Commun. Math. Phys. 220, 429 – 451 (2001)

Communications in
Mathematical Physics

On the Characteristic Polynomial of a Random Unitary Matrix

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

[1] BRIMS, Hewlett-Packard Labs, Bristol, BS34 8QZ, UK
[2] School of Mathematics, University of Bristol, University Walk, Bristol, BS8 1TW, UK

Received: 27 June 2000 / Accepted: 30 January 2001

Abstract: We present a range of fluctuation and large deviations results for the logarithm of the characteristic polynomial Z of a random $N \times N$ unitary matrix, as $N \to \infty$. First we show that $\ln Z/\sqrt{\frac{1}{2}\ln N}$, evaluated at a finite set of distinct points, is asymptotically a collection of i.i.d. complex normal random variables. This leads to a refinement of a recent central limit theorem due to Keating and Snaith, and also explains the covariance structure of the eigenvalue counting function. Next we obtain a central limit theorem for $\ln Z$ in a Sobolev space of generalised functions on the unit circle. In this limiting regime, lower-order terms which reflect the global covariance structure are no longer negligible and feature in the covariance structure of the limiting Gaussian measure. Large deviations results for $\ln Z/A$, evaluated at a finite set of distinct points, can be obtained for $\sqrt{\ln N} \ll A \ll \ln N$. For higher-order scalings we obtain large deviations results for $\ln Z/A$ evaluated at a single point. There is a phase transition at $A = \ln N$ (which only applies to negative deviations of the real part) reflecting a switch from global to local conspiracy.

1. Introduction and Summary

Let U be an $N \times N$ unitary matrix, chosen uniformly at random from the unitary group $\mathcal{U}(N)$, and denote its eigenvalues by $\exp(i\theta_1), \ldots, \exp(i\theta_N)$. In order to develop a heuristic understanding of the value distribution and moments of the Riemann zeta function, Keating and Snaith [21] considered the characteristic polynomial (normalised so that its logarithm has zero mean)

$$Z(\theta) = \det(I - Ue^{-i\theta}) = \prod_{n=1}^{N} \left(1 - e^{i(\theta_n - \theta)}\right). \tag{1.1}$$

This is believed to be a good statistical model for the zeta function at (large but finite) height T up the critical line when the mean density of the non-trivial zeros (which equals

$(1/2\pi)\ln(T/2\pi))$ is set equal to the mean density of eigenangles (which is $N/2\pi$). (For additional evidence of this, concerning other statistics, see [9].)

Note that the law of $Z(\theta)$ is independent of $\theta \in \mathbb{T}$ (the unit circle). In [21] it is shown that as $N \to \infty$, $\ln Z(0)/\sigma$ converges in distribution to a standard complex normal random variable, where $2\sigma^2 = \ln N$. That is

$$\frac{\ln Z(0)}{\sqrt{\frac{1}{2}\ln N}} \Longrightarrow X + iY, \tag{1.2}$$

where X and Y are independent normal random variables with mean zero and variance one[1], and $\Longrightarrow$ denotes convergence in distribution. (A similar result can be found in [2], but there the real and imaginary parts of $\ln Z/\sigma$ are treated separately.) In order to make the imaginary part of the logarithm well-defined, the branch is chosen so that

$$\ln Z(\theta) = \sum_{n=1}^{N} \ln\left(1 - e^{i(\theta_n - \theta)}\right) \tag{1.3}$$

and

$$-\tfrac{1}{2}\pi < \mathfrak{Im}\ln\left(1 - e^{i(\theta_n-\theta)}\right) \le \tfrac{1}{2}\pi. \tag{1.4}$$

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, §2.11 of [24] or §4 of [22]) that, for rectangles $B \subseteq \mathbb{C}$,

$$\lim_{T\to\infty} \frac{1}{T}\left|\left\{T \le t \le 2T \;:\; \frac{\ln\zeta(\frac{1}{2}+it)}{\sqrt{\frac{1}{2}\ln\ln T}} \in B\right\}\right| = \frac{1}{2\pi}\iint_B e^{-(x^2+y^2)/2}\,\mathrm{d}x\,\mathrm{d}y. \tag{1.5}$$

Equating the mean density of the Riemann zeros at height T with the mean density of eigenangles of an $N \times N$ unitary matrix, we have $N = \ln(T/2\pi)$ and thus we see that these two central limit theorems are consistent.

In this paper we obtain more detailed fluctuation theorems for $\ln Z$ as $N \to \infty$, and a range of large and moderate deviations results.

First we show that $\ln Z/\sigma$, evaluated at a finite set of distinct points, is asymptotically a collection of i.i.d. complex normal random variables. This leads to a refinement of the above central limit theorem, and also explains the mysterious covariance structure which has been observed, by Costin and Lebowitz [10] and Wieand [32,33], in the eigenvalue counting function.

We also obtain a central limit theorem for $\ln Z$ in a Sobolev space of generalised functions on the unit circle. In this limiting regime, lower-order terms which reflect the global covariance structure are no longer negligible and feature in the covariance structure of the limiting Gaussian measure. The limiting process is not in $L_2(\mathbb{T})$. It is, however, when integrated, Hölder continuous with parameter $1-\delta$, for any $\delta > 0$.

Large deviations results for $\ln Z/A$, evaluated at a finite set of distinct points, are obtained for $\sqrt{\ln N} \ll A \ll \ln N$. For higher-order scalings we obtain large deviations results for $\ln Z/A$ evaluated at a single point. For the imaginary part, all scalings $A \ll N$

[1] Perhaps we should warn the reader at this point that some authors use the term "standard complex normal" to refer to the case where the variance of each component is $1/2$.

lead to quadratic rate functions. At $A = N$, the speed is N^2, and the rate function is a convex function for which we give an explicit formula. For the real part, only scalings up to $A = \ln N$ lead to quadratic rate functions. At this critical scaling one observes a phase transition, and beyond it deviations to the left and right occur at different speeds. For deviations to the left, the rate function becomes linear; for deviations to the right, the rate function remains quadratic up to but not including the scaling $A = N$. At the scaling $A = N$, deviations to the left occur at speed N, while deviations to the right occur at speed N^2, and the rate function is again a convex function for which we give an explicit formula. The phase transition reflects a switch from global to local conspiracy.

Related fluctuation theorems for random matrices can be found in [10, 13, 12, 19, 14, 27] and references therein. In particular, Diaconis and Evans [12] give an alternative proof of Theorem 2.2 below. The large deviation results at speed N^2 are partially consistent with (but do not follow from) a higher-level large deviation principle due to Hiai and Petz [16]. High-level large deviations results and concentration inequalities for other ensembles can be found in [5, 6, 15].

2. Fluctuation Results

Our first main result is that the law of $\ln Z(0)$ obtained by averaging over the unitary group is asymptotically the same as the value distribution of $\ln Z(\theta)$ obtained by averaging over θ for a typical realisation of U:

Theorem 2.1. *Set $W_N(\theta) = \ln Z(\theta)/\sigma$, and denote by m the uniform probability measure on $\mathbb{T}$ (so that $m(\mathrm{d}\theta) = \mathrm{d}\theta/2\pi$). As $N \to \infty$, the sequence of laws $m \circ W_N^{-1}$ converges weakly in probability to a standard complex normal variable.*

This will follow from Theorem 2.2 below, so we defer the proof.

Theorem 2.1 hints at the possibility that the t-range in (1.5) can be significantly reduced.

The characteristic polynomial can also be used to explain the mysterious "white noise" process which appears in recent work of Wieand [32, 33] on the counting function (and less explicitly in earlier work of Costin and Lebowitz [10]). A *Gaussian process* is defined to be a collection of real (complex) random variables $\{X(\alpha),\ \alpha \in I\}$, with the property that, for any $\alpha_1, \dots, \alpha_m$, the joint distribution of $X(\alpha_1), \dots, X(\alpha_m)$ is multivariate (complex) normal. For $-\pi < s < t \le \pi$, let $C_N(s,t)$ denote the number of eigenangles of U that lie in the interval (s,t). Wieand proves that the finite dimensional distributions of the process $\widetilde{C}_N$ defined by

$$\widetilde{C}_N(s,t) = \frac{C_N(s,t) - (t-s)N/2\pi}{\frac{1}{\pi}\sqrt{\frac{1}{2}\ln N}} \tag{2.1}$$

converge to those of a Gaussian process C which can be realised in the following way: let Y be a centered Gaussian process indexed by $\mathbb{T}$ with covariance function $\mathbb{E}Y(s)Y(t) = \mathbb{1}_{\{s=t\}}$ (where $\mathbb{1}$ is the indicator function) and set $C(s,t) = Y(t) - Y(s)$. What is the origin of this process Y? The answer is as follows. First, it is not hard to show that for each N,

$$\widetilde{C}_N(s,t) = Y_N(t) - Y_N(s), \tag{2.2}$$

where $Y_N(\theta) = \mathfrak{Im} \ln Z(\theta)/\sigma$. This follows from the identity

$$\mathbb{1}_{\{\theta\in(s,t)\}} = \frac{t-s}{2\pi} + \frac{1}{\pi}\mathfrak{Im}\ln(1-e^{i(\theta-t)}) - \frac{1}{\pi}\mathfrak{Im}\ln(1-e^{i(\theta-s)}), \tag{2.3}$$

where, as always, the principal branch of the logarithm is chosen as in (1.4).

Moreover:

Theorem 2.2. *Set $W_N(\theta) = \ln Z(\theta)/\sigma$. If $r_1, \dots, r_k \in \mathbb{T}$ are distinct, the joint law of $(W_N(r_1), \dots, W_N(r_k))$ converges as $N \to \infty$ to that of k i.i.d. standard complex normal random variables. In particular, the finite dimensional distributions of Y_N converge to those of Y.*

This suggests that the analogous extension of Selberg's theorem (1.5) might hold for the zeta function.

Proof. Let f be a real-valued function in $L_1(\mathbb{T})$, and denote by

$$\hat{f}_k = \int_0^{2\pi} f(\theta)e^{-ik\theta}m(\mathrm{d}\theta) \tag{2.4}$$

its Fourier coefficients. The N^{th} order Toeplitz determinant with symbol f is defined by

$$D_N[f] = \det(\hat{f}_{j-k})_{1\le j,k\le N}. \tag{2.5}$$

Heine's identity (see, for example, [28]) states that

$$D_N[f] = \mathbb{E}\prod_{n=1}^{N} f(\theta_n). \tag{2.6}$$

The following lemma is more general than we need here, but we record it for later reference.

Lemma 2.3. *For any $d(N) \gg 1$ as $N \to \infty$, $s, t \in \mathbb{R}^k$ with N sufficiently large such that $s_j > -d(N)$ for all j, and r_j distinct in $\mathbb{T}$,*

$$\mathbb{E}\exp\left(\sum_{j=1}^{k} s_j \mathfrak{Re}\ln Z(r_j)/d + t_j\mathfrak{Im}\ln Z(r_j)/d\right) \tag{2.7}$$

$$\sim \prod_{j=1}^{k} \mathbb{E}\exp\left(s_j \mathfrak{Re}\ln Z(r_j)/d + t_j\mathfrak{Im}\ln Z(r_j)/d\right) \tag{2.8}$$

$$\sim \exp\left(\sum_{j=1}^{k}\frac{\ln N}{4d^2}(s_j^2+t_j^2)\right). \tag{2.9}$$

Proof. This follows from Heine's identity and a result of Basor [4] on the asymptotic behaviour of Toeplitz determinants with Fisher–Hartwig symbols. The Fisher–Hartwig symbol we require has the form

$$f(\theta) = \prod_{j=1}^{k}(1-e^{i(\theta-r_j)})^{\alpha_j+\beta_j}(1-e^{i(r_j-\theta)})^{\alpha_j-\beta_j}. \tag{2.10}$$

Taking $\alpha_j = s_j/2d$ and $\beta_j = -it_j/2d$, we have, by Heine's identity,

$$\mathbb{E}\exp\left(\sum_{j=1}^{k} s_j \mathfrak{Re}\ln Z(r_j)/d + t_j \mathfrak{Im}\ln Z(r_j)/d\right) = D_N[f]. \tag{2.11}$$

Note that the α_j's are real and the β_j's purely imaginary. Basor [4] proves that, as $N \to \infty$, for r_j distinct,

$$D_N[f] \sim E(\alpha_1, \beta_1, r_1, \ldots, \alpha_k, \beta_k, r_k) \prod_{j=1}^{k} N^{\alpha_j^2 - \beta_j^2}, \tag{2.12}$$

for $\alpha_j > -1/2$, where

$$\begin{aligned} E(\alpha_1, \beta_1, r_1, \ldots, \alpha_k, \beta_k, r_k) = &\prod_{\substack{1\le m,n\le k \\ m\neq n}} \left(1 - e^{i(r_m - r_n)}\right)^{-(\alpha_m - \beta_m)(\alpha_n + \beta_n)} \\ &\times \prod_{j=1}^{k} \frac{G(1+\alpha_j+\beta_j)G(1+\alpha_j-\beta_j)}{G(1+2\alpha_j)}, \end{aligned} \tag{2.13}$$

where G is the Barnes G-function, and $\left|\arg\left(1 - e^{i(r_m - r_n)}\right)\right| \le \pi/2$. By closer inspection of the proof given in [4] it can be seen that (2.12) holds uniformly for $|\alpha_j| < 1/2 - \delta$, and $|\beta_j| < \gamma$, for any fixed $\delta, \gamma > 0$[2]. We remark that uniformity in β is worked out carefully in [32] for the case $\alpha_j = 0$ for each j, and uniformity in α is discussed in [4]. The statement of the lemma follows from noting that $E(0, 0, r_1, \ldots, 0, 0, r_k) = 1$. □

Setting $d = \sqrt{\frac{1}{2}\ln N} = \sigma$ completes the proof of Theorem 2.2. □

Proof of Theorem 2.1. Set $X_N(\theta) = \mathfrak{Re}\ln Z(\theta)/\sigma$, $Y_N(\theta) = \mathfrak{Im}\ln Z(\theta)/\sigma$ and

$$\phi_N(s,t) = \int_{\mathbb{T}} \exp\left(sX_N(\theta) + tY_N(\theta)\right) m(\mathrm{d}\theta). \tag{2.14}$$

By the central limit theorem derived in [21] (which we note, in passing, also follows from Theorem 2.2),

$$\mathbb{E}\phi_N(s,t) = \mathbb{E}\exp\left(sX_N(0) + tY_N(0)\right) \to e^{(s^2+t^2)/2}. \tag{2.15}$$

We also have

$$\mathbb{E}\phi_N(s,t)^2 = \int_{\mathbb{T}} \mathbb{E}\exp\left(sX_N(\theta) + tY_N(\theta) + sX_N(0) + tY_N(0)\right) m(\mathrm{d}\theta). \tag{2.16}$$

By Cauchy–Schwartz, the integrand is bounded above by

$$\sup_{N\ge N_0} \mathbb{E}\exp\left(2sX_N(0) + 2tY_N(0)\right), \tag{2.17}$$

---

[2] This was pointed out to us by Harold Widom.

where N_0 is chosen such that $2s > -\sigma(N_0)$. Thus, by Theorem 2.2 and the bounded convergence theorem,

$$\mathbb{E}\phi_N(s,t)^2 \to e^{s^2+t^2}, \tag{2.18}$$

and hence

$$\mathbb{P}(|\phi_N(s,t) - e^{(s^2+t^2)/2}| > \epsilon) \leq \operatorname{Var}\phi_N(s,t)/\epsilon^2 \to 0, \tag{2.19}$$

for any $\epsilon > 0$, by Chebyshev's inequality. Thus, for each s, t, the sequence $\phi_N(s,t)$ converges in probability to $e^{(s^2+t^2)/2}$. The result now follows from the fact that moment generating functions are convergence-determining. □

We note that Szegö's asymptotic formula for Toeplitz determinants does not apply in the above context. Szegö's theorem for real-valued functions states that if

$$A(h) = \sum_{k=1}^{\infty} k|\hat{h}_k|^2 < \infty, \tag{2.20}$$

then

$$D_N[e^h] = \exp\left(N\hat{h}_0 + A(h) + o(1)\right) \tag{2.21}$$

as $N \to \infty$. Combining this with Heine's identity, 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 $2A(h)$. Now, we can write $\mathfrak{Re}\ln Z(\theta) = \operatorname{Tr} h(U)$, where $h(t) = \mathfrak{Re}\ln(1 - e^{i(t-\theta)})$, but the Fourier coefficients $\hat{h}_k$ are of order $1/k$ in this case and $A(h) = +\infty$.

We can, however, apply Szegö's theorem to obtain a functional central limit theorem for $\ln Z$. Actually, we will use the following fact, due to Diaconis and Shahshahani [13], which can be deduced from Szegö's theorem.

Lemma 2.4. *For each l, the collection of random variables*

$$\left\{\sqrt{\frac{2}{j}}\operatorname{Tr} U^{-j},\ j = 1, \ldots, l\right\} \tag{2.22}$$

converges in distribution to a collection of i.i.d. standard complex normal random variables.

(In fact, it is shown in [13] that there is exact agreement of moments up to high order for each N. See also [18], where superexponential rates of convergence are established.)

Denote by H_0^a the space of generalised real-valued functions f on $\mathbb{T}$ with $\hat{f}_0 = 0$ and

$$\|f\|_a^2 = \sum_{k=-\infty}^{\infty} |k|^{2a}|\hat{f}_k|^2 = 2\sum_{k=1}^{\infty} k^{2a}|\hat{f}_k|^2 < \infty. \tag{2.23}$$

This is a Hilbert space with the inner product

$$\langle f, g\rangle_a = \sum_{k=-\infty}^{\infty} |k|^{2a}\hat{f}_k\hat{g}_k^*. \tag{2.24}$$

It is also a closed subspace of the Sobolev space H^a, which is defined similarly but without the restriction $\hat{f}_0 = 0$. Sobolev spaces have the following useful property: the unit ball in H^a is compact in H^b, whenever $a > b$. It follows that the unit ball in H_0^a is compact in H_0^b for $a > b$. We shall make use of this fact later. Note that, when $a = 0$, $\langle \cdot, \cdot \rangle_a$ is just the usual inner product on $L_2(\mathbb{T})$; in this case we will drop the subscript.

Fix $a < 0$, and define a Gaussian measure μ on $H_0^a \times H_0^a$ as follows. First, let $X_1, X_2, \dots$ be a sequence of i.i.d. standard complex normal random variables, $X_0 = 0$ and $X_{-k} = X_k^*$, and define a random element $F \in H_0^a$ by

$$F(\theta) = \sum_{k=-\infty}^{\infty} \left(\frac{X_k}{2\sqrt{2|k|}} \right) e^{ik\theta}. \tag{2.25}$$

(To see that $F \in H_0^a$, note that in fact $\mathbb{E}\|F\|_a^2 < \infty$ for $a < 0$.) Now define μ to be the law of (F, AF), where A is the Hilbert transform on $H_0^{-1/2}$:

$$\widehat{Af}_k = \begin{cases} i\hat{f}_k & k > 0 \\ -i\hat{f}_k & k < 0. \end{cases} \tag{2.26}$$

We will describe some properties of μ later. First we will prove:

Theorem 2.5. *The law of* $(\mathfrak{Re} \ln Z, \mathfrak{Im} \ln Z)$ *converges weakly to* μ.

Proof. First note that $\mathfrak{Im} \ln Z = A(\mathfrak{Re} \ln Z)$ and, for $k \neq 0$,

$$\left(\widehat{\mathfrak{Re} \ln Z} \right)_k = \frac{-\operatorname{Tr} U^{-k}}{2|k|}. \tag{2.27}$$

Convergence on cylinder sets (in the Fourier representation) therefore follows from Lemma 2.4.

To prove tightness in $H_0^a \times H_0^a$ we will use the fact that the unit ball in H_0^b is compact in H_0^a for $a < b < 0$, and the uniform bound

$$\mathbb{E}\|\mathfrak{Re} \ln Z\|_b^2 = 2\mathbb{E} \sum_{k=1}^{\infty} k^{2b} \left| \left(\widehat{\mathfrak{Re} \ln Z} \right)_k \right|^2 \tag{2.28}$$

$$= \frac{1}{2} \sum_{k=1}^{\infty} k^{2b-2} \mathbb{E} |\operatorname{Tr} U^{-k}|^2 \tag{2.29}$$

$$= \frac{1}{2} \sum_{k=1}^{\infty} k^{2b-2} \min(k, N) \tag{2.30}$$

$$\leq \frac{1}{2} \sum_{k=1}^{\infty} k^{2b-1}; \tag{2.31}$$

a similar bound holds for $\mathfrak{Im} \ln Z$. We have used the fact that $\mathbb{E}|\operatorname{Tr} U^k|^2 = \min(|k|, N)$ for $k \neq 0$ (see, for example, [25]). Thus,

$$\sup_N \mathbb{P}\left(\max\{\|\mathfrak{Re} \ln Z\|_b, \|\mathfrak{Im} \ln Z\|_b\} > q\right) \tag{2.32}$$

$$\leq \sup_N \{\mathbb{P}\left(\|\mathfrak{Re} \ln Z\|_b > q\right) + \mathbb{P}\left(\|\mathfrak{Im} \ln Z\|_b > q\right)\} \tag{2.33}$$

$$\leq \sup_N \left\{\mathbb{E}\|\mathfrak{Re} \ln Z\|_b^2 + \mathbb{E}\|\mathfrak{Im} \ln Z\|_b^2\right\} / q^2 \tag{2.34}$$

$$\to 0 \tag{2.35}$$

as $q \to \infty$, so we are done. □

We will now discuss some properties of the limiting measure μ. Let (F, AF) be a realisation of μ. First note that F and AF have the same law. Recalling the construction of F, we note that for $k > 0$ the random variables $|\hat{F}_k|^2$ are independent and $|\hat{F}_k|^2$ is exponentially distributed with mean $1/4k$. It follows that $\|F\|_a < \infty$ if, and only if, $a < 0$. In particular, F is almost surely not in $L_2(\mathbb{T})$.

Nevertheless, we can characterise the law of F by stating that, for $f \in H_0^{-1/2}$,

$$2\langle f, F\rangle / \|f\|_{-1/2} \tag{2.36}$$

is a standard normal random variable. The covariance is given by

$$\mathbb{E}\langle f, F\rangle\langle g, F\rangle = \frac{1}{4}\langle f, g\rangle_{-1/2}. \tag{2.37}$$

We note that

$$\langle f, g\rangle_{-1/2} = -2\int_{\mathbb{T}^2} \ln|e^{i\theta} - e^{i\phi}| f(\phi) g(\theta)\ m(\mathrm{d}\phi) m(\mathrm{d}\theta). \tag{2.38}$$

In the language of potential theory, if f is a charge distribution, then $\|f\|_{-1/2}^2$ is the *logarithmic energy* of f. The logarithmic energy functional also shows up as a large deviation rate function for the sequence of eigenvalue distributions: see Sect. 3.5 below.

We can also write down a stochastic integral representation for the process F. If we set

$$S(\phi) = \int^{\phi} F(\theta) \mathrm{d}\theta, \tag{2.39}$$

then S has the same law as

$$\widetilde{S}(\phi) = \frac{1}{2\pi}\int_0^{2\pi} b(\phi - \theta) \mathrm{d}B(\theta), \tag{2.40}$$

where B is a standard Brownian motion and

$$b(\theta) = \frac{1}{\sqrt{8\pi}} \sum_{k=1}^{\infty} k^{-3/2} \cos(k\theta). \tag{2.41}$$

To see this, compare covariances using the identity

$$\frac{1}{4}\left\| \mathbb{1}_{[0,t]} - \frac{t}{2\pi} \right\|_{-1/2}^2 = \frac{1}{4\pi^2} \sum_{k=1}^{\infty} \frac{1 - \cos(kt)}{k^3}. \tag{2.42}$$

Finally, we observe:

Lemma 2.6. *Let* $\delta > 0$. *The process* S *has a modification which is almost surely Hölder continuous with parameter* $1 - \delta$.

Proof. This follows from Kolmogorov's criterion (see, for example, [26, Theorem 2.1]) and the fact that

$$\mathbb{E}|S(t) - S(0)|^2 = \frac{1}{4} \left\| \mathbb{1}_{[0,t]} - \frac{t}{2\pi} \right\|_{-1/2}^2 \tag{2.43}$$

$$= \frac{1}{4\pi^2} \sum_{k=1}^{\infty} \frac{1 - \cos(kt)}{k^3} \tag{2.44}$$

$$\sim -\frac{1}{8\pi^2} t^2 \ln t, \tag{2.45}$$

as $t \to 0^+$. To see that this asymptotic formula is valid, one can use the fact that the expression (2.44) is related to Claussen's integral (see, for example, [1, §27.8]). □

We conclude this section with two remarks on Theorem 2.5. First, Rains [25] showed that, for each $\theta \neq 0$,

$$\operatorname{Var} C_N(0, \theta) = \frac{1}{\pi^2} (\ln N + \gamma + 1 + \ln |2 \sin(\theta/2)|) + o(1), \tag{2.46}$$

where $C_N(0, \theta)$ is the number of eigenangles lying in the interval $(0, \theta)$. Comparing this with (2.2) we see that

$$\mathbb{E}\Im\mathfrak{m} \ln Z(\theta) \Im\mathfrak{m} \ln Z(0) = -\frac{1}{2} \ln |2 \sin(\theta/2)| + o(1). \tag{2.47}$$

This is consistent with the fact that (formally)

$$\mathbb{E}F(\theta)F(0) = -\frac{1}{2} \ln |2 \sin(\theta/2)|. \tag{2.48}$$

The formal identity (2.48) in fact contains all of the information needed to determine the covariance structure of the process F. The fluctuation theorem (2.5) is therefore a statement which contains information about the global covariance structure of $\ln Z$. The covariance (2.47) is too small to feature in the scaling of Theorem 2.2.

Finally, the following observation arose in discussions with Marc Yor. The process F also appears in the following context. Let B be a standard complex Brownian motion, and $f : \mathbb{C} \to \mathbb{R}$ defined by $f(z) = h(\arg z)\delta(|z| = 1)$ for some $h : \mathbb{T} \to \mathbb{R}$ with $\hat{h}_0 = 0$. Then, as $t \to \infty$,

$$\frac{1}{\pi \sqrt{\ln t}} \int_0^t f(B_s) ds \Longrightarrow \langle h, F \rangle. \tag{2.49}$$

This can be deduced from a result of Kasahara and Kotani given in [20].

3. Large Deviations

In this section we present large and moderate deviations results for $\ln Z(0)$. We begin with a quick review of one-dimensional large deviation theory (see, for example, [8, 11]).

We are concerned with the log-asymptotics of the probability distribution of $R_N/A(N)$, where R_N is some one-dimensional real random variable and $A(N)$ is a scaling that is much greater than the square root of the variance of R_N (so we are outside the regime of the central limit theorem).

Suppose that there exists a function $B(N)$ (which tends to infinity as $N \to \infty$), such that

$$\Lambda(\lambda) := \lim_{N\to\infty} \frac{1}{B(N)} \ln \mathbb{E} \exp\left(\lambda \frac{B(N)}{A(N)} R_N\right) \tag{3.1}$$

exists as an extended real number, for each λ (i.e. the pointwise limit exists in the extended reals). The *effective domain* of $\Lambda(\cdot)$ is the set

$$\mathcal{D} = \{\lambda \in \mathbb{R} : \Lambda(\lambda) < \infty\} \tag{3.2}$$

and its interior is denoted by $\mathcal{D}^\circ$. The *convex dual* of $\Lambda(\cdot)$ is given by

$$\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}}\{\lambda x - \Lambda(\lambda)\}. \tag{3.3}$$

Theorem 3.1. *For $a < b$, if $\Lambda(\cdot)$ is differentiable in $\mathcal{D}^\circ$ and if*

$$(a, b) \subseteq \{\Lambda'(\lambda) : \lambda \in \mathcal{D}^\circ\}, \tag{3.4}$$

then

$$\lim_{N\to\infty} \frac{1}{B(N)} \ln \mathbb{P}\left\{\frac{R_N}{A(N)} \in (a, b)\right\} = - \inf_{x\in(a,b)} \Lambda^*(x). \tag{3.5}$$

If (3.5) holds we say that $R_N/A(N)$ satisfies the *large deviation principle* (LDP) with speed $B(N)$ and rate function Λ^*.

Some partial moderate deviations results can be obtained using Lemma 2.3; however, for many of the results presented here we will need more detailed information. In particular, we will make use of the following explicit formula (see, for example, [2, 7, 21]):

$$\begin{aligned}\mathbb{E}\exp&\left(s\mathfrak{Re}\ln Z(\theta) + t\mathfrak{Im}\ln Z(\theta)\right)\\ &= \frac{G(1+s/2+it/2)G(1+s/2-it/2)G(1+N)G(1+N+s)}{G(1+N+s/2+it/2)G(1+N+s/2-it/2)G(1+s)},\end{aligned} \tag{3.6}$$

valid for $\mathfrak{Re}(s \pm it) > -1$, where $G(\cdot)$ is the Barnes G-function, described in Appendix A. We will find the single moment generating functions useful, which we record here as

$$M_N(s) := \mathbb{E}\exp(s\mathfrak{Re}\ln Z(0)) \tag{3.7}$$

$$= \frac{G^2\left(1+\frac{1}{2}s\right)G(N+1)G(N+1+s)}{G(1+s)G^2\left(N+1+\frac{1}{2}s\right)}, \tag{3.8}$$

and

$$L_N(t) := \mathbb{E}\exp(it\mathfrak{Im}\ln Z(0)) \tag{3.9}$$

$$= \frac{G\left(1+\frac{1}{2}t\right)G\left(1-\frac{1}{2}t\right)G^2(N+1)}{G\left(N+1+\frac{1}{2}t\right)G\left(N+1-\frac{1}{2}t\right)}. \tag{3.10}$$

Theorem 3.2. *For any $A(N) \gg \ln N$, and $a < b < 0$,*

$$\lim_{N\to\infty}\frac{1}{A}\ln\mathbb{P}\left\{\frac{\mathfrak{Re}\ln Z(0)}{A} \in (a,b)\right\} = b. \tag{3.11}$$

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

$$\lim_{N\to\infty}\frac{1}{\ln N}\ln\mathbb{P}\left\{\frac{\mathfrak{Re}\ln Z(0)}{\ln N} \in (a,b)\right\} = b + 1/4. \tag{3.12}$$

Proof. From Theorem 3.9 we have that if $\limsup_{N\to\infty} x/\ln N < -1/2$, then

$$p(x) \sim e^x \exp\left(3\zeta'(-1) + \tfrac{1}{12}\ln 2 - \tfrac{1}{2}\ln\pi\right) N^{1/4}, \tag{3.13}$$

where $p(x)$ is the probability density function of $\mathfrak{Re}\ln Z(0)$.

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

$$\mathbb{P}\left\{\frac{\mathfrak{Re}\ln Z(0)}{\ln N} \in (a,b)\right\} = \int_{a\ln N}^{b\ln N} p(x)\,\mathrm{d}x$$

$$\sim \exp\left(3\zeta'(-1) + \tfrac{1}{12}\ln 2 - \tfrac{1}{2}\ln\pi\right) N^{1/4}\left(N^b - N^a\right) \tag{3.14}$$

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

3.1. Large deviations at the scaling $A = N$. Since $\mathfrak{Re}\ln Z(0) \leq N\ln 2$ and $|\mathfrak{Im}\ln Z(0)| \leq N\pi/2$, the scaling $A = N$ is the maximal non-trivial scaling.

Theorem 3.3. *The sequence $\mathfrak{Re}\ln Z(0)/N$ satisfies the LDP with speed N^2 and rate function given by the convex dual of*

$$\Lambda(s) = \begin{cases} \frac{1}{2}(1+s)^2\ln(1+s) - \left(1+\frac{1}{2}s\right)^2\ln\left(1+\frac{1}{2}s\right) - \frac{1}{4}s^2\ln 2s & \text{for } s \geq 0 \\ \infty & \text{for } s < 0. \end{cases} \tag{3.15}$$

Proof. $\ln\mathbb{E}\exp(sN\mathfrak{Re}\ln Z(0)) = \ln M_N(Ns)$, the asymptotics of which are given in Appendix C, and so

$$\Lambda(s) = \lim_{N\to\infty}\frac{1}{N^2}\ln M_N(Ns) \tag{3.16}$$

$$= \tfrac{1}{2}(1+s)^2\ln(1+s) - \left(1+\tfrac{1}{2}s\right)^2\ln\left(1+\tfrac{1}{2}s\right) - \tfrac{1}{4}s^2\ln 2s \tag{3.17}$$

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

If $x > 0$, then Theorem 3.1 implies that the rate function, $I(x)$, is given by the convex dual of $\Lambda(s)$. If $x < 0$, then Theorem 3.2 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 3.3. □

One can also obtain an LDP for the imaginary part:

Theorem 3.4. *The sequence* $\Im\mathfrak{m} \ln Z(0)/N$ *satisfies the LDP with speed* N^2 *and rate function given by the convex dual of*

$$\Lambda(t) = \tfrac{1}{8}t^2 \ln\left(1+\frac{4}{t^2}\right) - \tfrac{1}{2}\ln\left(1+\tfrac{1}{4}t^2\right) + t \arctan\left(\frac{1}{2}t\right). \tag{3.18}$$

Proof. $\ln \mathbb{E} \exp(tN\Im\mathfrak{m} \ln Z(0)) = \ln L_N(-iNt)$, and the asymptotics (given in Appendix D) imply that

$$\Lambda(t) = \lim_{N\to\infty} \frac{1}{N^2} \ln L_N(-iNt) \tag{3.19}$$

$$= \tfrac{1}{8}t^2 \ln\left(1+\frac{4}{t^2}\right) - \tfrac{1}{2}\ln\left(1+\tfrac{1}{4}t^2\right) + t \arctan\left(\frac{1}{2}t\right) \tag{3.20}$$

Theorem 3.1 implies that $J(y)$, the rate function, is given by the convex dual of $\Lambda(t)$, for all $y \in \mathbb{R}$. □

3.2. Moderate Deviations. At other scalings, one finds that the rate function is either quadratic or linear.

Theorem 3.5. *For scalings* $\sqrt{\ln N} \ll A \ll N$, *the sequence* $\Re\mathfrak{e} \ln Z(0)/A$ *satisfies the LDP with speed* $B = -A^2/W_{-1}(-A/N)$ *(where* W_{-1} *is Lambert's W-function, described in Appendix B) and rate function given by*

$$I(x) = \begin{cases} x^2 & \text{if } \sqrt{\ln N} \ll A \ll \ln N \\ \begin{cases} x^2 & x \geq -1/2 \\ -x - 1/4 & x < -1/2 \end{cases} & \text{if } A = \ln N \\ \begin{cases} x^2 & x \geq 0 \\ 0 & x < 0 \end{cases} & \text{if } \ln N \ll A \ll N. \end{cases} \tag{3.21}$$

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

$$\lim_{N\to\infty} \frac{1}{B} \ln M_N(sB/A) \tag{3.22}$$

exists as a non-trivial pointwise limit.

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

$$\frac{1}{B} \ln M_N(sN/\chi) = \begin{cases} \frac{1}{4}s^2 \frac{N^2 \ln \chi}{B\chi^2} + O_s\left(\frac{N^2}{\chi^2}\right) & \text{if } Ns/\chi > -1 \\ \infty & \text{if } Ns/\chi \leq -1 \end{cases} \tag{3.23}$$

which follows from results summarized in Appendix C.

Therefore a non-trivial limit of (3.22) occurs if $B = N^2 \ln \chi/\chi^2$, where $\chi = NA/B$, that is, if

$$B = \frac{A^2}{-W_{-1}\left(-\frac{A}{N}\right)}. \tag{3.24}$$

Note that the restriction $\chi \to \infty$ implies $A \ll N$, and that the restriction that $B \to \infty$ implies $A \gg \sqrt{\ln N}$.

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

$$\Lambda(s) = \lim_{N\to\infty} \frac{1}{B} \ln M_N(sB/A) \tag{3.25}$$

$$= \begin{cases} \frac{1}{4}s^2 & \text{for } s > -\delta \\ \infty & \text{for } s < -\delta. \end{cases} \tag{3.26}$$

If $\sqrt{\ln N} \ll A \ll \ln N$ then $\delta = +\infty$ and Theorem 3.1 implies that $I(x) = x^2$ for all $x \in \mathbb{R}$.

If $A = \ln N$, then $\delta = 1/2$, and Theorem 3.1 applies only for $x > -1/2$, where we have $I(x) = x^2$. However, since $B \sim \ln N$ at this scaling, Theorem 3.2 implies that, for $x < -1/2$, $I(x) = |x| - 1/4$.

Finally, if $\ln N \ll A \ll N$, then $\delta = 0$, and $I(x) = x^2$ for $x > 0$ by Theorem 3.1 and $I(x) = 0$ for $x < 0$ by Theorem 3.2 (since $B \gg A$ for $A \gg \ln N$).

This completes the proof of Theorem 3.5, □

Remark. For all $\sqrt{\ln N} \ll A \ll N$ it turns out that $I(x)$ is the convex dual of $\Lambda(s)$.

Once again, a similar result is true for the imaginary part, but this time the rate function is always quadratic.

Theorem 3.6. *For scalings $\sqrt{\ln N} \ll A \ll N$, the sequence $\mathfrak{Im} \ln Z(0)/A$ satisfies the LDP with speed $B = -A^2/W_{-1}(-A/N)$ and rate function $J(y) = y^2$.*

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

$$\lim_{N\to\infty} \frac{1}{B} \ln L_N(-itB/A) \tag{3.27}$$

exists as a non-trivial pointwise limit. Applying results from Appendix D we have

$$\ln L_N(-itB/A) = \tfrac{1}{4}t^2 N^2 \frac{\ln \chi}{\chi^2} + O_t\left(\frac{N^2}{\chi^2}\right) \tag{3.28}$$

for all $t \in \mathbb{R}$.

So, as in the proof of Theorem 3.5, we need B to be as in (3.24) (which will be valid for $\sqrt{\ln N} \ll A \ll N$), and the rate function will be given by the convex dual of $\frac{1}{4}t^2$, i.e. $J(y) = y^2$. □

3.3. *Large deviations of* $\ln Z(\theta)$ *evaluated at distinct points.*

Theorem 3.7. *For $\sqrt{\ln N} \ll A \ll \ln N$, and for any $r_1, \ldots, r_k$ (distinct), the sequence*

$$(\mathfrak{Re} \ln Z(r_1)/A, \mathfrak{Im} \ln Z(r_1)/A, \ldots, \mathfrak{Re} \ln Z(r_k)/A, \mathfrak{Im} \ln Z(r_k)/A) \tag{3.29}$$

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

$$I(x_1, y_1 \ldots, x_k, y_k) = \sum_{j=1}^{k} x_j^2 + y_j^2. \tag{3.30}$$

Proof. By Theorem 2.3, if $B/A \ll 1$,

$$\ln \mathbb{E} \exp\left(\sum_{j=1}^{k} s_j \mathfrak{Re} \ln Z(r_j) B/A + t_j \mathfrak{Im} \ln Z(r_j) B/A\right)$$
$$\sim \left(\sum_{j=1}^{N} (s_j^2 + t_j^2)/4\right) \frac{B^2 \ln N}{A^2}, \quad (3.31)$$

so choosing the speed $B = A^2/\ln N$, the stated result follows from a multidimensional analogue of Theorem 3.1 (see, for example, [11]). □

Remark. If B is given by (3.24), then for $\sqrt{\ln N} \ll A \ll \ln N$, $B \sim \frac{A^2}{\ln N}$. So for A in this restricted range, this theorem generalizes Theorems 3.5 and 3.6.

From this we can deduce large deviations results for the counting function, using the identity (2.2). For example:

Theorem 3.8. *For* $\sqrt{\ln N} \ll A \ll \ln N$, *and* $-\pi < s < t \leq \pi$, *the sequence* $(C_N(s,t) - (t-s)N/2\pi)/A$ *satisfies the LDP in* $\mathbb{R}$ *with speed* $B = A^2/\ln N$ *and rate function* $L(x) = \pi^2 x^2/2$.

3.4. Refined large deviations estimates. By Fourier inversion, the probability density of $\mathfrak{Re} \ln Z(0)$ is given by

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iyx} M_N(iy)\, \mathrm{d}y, \quad (3.32)$$

where $M_N(iy) = \mathbb{E} e^{iy \mathfrak{Re} \ln Z(0)}$ is given by (3.8).

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

$$p(x) \sim e^x \exp\left(3\zeta'(-1) + \tfrac{1}{12} \ln 2 - \tfrac{1}{2} \ln \pi\right) N^{1/4}. \quad (3.33)$$

Proof. We evaluate

$$\frac{1}{2\pi} \int_C e^{-iyx} M_N(iy)\, \mathrm{d}y, \quad (3.34)$$

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

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

$$p(x) = i \operatorname*{Res}_{y=i} \left\{ e^{-iyx} M_N(iy) \right\} + E, \quad (3.35)$$

where

$$E = \frac{e^{x+\epsilon x}}{2\pi} \int_{-\infty}^{\infty} e^{-itx} M_N(it - 1 - \epsilon)\, \mathrm{d}t. \quad (3.36)$$

It is not hard to show that

$$i \operatorname*{Res}_{y=i} \left\{ e^{-iyx} M_N(iy) \right\} \sim e^x \exp\left(3\zeta'(-1) + \tfrac{1}{12}\ln 2 - \tfrac{1}{2}\ln \pi\right) N^{1/4}, \tag{3.37}$$

and

$$|E| \le \frac{e^{x+\epsilon x}}{2\pi} \int_{-\infty}^{\infty} |M_N(it - 1 - \epsilon)| \mathrm{d}t \tag{3.38}$$

$$\sim \frac{e^{x+\epsilon x}}{\sqrt{\pi}} \left| \frac{G^2\left(\frac{1}{2} - \frac{1}{2}\epsilon\right)}{G(-\epsilon)} \right| N^{1/4+\epsilon/2+\epsilon^2/4} (\ln N)^{-1/2}. \tag{3.39}$$

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

$$e^{x\epsilon} N^{\epsilon/2+\epsilon^2/4} (\ln N)^{-1/2} \ll 1. \tag{3.40}$$

Thus the error term can be made subdominant if

$$\limsup_{N\to\infty} \frac{x}{\ln N} < -\frac{1}{2} \tag{3.41}$$

by choosing

$$0 < \epsilon < \min\left\{ -2 - 4 \limsup_{N\to\infty} \frac{x}{\ln N} \ ,\ 1 \right\}, \tag{3.42}$$

which completes the proof of the theorem. □

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

$$p(x) = \sum_{n=1}^{\infty} e^{(2n-1)x} \operatorname*{Res}_{s=0} \left\{ e^{-sx} M_N(s - (2n-1)) \right\}. \tag{3.43}$$

The problem with this evaluation of $p(x)$ is that it is hard to evaluate the residues of the non-simple poles in the sum, and when one does so the sum is asymptotic (in x) only for $x \ll -\ln N$.

Using Appendix C on the asymptotics of $M_N(t)$, the saddle point method gives

- For $|x| \ll \ln N$,

$$p(x) \sim \frac{1}{\sqrt{\pi(\ln N + 1 + \gamma)}} \exp\left(\frac{-x^2}{\ln N + 1 + \gamma} \right) \tag{3.44}$$

(This result was first found in [21] for $x = O(\sqrt{\ln N})$ – the central limit theorem.)

- For $\ln N \ll x \ll N^{1/3}$, writing W for $W_{-1}\left(-\frac{4x}{eN}\right)$,

$$p(x) \sim \frac{1}{\sqrt{\pi}} \exp\left(\frac{-x^2}{-W} + \frac{x^2}{2W^2} - \tfrac{5}{12}\ln(-W) - \tfrac{1}{12}\ln x + \tfrac{1}{12}\ln 2 + \zeta'(-1) \right). \tag{3.45}$$

The probability density of $\Im\mathfrak{m} \ln Z(0)$ is given by

$$q(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iyx} L_N(y)\, \mathrm{d}y \tag{3.46}$$

which we note is an even function.

Applying the results from Appendix D to calculate $L_N(s)$, the saddle point method gives

- For $|x| \ll \ln N$

$$q(x) \sim \frac{1}{\sqrt{\pi(\ln N + 1 + \gamma)}} \exp\left(\frac{-x^2}{\ln N + 1 + \gamma}\right). \tag{3.47}$$

- For $\ln N \ll |x| \ll \sqrt{N}$, writing W for $W_{-1}\left(\frac{-x}{Ne}\right)$,

$$q(x) \sim \frac{1}{\sqrt{\pi}} \exp\left(\frac{-x^2}{-W} + \frac{x^2}{W^2} - \tfrac{1}{3}\ln(-W) - \tfrac{1}{6}\ln x + 2\zeta'(-1)\right). \tag{3.48}$$

3.5. Inside the circle. The sequence of spectral measures

$$S_N = \frac{1}{N} \sum_{n=1}^{N} \delta_{\theta_n} \tag{3.49}$$

satisfies the LDP in $M_1(\mathbb{T})$ with speed N^2 and good convex rate function given by the logarithmic energy functional

$$\Sigma(\nu) = -\int_0^{2\pi}\int_0^{2\pi} \ln|e^{is} - e^{it}|\nu(\mathrm{d}s)\nu(\mathrm{d}t). \tag{3.50}$$

For a proof of this fact, see [16].

In this context, Varadhan's lemma (see, for example, [11]) can be stated as follows.

Theorem 3.10. *For any continuous $\phi : M_1(\mathbb{T}) \longrightarrow \mathbb{R}$ satisfying the condition*

$$\limsup_{N\to\infty} \frac{1}{N^2} \ln \mathbb{E} e^{\lambda N^2 \phi(S_N)} < \infty \tag{3.51}$$

for some $\lambda > 1$, then

$$\lim_{N\to\infty} \frac{1}{N^2} \ln \mathbb{E} e^{N^2 \phi(S_N)} = \sup_{\nu \in M_1(\mathbb{T})} \{\phi(\nu) - \Sigma(\nu)\}. \tag{3.52}$$

Now, we can write $\mathfrak{Re} \ln Z(0)/N = F_0(S_N)$, where

$$F_0(\nu) := \int_0^{2\pi} \mathfrak{Re} \ln\left(1 - e^{i\theta}\right) \nu(\mathrm{d}\theta); \tag{3.53}$$

however, F_0 is not weakly continuous, and Varadhan's lemma does not apply. Nevertheless, it is interesting to see if it gives the correct answer. That is, does the asymptotic cumulant generating function of Theorem 3.3 satisfy

$$\Lambda(s) = \sup_{\nu \in M_1(\mathbb{T})} \{s F_0(\nu) - \Sigma(\nu)\}? \tag{3.54}$$

If so, this variational formula would contain information about how large deviations for $\mathfrak{Re} \ln Z(0)/N$ actually occur. A similar variational problem can be written down for the imaginary part. Unfortunately, we are not able to even formally verify this except in very restricted and degenerate cases.

Consider first, for $\epsilon > 0$, the *continuous* function

$$F_\epsilon(\nu) := \int_0^{2\pi} \mathfrak{Re} \ln \left(1 - e^{-\epsilon} e^{i\theta}\right) \nu(\mathrm{d}\theta). \tag{3.55}$$

Then $\mathfrak{Re} \ln Z_\epsilon / N = F_\epsilon(S_N)$, where

$$Z_\epsilon = \prod_{n=1}^{N} \left(1 - e^{-\epsilon} e^{i\theta_n}\right). \tag{3.56}$$

Applying Varadhan's lemma, we obtain

$$\begin{aligned} \Lambda_\epsilon(s) &:= \lim_{N\to\infty} \frac{1}{N^2} \ln \mathbb{E} e^{Ns\mathfrak{Re} \ln Z_\epsilon)} \\ &= \sup_{\nu \in M_1(\mathbb{T})} \{s F_\epsilon(\nu) - \Sigma(\nu)\}. \end{aligned}$$

It is possible to solve this variational problem in the restricted range $-e^\epsilon - 1 \le s \le e^\epsilon - 1$, where we obtain:

$$\Lambda_\epsilon(s) = \tfrac{1}{4} s^2 \ln \left(\frac{1}{1 - e^{-2\epsilon}}\right). \tag{3.57}$$

Outside this range, it is much harder to solve.

Note that, letting $\epsilon \to 0$, we formally obtain $\Lambda(s) = \infty$ for $-2 \le s < 0$, which agrees (in this very restricted range) with the $\Lambda(s)$ of Theorem 3.3.

Similarly, for

$$G_\epsilon := \int_0^{2\pi} \mathfrak{Im} \ln \left(1 - e^{-\epsilon} e^{i\theta}\right) \nu(\mathrm{d}\theta), \tag{3.58}$$

we get that for $|t| \le e^\epsilon - e^{-\epsilon}$,

$$\lim_{N\to\infty} \frac{1}{N^2} \ln \mathbb{E} e^{Nt\mathfrak{Im} \ln Z_\epsilon} = \tfrac{1}{4} t^2 \ln \left(\frac{1}{1 - e^{-2\epsilon}}\right), \tag{3.59}$$

so letting $\epsilon \to 0$ all we could possibly obtain is $\Lambda(0) = 0$.

In both cases, the problem (of extending s and t beyond the ranges given) comes from finding the maximum over the set of all *non-negative* functions; only within the ranges given does the infimiser lie away from the boundary of this set.

Finally, we remark that

$$\mathfrak{Re}\ln\left(1-e^{-\epsilon}e^{i\theta_n}\right)=\sum_{\substack{k=-\infty\\k\neq 0}}^{\infty}\frac{-e^{-|k|\epsilon}}{2|k|}e^{ik\theta_n},\tag{3.60}$$

and

$$\mathfrak{Im}\ln\left(1-e^{-\epsilon}e^{i\theta_n}\right)=\sum_{\substack{k=-\infty\\k\neq 0}}^{\infty}\frac{ie^{-|k|\epsilon}}{2k}e^{ik\theta_n},\tag{3.61}$$

so Szegö's theorem implies that both $\mathfrak{Re}\ln Z_\epsilon$ and $\mathfrak{Im}\ln Z_\epsilon$ converges in distribution to normal random variables, with mean 0 and variance $-\frac{1}{2}\ln\left(1-e^{-2\epsilon}\right)$. Note the lack of $\sqrt{\ln N}$ normalization, as required in the case $\epsilon=0$.

3.6. The phase transition. The phase transition of Theorem 3.5 can be understood in terms of how deviations to the left for the real part actually occur, given that they do occur: here we present some heuristic arguments.

For $\sqrt{\ln N}\ll A\ll\ln N$ and $B=A^2/\ln N$, we have $(x>0)$

$$\frac{1}{B}\ln\mathbb{P}(\mathfrak{Re}\ln Z(0)<-xA)\sim -x^2.\tag{3.62}$$

On the other hand, if $A\gg\ln N$,

$$\frac{1}{A}\ln\mathbb{P}(\mathfrak{Re}\ln Z(0)<-xA)\sim -x.\tag{3.63}$$

Fix $\epsilon>0$ and consider the lower bound

$$\begin{aligned}&\mathbb{P}(\mathfrak{Re}\ln Z(0)<-xA)\\&\qquad\geq\mathbb{P}\left(\ln|1-e^{i\theta_1}|<-(x+\epsilon)A,\ \sum_{n=2}^{N}\ln|1-e^{i\theta_n}|<\epsilon A\right).\end{aligned}\tag{3.64}$$

Assuming the two events on the right hand side are approximately independent, and using the facts that θ_1 is uniformly distributed on $\mathbb{T}$ and

$$\mathbb{P}\left(\sum_{n=2}^{N}\ln|1-e^{i\theta_n}|<\epsilon A\right)\to 1,\tag{3.65}$$

this yields, for $A\gg\ln N$, the lower bound

$$\liminf_{N\to\infty}\frac{1}{A}\ln\mathbb{P}(\mathfrak{Re}\ln Z(0)<-xA)\geq\liminf_{N\to\infty}\frac{1}{A}\ln\mathbb{P}\left(\ln|1-e^{i\theta_1}|<-(x+\epsilon)A\right)\tag{3.66}$$

$$=-(x+\epsilon);\tag{3.67}$$

since ϵ is arbitrary, we obtain

$$\liminf_{N\to\infty} \frac{1}{A} \ln \mathbb{P}(\mathfrak{Re} \ln Z(0) < -xA) \geq -x. \tag{3.68}$$

On the other hand, if $\sqrt{\ln N} \ll A \ll \ln N$, the same estimate leads to

$$\liminf_{N\to\infty} \frac{1}{B} \ln \mathbb{P}(\mathfrak{Re} \ln Z(0) < -xA) \geq -\infty. \tag{3.69}$$

The fact that this simple estimate gives the right answer when $A \gg \ln N$, suggests that if the deviation

$$\{\mathfrak{Re} \ln Z(0) < -xA\} \tag{3.70}$$

occurs, it occurs simply because there is an eigenvalue too close to 1 (the other eigenvalues are “free to follow their average behaviour”). This is what we mean by a *local conspiracy*.

The fact that it leads to a gross underestimate when $\sqrt{\ln N} \ll A \ll \ln N$, suggests that in this case the deviation must involve the cooperation of many eigenvalues. A similar estimate based on only a (fixed) finite number of eigenvalues deviating from their mean behaviour leads to a similarly gross underestimate. Clearly it is more efficient in this case for many eigenvalues to arrange themselves and “share the load”, so to speak, than it is for one to bear it alone.

A. Barnes’ *G*-Function

Barnes’ G-function is defined [3] 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.

The G-function has the following properties [3,30]:

Recurrence relation: $G(z+1) = \Gamma(z)G(z)$.
Complex conjugation: $G^*(z) = G(z^*)$.
Asymptotic formula, valid for $|z| \to \infty$ with $|\arg(z)| < \pi$,

$$\ln G(z+1) \sim z^2\left(\tfrac{1}{2}\ln z - \tfrac{3}{4}\right) + \tfrac{1}{2} z \ln 2\pi - \tfrac{1}{12}\ln z + \zeta'(-1) + O\left(\frac{1}{z}\right). \tag{A.2}$$

Taylor expansion for $|z| < 1$,

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

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, $\Psi^{(n)}(z)$,

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

and

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

See, for example, [1] for properties of the gamma and polygamma functions.

B. Lambert's 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

$$\ln x = vx^{\beta} \tag{B.2}$$

has solution

$$x = \exp\left(\frac{-W(-\beta v)}{\beta}\right). \tag{B.3}$$

There are various asymptotic expansions of the W function:

- As $x \to \infty$,

$$W_0(x) \sim \ln x - \ln\ln x + \frac{\ln\ln x}{\ln x}. \tag{B.4}$$

- As $x \to 0$ on the principal branch,

$$W_0(x) \sim x - x^2 + \tfrac{3}{2}x^3. \tag{B.5}$$

- As $x \to 0^-$ on the -1 branch,

$$W_{-1}(x) \sim \ln|x| - \ln|\ln|x|| + \frac{\ln|\ln|x||}{\ln|x|}. \tag{B.6}$$

C. Asymptotics of $\ln M_N(x)$

From the asymptotics for the G-function, (A.2), we have for $x > -1$,

$$\begin{aligned}\ln M_N(x) = 2\ln G\left(1+\tfrac{1}{2}x\right) - \ln G(1+x) - \tfrac{3}{8}x^2 + \tfrac{1}{2}N^2\ln N \\ + \tfrac{1}{2}(N+x)^2\ln(N+x) - \left(N+\tfrac{1}{2}x\right)^2\ln\left(N+\tfrac{1}{2}x\right) \\ + \tfrac{1}{6}\ln\left(N+\tfrac{1}{2}x\right) - \tfrac{1}{12}\ln(N+x) - \tfrac{1}{12}\ln N + O\left(\frac{1}{N}\right),\end{aligned} \tag{C.1}$$

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

$$\ln M_N(x) = \tfrac{1}{4}x^2(\ln N + 1 + \gamma) + O\left(x^3\right) + O\left(\frac{1}{N}\right). \tag{C.2}$$

- If $x = O(1)$ and $x > -1$, then

$$\ln M_N(x) = \tfrac{1}{4}x^2\ln N + 2\ln G\left(1+\tfrac{1}{2}x\right) - \ln G(1+x) + O\left(\frac{1}{N}\right). \tag{C.3}$$

- If $1 \ll x \ll \sqrt[3]{N}$, then

$$\begin{aligned}\ln M_N(x) = \tfrac{1}{4}x^2\left(\ln N - \ln x - \ln 2 + \tfrac{3}{2}\right) + \tfrac{1}{6}\ln 2 - \tfrac{1}{12}\ln x + \zeta'(-1) \\ + O\left(\frac{x^3}{N}\right) + O\left(\frac{1}{x}\right).\end{aligned} \tag{C.4}$$

- If $x = \lambda N$ with $\lambda = O(1)$ and $\lambda > 0$, then

$$\begin{aligned}\ln M_N(x) = N^2\left\{\tfrac{1}{2}(1+\lambda)^2\ln(1+\lambda) - \left(1+\tfrac{1}{2}\lambda\right)^2\ln\left(1+\tfrac{1}{2}\lambda\right)\right. \\ \left.-\tfrac{1}{4}\lambda^2\ln(2\lambda)\right\} - \tfrac{1}{12}\ln N - \tfrac{1}{12}\ln\lambda + \zeta'(-1) \\ + \tfrac{1}{6}\ln(2+\lambda) - \tfrac{1}{12}\ln(1+\lambda) + O\left(\frac{1}{N}\right).\end{aligned} \tag{C.5}$$

D. Asymptotics of $\ln L_N(ix)$

We consider $x \in \mathbb{R}$. From the asymptotics for the G-function, (A.2), we have

$$\begin{aligned}\ln L_N(ix) = \ln G\left(1+\tfrac{1}{2}ix\right) + \ln G\left(1-\tfrac{1}{2}ix\right) - \tfrac{3}{8}x^2 + N^2\ln N \\ - \tfrac{1}{2}\left(N+\tfrac{1}{2}ix\right)^2\ln\left(N+\tfrac{1}{2}ix\right) - \tfrac{1}{2}\left(N-\tfrac{1}{2}ix\right)^2\ln\left(N-\tfrac{1}{2}ix\right) \\ - \tfrac{1}{6}\ln N + \tfrac{1}{12}\ln\left(N+\tfrac{1}{2}ix\right) + \tfrac{1}{12}\ln\left(N-\tfrac{1}{2}ix\right) + O\left(\frac{1}{N}\right).\end{aligned} \tag{D.1}$$

Constraining $x(N)$ to lie in various regimes simplifies the above considerably:

- If $|x| \ll 1$, then

$$\ln L_N(ix) = \tfrac{1}{4}x^2(\ln N + 1 + \gamma) + O(x^4) + O\left(\frac{1}{N}\right). \tag{D.2}$$

- If $x = O(1)$, then

$$\ln L_N(ix) = \ln G\left(1 + \tfrac{1}{2}ix\right) + \ln G\left(1 - \tfrac{1}{2}ix\right) + \tfrac{1}{4}x^2 \ln N + O\left(\frac{1}{N}\right). \tag{D.3}$$

- If $1 \ll |x| \ll \sqrt{N}$, then

$$\begin{aligned} \ln L_N(ix) = {} & \tfrac{1}{4}x^2\left(\ln N - \ln x + \ln 2 + \tfrac{3}{2}\right) - \tfrac{1}{6}\ln x + \tfrac{1}{6}\ln 2 + 2\zeta'(-1) \\ & + O\left(\frac{x^4}{N^2}\right) + O\left(\frac{1}{x^2}\right). \end{aligned} \tag{D.4}$$

- If $x = \lambda N$ with $\lambda = O(1)$, then

$$\begin{aligned} \ln L_N(ix) = {} & N^2\left\{\tfrac{1}{8}\lambda^2 \ln\left(1 + 4\lambda^{-2}\right) - \tfrac{1}{2}\ln\left(1 + \tfrac{1}{4}\lambda^2\right) + \lambda \tan^{-1}\tfrac{1}{2}\lambda\right\} \\ & - \tfrac{1}{6}\ln N + \tfrac{1}{12}\ln\left(1 + 4\lambda^{-2}\right) + 2\zeta'(-1) + O\left(\frac{1}{N}\right). \end{aligned} \tag{D.5}$$

Acknowledgements. We are grateful to Persi Diaconis and Steve Evans for their suggestions and for making the preprint [12] available to us. Thanks also to Harold Widom for helpful correspondence and Marc Yor for fascinating discussions on the connection between Theorem 2.5 and planar Brownian motion.

References

1. Abramowitz, M. and Stegun, I.A.: *Handbook of Mathematical Functions*. Dover, 1965
2. Baker, T.H. and Forrester, P.J.: Finite-N fluctuations formulas for random matrices. J. Stat. Phys. **88**, nos. 5/6, 1371–1386 (1997)
3. Barnes, E.W.: The theory of the G-function. Quart. J. Pure Appl. Math. **31**, 264–314 (1899)
4. Basor, E.: Asymptotic formulas for Toeplitz determinants. Trans. Am. Math. Soc. **239**, 33–65 (1978)
5. Ben Arous, G. and Guionnet, A.: Large deviations for Wigner's law and Voiculescu's non-commutative entropy. Probab. Theor. Rel. Fields **108**, no. 4, 517–542 (1997)
6. Ben Arous, G. and Zeitouni, O.: Large deviations from the circular law. ESAIM: Probability and Statistics **2**, 123–134 (1998)
7. Böttcher, A. and Silbermann, B.: *Introduction to Large Truncated Toeplitz Matrices*. Berlin–Heidelberg–New York: Springer-Verlag, 1999
8. Bucklew, J.: *Large Deviation Techniques in Decision, Simulation, and Estimation*. New York: Wiley Interscience, 1990
9. Coram, M. and Diaconis, P.: New tests of the correspondence between unitary eigenvalues and the zeros of Riemann's zeta function. Preprint (2000)
10. Costin, O. and Lebowitz, J.L.: Gaussian fluctuation in random matrices. Phys. Rev. Lett. **75**, 69–72 (1995)
11. Dembo, A. and Zeitouni, O.: *Large Deviations Techniques and Applications,* 2nd Ed. Berlin–Heidelberg–New York: Springer-Verlag, 1998
12. Diaconis, P. and Evans, S.N.: Linear functionals of eigenvalues of random matrices. Preprint (2000)
13. Diaconis, P. and Shahshahani, M.: On the eigenvalues of random matrices. J. Appl. Probab. A **31**, 49–62 (1994)
14. Guiounnet, A.: Fluctuations for strongly interacting random variables and Wigner's law. Preprint (1999)
15. Guiounnet, A. and Zeitouni, O.: Concentration of the spectral measure for large matrices. Preprint (2000)

16. Hiai, F. and Petz, D.: A large deviation theorem for the empirical eigenvalue distribution of random unitary matrices. Preprint (2000)
17. Hida, T.: *Brownian Motion*. Berlin–Heidelberg–New York: Springer, 1980
18. Johansson, K.: On random matrices from the classical compact groups. Ann. Math. **145**, 519–545 (1997)
19. Johansson, K.: On fluctuations of eigenvalues of random Hermitian matrices. Duke J. Math. **91**, 151–204 (1998)
20. Kasahara, Y. and Kotani, S.: On limit processes for a class of additive functionals of recurrent diffusion processes. Z. Wahrsch. verw. Gebiete. **49**, 133–153 (1979)
21. Keating, J.P. and Snaith, N.C.: Random matrix theory and $\zeta(1/2+it)$. Commun. Math. Phys. **214**, 57–89 (2000)
22. Laurinčikas, A.: *Limit Theorems for the Riemann Zeta Function*. Dordrecht: Kluwer Academic Publishers, 1996
23. Mehta, M.L.: *Random Matrices*. New York: Academic Press, 1991
24. Odlyzko, A.M.: The 10^{20}-th zero of the Riemann zeta function and 175 million of its neighbors. Unpublished
25. Rains, E.: High powers of random elements of compact Lie groups. Prob. Th. Rel. Fields **107**, 219–241 (1997)
26. Revuz, D. and Yor, M.: *Continuous Martingales and Brownian Motion*. Berlin–Heidelberg–New York: Springer-Verlag, 1990
27. Shoshnikov, A.B.: Gaussian fluctuation for the number of particles in Airy, Bessel, sine, and other determinantal random point fields. Preprint (1999)
28. Szegö, G.: *Orthogonal Polynomials*. AMS Colloquium Publications XXII, 1939
29. Titchmarsh, E.C.: *The Theory of the Riemann Zeta-Function*. Oxford: Oxford Science Publications, 1986
30. Voros, A.: Spectral functions, special functions and the Selberg zeta function. Commun. Math. Phys. **110**, 439–465 (1987)
31. Weyl, H.: *Classical Groups*. Princeton, NJ: Princeton University Press, 1946
32. Wieand, K.: *Eigenvalue Distributions of Random Matrices in the Permutation Group and Compact Lie Groups*. PhD Thesis, Harvard University, 1998
33. Wieand, K.: Eigenvalue distributions of random unitary matrices. Preprint (2000)

Communicated by P. Sarnak