[84] A. Selberg, "Bemerkninger om et multiplet integral", *Norsk Matematisk Tidsskrift* **26** (1944) 71–78

[85] A. Selberg, "Contributions to the theory of the Riemann zeta-function", *Arch. Math. Natur. B* **48**, No. 5, (1946) 89–155

[86] N.C. Snaith, *Random Matrix Theory and Zeta Functions*, PhD Thesis, (University of Bristol, 2000)

[87] A. Soshnikov, "Level spacings distribution for large random matrices: Gaussian fluctuations", *Ann. of Math. (2)* **148** (1998) 573–617

[88] A. Soshnikov, "Central limit theorem for local linear statistics in classical compact groups and related combinatorial identities", *Ann. Probab.* **28** (2000) 1353–1370

[89] G. Szegő, *Orthogonal Polynomials* (AMS Colloquium Publications XXII, 1939)

[90] E.C. Titchmarsh, *The Theory of the Riemann Zeta-Function* 2nd ed., revised by D.R. Heath-Brown (Oxford Science Publications, 1986)

[91] C.A. Tracy and H. Widom, "Introduction to random matrices", *Geometric and Quantum Aspects of Integrable Systems* (Ed. G. F. Helminck), Springer-Verlag Lecture Notes in Physics **424** (1993) 103–130

[92] K.-M. Tsang, "Some Ω-theorems for the Riemann zeta-function", *Acta Arith.* **46** (1986) 369–395

[93] C.-J. de la Vallée Poussin "Recherches analytiques sur la théorie des nombres (première partie)", *Ann. Soc. Sci. Bruxelles [I]* **20** (1896) 183–256

[94] A. Voros, "Spectral functions, special functions and the Selberg zeta function", *Comm. Math. Phys.* **110** (1987) 439–465

[95] H. Weyl, *Classical Groups*, (Princeton University Press, 1946)

[96] K. Wieand, *Eigenvalue Distributions of Random Matrices in the Permutation Group and Compact Lie Groups*, PhD Thesis, (Harvard University, 1998)

[97] K. Wieand, "Eigenvalue distributions of random unitary matrices", preprint (2000)