

Finite- N Fluctuation Formulas for Random Matrices

T. H. Baker¹ and P. J. Forrester¹

Received January 7, 1997

For the Gaussian and Laguerre random matrix ensembles, the probability density function (p.d.f.) for the linear statistic $\sum_{j=1}^N (x_j - \langle x \rangle)$ is computed exactly and shown to satisfy a central limit theorem as $N \rightarrow \infty$. For the circular random matrix ensemble the p.d.f.'s for the statistics $\frac{1}{2} \sum_{j=1}^N (\theta_j - \pi)$ and $-\sum_{j=1}^N \log 2 |\sin \theta_j/2|$ are calculated exactly by using a constant term identity from the theory of the Selberg integral, and are also shown to satisfy a central limit theorem as $N \rightarrow \infty$.

KEY WORDS: Random matrices; central limit theorem; fluctuation formulas; Toeplitz determinants; Selberg integral.

1. INTRODUCTION

A number of rigorous results have recently been established regarding probability density functions and associated fluctuation formulas for linear statistics in random matrix ensembles. We recall that with $\lambda_1, \dots, \lambda_N$ denoting the eigenvalues of a random matrix, a linear statistic is any stochastic function A which can be written in the form $A = \sum_{j=1}^N a(\lambda_j)$. Corresponding to the statistic A is the probability function (p.d.f.)

$$P(u) := \prod_{j=1}^N \int_{-\infty}^{\infty} d\lambda_1 \dots \int_{-\infty}^{\infty} d\lambda_N \delta\left(u - \sum_{j=1}^N a(\lambda_j)\right) W(\lambda_1, \dots, \lambda_N) \quad (1.1)$$

where W denotes the p.d.f. for the eigenvalue distribution, for the event that $u = \sum_{j=1}^N a(\lambda_j)$. An evaluation of the variance of $P(u)$ is referred to as a

¹ Department of Mathematics, University of Melbourne, Parkville, Victoria 3052, Australia; E-mail: tbaker@maths.mu.oz.au, matpjf@maths.mu.oz.au.

fluctuation formula, although this term is sometimes used to refer to an evaluation of $P(u)$ itself.

When $a(x) = \chi_{[-l/2, l/2]} - \langle \chi_{[-l/2, l/2]} \rangle$, where $\langle \rangle$ denotes the mean and $\chi_B = 1$ if $x \in B$, $\chi_B = 0$ otherwise, we have that the linear statistic A represents the deviation in the number of eigenvalues from the mean number in the interval $[-l/2, l/2]$. For $N \times N$ Gaussian random matrices, scaled so that the mean eigenvalue spacing in the bulk of the spectrum is unity, it has been proved by Costin and Lebowitz⁽¹⁾ that

$$\left(\frac{2}{\pi^2\beta} \log l\right)^{1/2} P\left(\left(\frac{2}{\pi^2\beta} \log l\right)^{1/2} u\right) \sim \frac{1}{(2\pi)^{1/2}} e^{-u^2/2} \quad \text{as } l \rightarrow \infty \quad (1.2)$$

where $\beta = 1, 2$ and 4 for random symmetric, Hermitian and self-dual real quaternion matrices respectively. A related rigorous result has been recently established by Soshnikov,⁽²⁾ in which a central limit theorem for the fluctuations in the nearest neighbour spacing for the Dyson circular ensembles of unitary random matrices is derived (although the nearest neighbour spacing is not a linear statistic). Before scaling, the eigenvalue p.d.f. for Gaussian random matrices is proportional to

$$\prod_{l=1}^N e^{-\beta\lambda_l^2/2} \prod_{1 \leq j < k \leq N} |\lambda_k - \lambda_j|^\beta \quad (1.3)$$

This can be interpreted as the Boltzmann factor for a log-gas system in equilibrium at inverse temperature β . From this interpretation, heuristic arguments based on macroscopic electrostatics have been devised^(3,4) which support the validity of (1.2) for all $\beta > 0$.

Another, more general class of rigorous results for the evaluation of (1.1) for random matrix ensembles has been obtained by Johansson^(5,6) (see also ref. 7 in the case $\beta = 2$). In these results, instead of scaling the eigenvalues so that the mean spacing is unity, the scale is chosen so that as $N \rightarrow \infty$ the support of the density of the eigenvalues is the finite interval $(-1, 1)$. Since for large N the density of eigenvalues $\rho(x)$ implied by (1.3) is given by the Wigner semi-circle law (see e.g., ref. 8)

$$\rho(x) \sim \frac{(2N)^{1/2}}{\pi} \left(1 - \frac{x^2}{2N}\right)^{1/2} \quad (1.4)$$

this is achieved by the scaling $\lambda_j \mapsto (2N)^{1/2} \lambda_j$ in (1.3). More generally, with the exponent λ_j^2 in (1.3) replaced by an even-degree polynomial which is positive for large λ_j , a scale can always be chosen so that the support is the finite interval $(-1, 1)$.⁽⁵⁾ Moreover, independent of the details of the

polynomial, Johansson proved that the p.d.f. (1.1) in the scaled $N \rightarrow \infty$ limit tends to a Gaussian:

$$\lim_{\substack{N \rightarrow \infty \\ \text{support } \rho(x) = (-1, 1)}} P(u) = \left(\frac{1}{2\pi\sigma^2} \right)^{1/2} e^{-u^2/2\sigma^2} \tag{1.5a}$$

where

$$\sigma^2 = \frac{1}{\beta\pi^2} \int_{-1}^1 dx \frac{a(x)}{(1-x^2)^{1/2}} \int_{-1}^1 dy \frac{a'(y)(1-y^2)^{1/2}}{x-y} \tag{1.5b}$$

provided σ^2 is finite. The value of the variance was predicted in earlier work due to Brézin and Zee⁽⁹⁾ in the case $\beta = 2$, and Beenakker⁽¹⁰⁾ for general $\beta > 0$.

Both (1.2) and (1.5) can be interpreted as central limit theorems. However, unlike the classical result the standard deviation is no longer proportional to \sqrt{N} (or \sqrt{l} in the case of (1.2)). Rather the fluctuations are strongly suppressed, being independent of N altogether in the case of (1.5).

It is the purpose of this work to extend the rigorous results relating to the p.d.f. (1.1) for random matrix ensembles. Attention will be focussed on two types of linear statistics: $a(x) = x$ and $a(x) = -\log|x|$, which both arise naturally within the log-gas interpretation of (1.3). Indeed $a(x) = x$ corresponds to the dipole moment while $a(x) = -\log|x|$ corresponds to the potential at the origin. This latter statistic has been considered for the two-dimensional one-component plasma (i.e., the two-dimensional generalization of the log-gas) by Alastuey and Jancovici.⁽¹¹⁾

In Section 2 we will show how (1.1) with W proportional to (1.3) can be computed exactly for $a(x) = x$. Also, we calculate (1.1) exactly for $a(x) = x - \langle x \rangle$ with W proportional to

$$\prod_{l=1}^N x_l^{\beta a_l/2} e^{-\beta x_l/2} \prod_{1 \leq j < k \leq N} |x_k - x_j|^\beta \tag{1.6}$$

$x_j \geq 0$ ($j = 1, \dots, N$), corresponding to the Laguerre random matrix ensemble. In the scaled $N \rightarrow \infty$ limit, the result (1.5) with $a(x) = x$ is reclaimed in both cases.

In Section 3 we consider (1.1) with W proportional to

$$\prod_{1 \leq j < k \leq N} (2|\sin(\theta_k - \theta_j)/2|)^\beta \tag{1.7}$$

$0 \leq \theta_j \leq 2\pi$ ($j = 1, \dots, N$), corresponding to Dyson's circular ensemble of unitary random matrices. In the case $\beta = 2$ the large- N behaviour of $P(u)$

for $a(\theta) = (\theta - \pi)/2$ and $a(\theta) = -\log 2 |\sin \theta/2|$ is computed using the so-called Fisher–Hartwig conjecture⁽¹²⁾ (which is now a theorem^(13, 14)) from the theory of Toeplitz determinants. For general β , $P(u)$ is evaluated for these statistics by using a constant term identity due to Morris⁽¹⁵⁾ which is equivalent to the well known Selberg integral.⁽¹⁶⁾ We conclude in Section 4 with an interpretation of the variance of the dipole moment statistic calculated in Section 2 as a susceptibility in macroscopic electrostatics.

2. DIPOLE MOMENT STATISTIC IN THE GAUSSIAN AND LAGUERRE ENSEMBLE

Rather than study the p.d.f. (1.1) directly, we consider instead its Fourier transform

$$\begin{aligned} \tilde{P}(k) &:= \int_{-\infty}^{\infty} dx e^{ikx} P(x) \\ &= \int_{-\infty}^{\infty} dx_1 e^{ika(x_1)} \dots \int_{-\infty}^{\infty} dx_N e^{ika(x_N)} W(x_1, \dots, x_N) \end{aligned} \tag{2.1}$$

Note that (2.1) can be interpreted as the canonical average of the Boltzmann factor for a one-body external potential $ika(x)/\beta$.

2.1. Gaussian Random Matrices

There are three distinct random matrix ensembles, in which the matrices X are real symmetric ($\beta = 1$), Hermitian ($\beta = 2$) and self-dual quaternion real ($\beta = 4$), and the joint distribution for their elements is proportional to $e^{-\beta \text{Tr}(X^2)/2}$. The corresponding eigenvalue p.d.f. is proportional to (1.3). Scaling the eigenvalues $\lambda_j \mapsto \sqrt{2N} \lambda_j$ so that the support of the density is $(-1, 1)$, we have from (1.3) and (2.1) that for the linear statistic corresponding to $a(x) = x$

$$\begin{aligned} \tilde{P}(k) &= \frac{1}{C} \int_{-\infty}^{\infty} d\lambda_1 e^{-N\beta\lambda_1^2 + ik\lambda_1} \dots \int_{-\infty}^{\infty} d\lambda_N e^{-N\beta\lambda_N^2 + ik\lambda_N} \prod_{1 \leq j < k \leq N} |\lambda_k - \lambda_j|^\beta \\ &= \frac{e^{-k^2/4\beta}}{C} \int_{-\infty}^{\infty} d\lambda_1 e^{-N\beta(\lambda_1 - ik/2\beta N)^2} \\ &\quad \dots \int_{-\infty}^{\infty} d\lambda_N e^{-N\beta(\lambda_N - ik/2\beta N)^2} \prod_{1 \leq j < k \leq N} |\lambda_k - \lambda_j|^\beta \end{aligned} \tag{2.2}$$

where the normalization C is such that $\tilde{P}(0) = 1$.

Now for β even the integrand in (2.2) is analytic and decays sufficiently fast at infinity that we can shift the contours of integration from the real line to $\lambda_j = ik/2\beta N + t_j$ ($j = 1, \dots, N$), which gives that the integral is independent of k . Since the integral in (2.2) divided by C is a bounded analytic function of β for $\text{Re}(\beta) > 0$, it follows from Carlson's theorem⁽¹⁷⁾ that the integral is independent of k for all $\text{Re}(\beta) > 0$. Thus for all $\beta > 0$ and each $N = 1, 2, \dots$ we have the simple result that

$$\tilde{P}(k) = e^{-k^2/4\beta} \tag{2.3a}$$

and so

$$P(u) = \left(\frac{1}{\pi\beta}\right)^{1/2} e^{-\beta u^2} \tag{2.3b}$$

Since (2.3b) is independent of N it trivially remains valid in the $N \rightarrow \infty$ limit, and agrees with (1.5) provided $\sigma^2 = 1/2\beta$.

The value of σ^2 is easily computed from (1.5b) by recalling (see, e.g., ref. 18) that the solution of the integral equation

$$x = \int_{-1}^1 dy \frac{\phi(y)}{x-y} \tag{2.4}$$

is $\phi(y) = (1/\pi) \sqrt{1-y^2}$. Thus with $a(x) = x$, substituting (2.4) in (1.5b) we have

$$\sigma^2 = \frac{1}{\beta\pi} \int_{-1}^1 \frac{x^2}{\sqrt{1-x^2}} dx = \frac{1}{2\beta} \tag{2.5}$$

as required.

The formula (2.3) can also be derived for the Gaussian random matrices ($\beta = 1, 2$ or 4) from the fact that the joint p.d.f. for the elements is proportional to

$$\prod_{j=1}^N e^{-N\beta x_{jj}^2} \prod_{j < k} e^{-N(\beta/2) \sum_{s=1}^{\beta} (x_{jk}^{(s)})^2}$$

while $e^{ik \sum_{j=1}^N \lambda_j} = e^{ik \text{Tr } X} = e^{ik \sum_{j=1}^N x_{jj}}$. The average of the latter function is thus simply proportional to

$$\left(\int_{-\infty}^{\infty} e^{ikx} e^{-N\beta x^2} dx \right)^N$$

which immediately implies (2.3a).

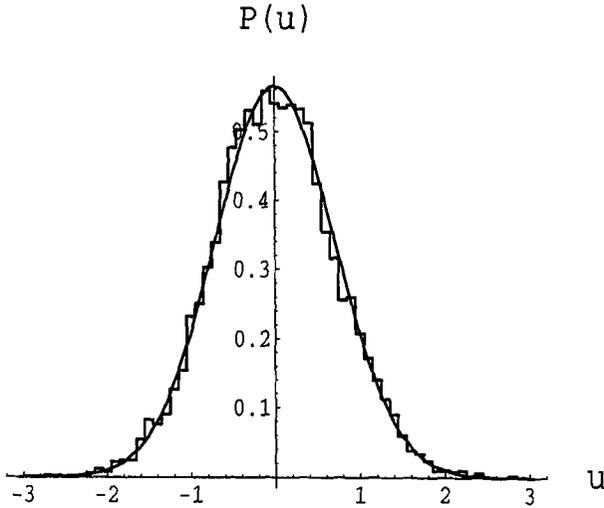


Fig. 1. Empirical and theoretical p.d.f. of the statistic $\sum_{j=1}^N \lambda_j$, where the λ_j are scaled eigenvalues of 2×2 GOE matrices.

In Fig. 1 we illustrate (2.3b) by empirically calculating $P(u)$ for the eigenvalues of 5,000 2×2 matrices from the Gaussian Orthogonal Ensemble of random real symmetric matrices ($\beta = 1$) and comparing the empirical p.d.f. to the theoretical value, Eq. (2.3b) with $\beta = 1$.

2.2. Laguerre Random Matrix Ensemble

If X is a random $n \times m$ ($n \geq m$) matrix with Gaussian entries, which are either all real ($\beta = 1$), complex ($\beta = 2$) or real quaternion ($\beta = 4$) with joint distribution of the elements proportional to $e^{-\beta \text{Tr}(X^\dagger X)/2}$, then the eigenvalue p.d.f. of $X^\dagger X$ is proportional to (1.6) with $N = m$ and $a = n - m + 1 - 2/\beta$. For this so called Laguerre ensemble it is known (see, e.g., ref. 8) that the support of the eigenvalue density is $(0, 4N)$. Scaling the eigenvalues $\lambda_j \mapsto 2N\lambda_j$, so that the support is $(0, 2)$ (the important point here is that the length of the interval is 2, as is required for the validity of (1.5b)), (2.1) with $a(x) = x - \langle x \rangle$ reads

$$\begin{aligned} \tilde{P}(k) = & \frac{1}{C} \int_0^\infty dx_1 x_1^{\beta a/2} e^{-\beta N x_1 + ik(x_1 - \langle x \rangle)} \dots \int_0^\infty dx_N x_N^{\beta a/2} e^{-\beta N x_N + ik(x_N - \langle x \rangle)} \\ & \times \prod_{1 \leq j < k \leq N} |x_k - x_j|^\beta \end{aligned} \tag{2.6}$$

where C is such that $\tilde{P}(0) = 1$ and

$$\begin{aligned} \langle x \rangle &:= \frac{1}{C} \int_0^\infty dx_1 x_1^{\beta a/2} e^{-\beta N x_1} \dots \int_0^\infty dx_N x_N^{\beta a/2} e^{-\beta N x_N} \left(\frac{1}{N} \sum_{j=1}^N x_j \right) \\ &\times \prod_{1 \leq j < k \leq N} |x_k - x_j|^\beta \end{aligned} \tag{2.7}$$

To evaluate (2.7), we note that a simple change of variables gives

$$\begin{aligned} &\int_0^\infty dx_1 x_1^\alpha e^{-\mu b x_1} \dots \int_0^\infty dx_N x_N^\alpha e^{-\mu b x_N} \prod_{1 \leq j < k \leq N} |x_k - x_j|^\beta \\ &= \mu^{-N(\alpha+1) - \beta N(N-1)/2} \int_0^\infty dx_1 x_1^\alpha e^{-b x_1} \dots \int_0^\infty dx_N x_N^\alpha e^{-b x_N} \\ &\times \prod_{1 \leq j < k \leq N} |x_k - x_j|^\beta \end{aligned} \tag{2.8}$$

and that the integral in (2.7) can be obtained from the left hand side of (2.8) by differentiation with respect to μ and setting $\alpha = \beta a/2$, $b = \beta N$ and $\mu = 1$. Applying the same operation to the right hand side of (2.8) gives

$$\langle x \rangle = \frac{1}{2} + \frac{1}{2N} \left(a - 1 + \frac{2}{\beta} \right) \tag{2.9}$$

Now, the non-trivial dependence on k in (2.6) occurs only in the factors $e^{-(\beta N - ik)x_j}$. This dependence can be taken outside the integral by the change of variables $(\beta N - ik)x_j = \beta N y_j$ (this operation is immediately valid for β even; it remains valid for $\text{Re}(\beta) > 0$ by Carlson's theorem) to give

$$\tilde{P}(k) = e^{-ikN\langle x \rangle} \left(\frac{1}{1 - ik/\beta N} \right)^{N^2 \beta \langle x \rangle} \tag{2.10}$$

Taking the inverse transform we have

$$P(u) = \frac{1}{\Gamma(\beta N^2 \langle x \rangle)} (\beta N)^{\beta N^2 \langle x \rangle} (u + N \langle x \rangle)^{\beta N^2 \langle x \rangle - 1} e^{-\beta N(u + N \langle x \rangle)} \tag{2.11}$$

for $u > -N \langle x \rangle$, $P(u) = 0$ for $u < -N \langle x \rangle$. In the limit $N \rightarrow \infty$ we see that (2.10) tends to (2.3a) as expected, thus explicitly demonstrating the universality feature of (1.5).

For the random matrix values $\beta = 1, 2$ and 4 , the result (2.10) can be derived by using the fact that

$$e^{ik \sum_{j=1}^N \lambda_j} = e^{ik \operatorname{Tr} X^\dagger X} = e^{ik \sum_{j=1}^n \sum_{k=1}^m (x_{jk}^{(s)})^2}$$

while the p.d.f. for the matrix X is proportional to

$$\prod_{j=1}^n \prod_{k=1}^m e^{-\beta N \sum_{s=1}^\beta (x_{jk}^{(s)})^2}$$

Hence the average of $e^{ik \sum_{j=1}^N (\lambda_j - \langle \lambda \rangle)}$ with respect to this p.d.f. is proportional to

$$e^{-ikN \langle \lambda \rangle} \left(\int_{-\infty}^{\infty} e^{ikx^2} e^{-\beta N x^2} dx \right)^{\beta mn} = e^{-ikN \langle \lambda \rangle} \left(\frac{1}{1 - ik\beta/N} \right)^{\beta mn/2}$$

which is in agreement with (2.10) (recall that in (2.9) for the random matrix application we have $N = m$ and $a = n - m + 1 - 2/\beta$ so that $\langle x \rangle = 1/2 + (1/2m)(n - m) = n/2m$).

In Fig. 2 we illustrate (2.11) by numerically calculating a histogram for the p.d.f. of $\sum_{j=1}^N (\lambda_j - \langle \lambda \rangle)$ for 3,000 matrices $X^T X$, where X is a 3×2 real rectangular matrix with entries chosen with p.d.f. $(1/\sqrt{2\pi}) e^{-x_{jk}^2/2}$, and comparing it against the theoretical prediction (2.11) with $\beta = 1, N = 2$ and $\langle x \rangle = 1/2 + 1/2N$.

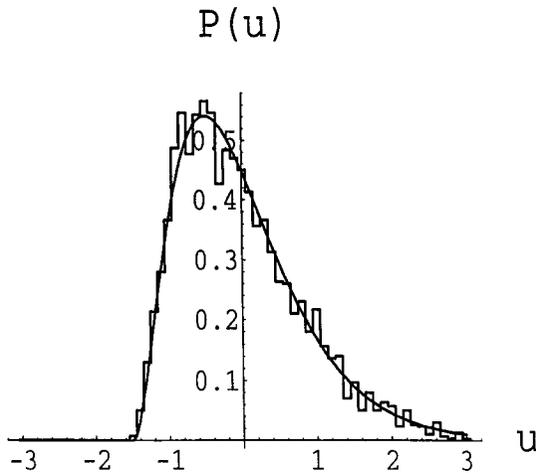


Fig. 2. Empirical and theoretical p.d.f. for the statistic $\sum_{j=1}^N (x_j - \langle x \rangle)$ where the x_j are scaled eigenvalues for the matrix $X^T X$. Here X is a 3×2 real rectangular random matrix with Gaussian entries.

3. DIPOLE MOMENT AND POTENTIAL STATISTIC IN THE CIRCULAR ENSEMBLE

3.1. The Variance

For a general linear statistic A ,

$$\begin{aligned} \sigma^2 &:= \langle (A - \langle A \rangle)^2 \rangle \\ &= \int_I dx \int_I dy a(x) a(y) (\rho_2^T(x, y) + \rho_1(x) \delta(x - y)) \end{aligned} \tag{3.1}$$

where $\rho_1(x)$ denotes the density, ρ_2^T the truncated two particle distribution and I the allowed domain for the particles (eigenvalues). For the circular ensemble we have $I = [0, 2\pi)$ and due to the periodicity we can write

$$\rho_2^T(x, y) + \rho_1(x) \delta(x - y) = \sum_{n=-\infty}^{\infty} r_n e^{i(x-y)n} \tag{3.2}$$

Substituting (3.2) in (3.1) gives

$$\sigma^2 = 2\pi \sum_{n=-\infty}^{\infty} |a_n|^2 r_n, \quad \text{where } a(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx}, \quad x \in [0, 2\pi) \tag{3.3}$$

Consider now the particular linear statistics with $a_1(x) = \frac{1}{2}(x - \pi)$ and $a_2(x) = -\log 2 |\sin x/2|$. We have

$$a_1(x) = \sum_{m=1}^{\infty} \frac{\sin mx}{m} \quad \text{and} \quad a_2(x) = \sum_{m=1}^{\infty} \frac{\cos mx}{m} \tag{3.4}$$

so to analyze σ^2 from the formula in (3.3) it remains to specify the behavior of r_n . This can be done using an heuristic electrostatic argument combined with linear response theory (see, e.g., ref. 19), which gives $r_n \sim |n|/\pi\beta$ for $0 \leq |n| \leq O(N)$. Substituting in (3.4) we therefore have

$$\sigma^2 \sim 2\pi \sum_{\substack{n=-cN \\ n \neq 0}}^{cN} \frac{1}{\pi\beta 2^2 |n|} \sim \frac{1}{\beta} \log N \tag{3.5}$$

for both the dipole moment and potential statistics. Linear response arguments^(20, 21) also give the prediction that in an appropriate macroscopic limit (which here corresponds to $N \rightarrow \infty$) the p.d.f. of any linear

statistic will be Gaussian, thus suggesting that for the dipole moment and potential statistics

$$\left(\frac{1}{\beta} \log N\right)^{1/2} P\left(\left(\frac{1}{\beta} \log N\right)^{1/2} u\right) \sim \left(\frac{1}{2\pi}\right)^{1/2} e^{-u^2/2} \quad \text{as } N \rightarrow \infty \quad (3.6)$$

3.2. The Case $\beta = 2$

From the Vandermonde determinant expansion

$$\det[z_j^{k-1}]_{j,k=1,\dots,N} = \prod_{1 \leq j < k \leq N} (z_k - z_j) \quad (3.7)$$

it is straightforward to show (see, e.g., ref. 22) that with W given by (1.7) and $\beta = 2$ (2.1) can be rewritten in terms of a Toeplitz determinant:

$$\tilde{P}(k) = D_N[e^{ika(\theta)}] := \det \left[\frac{1}{2\pi} \int_0^{2\pi} d\theta e^{ika(\theta)} e^{i\theta(m-n)} \right]_{m,n=1,\dots,N} \quad (3.8)$$

Using an asymptotic formula for the large- N form of D_N , known as the Fisher–Hartwig conjecture,^(12–14) we can rigorously establish (3.6).

Proposition 3.1 (Fisher–Hartwig conjecture). In (3.8) let

$$ika(\theta) = g(\theta) - i \sum_{r=1}^R \beta_r (\pi - (\theta - \theta_r)) + \sum_{r=1}^R \alpha_r \log |2 - 2 \cos(\theta - \theta_r)| \quad (3.9)$$

and assume $g(\theta) = \sum_{p=-\infty}^{\infty} g_p e^{i\theta p}$ where $\sum_{p=-\infty}^{\infty} |p| |g_p|^2 < \infty$. Then for $\text{Re}(\alpha_r) > -1/2$ and $\text{Im}(\beta_r) = 0$

$$D_N[e^{ika(\theta)}] \sim e^{g_0 N} e^{\sum_{r=1}^R (\alpha_r^2 - \beta_r^2) \log N} E \quad (3.10)$$

where E is independent of N . To specify E , write $g(\theta) - g_0 = g_+(e^{i\theta}) + g_-(e^{-i\theta})$ where $g_+(z) = \sum_{p=1}^{\infty} g_p z^p$ and $g_-(z) = \sum_{p=-\infty}^{-1} g_p z^p$. Then

$$E = e^{\sum_{k=1}^{\infty} k g_k g_{-k}} \prod_{r=1}^R e^{-(\alpha_r + \beta_r) g_-(e^{i\theta_r})} e^{-(\alpha_r - \beta_r) g_+(e^{-i\theta_r})} \\ \times \prod_{1 \leq r < s \leq R} |1 - e^{i(\theta_s - \theta_r)}|^{-(\alpha_r + \beta_r)(\alpha_s - \beta_s)} \prod_{r=1}^R \frac{G(1 + \alpha_r + \beta_r) G(1 + \alpha_r - \beta_r)}{G(1 + 2\alpha_r)} \quad (3.11)$$

where G is the Barnes G -function defined by

$$G(z + 1) = (2\pi)^{z/2} \exp(-z/2 - (\gamma + 1)z^2/2) \prod_{k=1}^{\infty} \left(1 + \frac{z}{k}\right)^k \exp(-z + z^2/k) \tag{3.12}$$

(γ denotes Euler’s constant), which has the special values $G(1) = G(2) = 1$ and satisfies the functional relation

$$G(z + 1) = \Gamma(z) G(z) \tag{3.13}$$

First consider the application of (3.9) to the calculation of the p.d.f. for the potential statistic. From (3.8) we have

$$\tilde{P}(k) = D_N[(2 - 2 \cos \theta)^{ik/2}] \tag{3.14}$$

so we take $g(\theta) = 0$, $R = 1$, $\beta_1 = 0$, $\alpha_1 = (ik/2)(1/2 \log N)^{-1/2}$ and $\theta_1 = 0$ in (3.9) and (3.10) to conclude that

$$\tilde{P}\left(\frac{k}{(1/2 \log N)^{1/2}}\right) \sim e^{-k^2/2} \tag{3.15}$$

which is equivalent to (3.6) with $\beta = 2$.

For the dipole moment statistic we have from (3.8) that

$$\tilde{P}(k) = D_N[e^{-(ik/2)(\pi - \theta)}] \tag{3.16}$$

Thus we take $g(\theta) = 0$, $R = 1$, $\alpha_1 = 0$, $\beta_1 = k/2$ and $\theta_1 = 0$ in (3.9) and (3.10) to conclude that (3.15) again holds, as predicted from (3.6).

3.3. General β

Here we will show that $\tilde{P}(k)$ for the dipole moment and potential statistics in the circular ensemble can be given in closed form for general β , and prove that the expected asymptotic behaviour (3.6) is valid for general rational β (at least). Our chief tool for this purpose is a constant term identity of Morris,⁽¹⁵⁾ written as the multidimensional integral evaluation⁽²³⁾

$$\begin{aligned} M_n(a, b, c) &:= \prod_{l=1}^n \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta_l e^{i(a-b)\theta_l/2} |1 + e^{i\theta_l}|^{a+b} \prod_{1 \leq j < k \leq n} |e^{i\theta_k} - e^{i\theta_j}|^{2c} \\ &= \prod_{j=0}^{n-1} \frac{\Gamma(c(j+1) + 1) \Gamma(cj + a + b + 1)}{\Gamma(cj + a + 1) \Gamma(cj + b + 1) \Gamma(1 + c)} \end{aligned} \tag{3.17}$$

First consider the dipole moment statistic. From (1.7) and (2.1) we have

$$\begin{aligned} \tilde{P}(k) &= \frac{1}{C} \prod_{l=1}^N \int_0^{2\pi} d\theta_l e^{ik(\theta_l - \pi)/2} \prod_{1 \leq j < k \leq N} |e^{i\theta_k} - e^{i\theta_j}|^\beta \\ &= \frac{M_n(k/2, -k/2, \beta/2)}{M_N(0, 0, \beta/2)} \\ &= \prod_{j=0}^{N-1} \frac{(\Gamma(\beta j/2 + 1))^2}{\Gamma(\beta j/2 + k/2 + 1) \Gamma(\beta j/2 - k/2 + 1)} \end{aligned} \tag{3.18}$$

Although (3.18) is a closed form evaluation, it is not convenient for the determination of the large- N asymptotics. This same problem has been faced in an earlier application of the Morris constant term identity.⁽²⁴⁾ Its resolution is to make use of the identity

$$\prod_{j=0}^{N-1} \Gamma(\alpha + 1 + jc) = c^{cN(N-1)/2 + N(\alpha + 1/2)} (2\pi)^{-N(c-1)/2} \prod_{l=1}^c \frac{G(N + (\alpha + l)/c)}{G((\alpha + l)/c)} \tag{3.19}$$

valid for c a positive integer. We can use (3.19) in (3.18) for all rational β , $\beta/2 = s/r$ with s and r relatively prime positive integers say. This is done by replacing N by rN and noting

$$\prod_{j=0}^{rN-1} \Gamma\left(\alpha + 1 + \frac{js}{r}\right) = \prod_{p=0}^{r-1} \prod_{k=0}^{N-1} \Gamma\left(\alpha + 1 + \frac{ps}{r} + sk\right) \tag{3.20}$$

Thus for $\beta/2 = s/r$ and with N replaced by rN , (3.18) can be rewritten to read

$$\begin{aligned} \tilde{P}(k) &= \prod_{p=0}^{r-1} \prod_{l=1}^s \frac{(G(N + ps/r + l/s))^2 G(ps/r + (l + k/2)/s)}{G(N + ps/r + (l + k/2)/s) G(N + ps/r + (l - k/2)/s)} \\ &\quad \times \frac{G(ps/r + (l - k/2)/s)}{(G(l/s + ps/r))^2} \end{aligned} \tag{3.21}$$

The large- N asymptotics of $\tilde{P}(k)$ can be deduced from (3.21) by making use of the asymptotic formula⁽²⁵⁾

$$\begin{aligned} \log \left(\frac{G(N + a + 1)}{G(N + b + 1)} \right) &\sim (b - a) N + \frac{a - b}{2} \log 2\pi \\ &\quad + \left((a - b) N + \frac{a^2 - b^2}{2} \right) \log N + o(1) \end{aligned} \tag{3.22}$$

This gives

$$\log \tilde{P}(k) \sim -\frac{k^2}{2\beta} \log N + \sum_{p=0}^{r-1} \sum_{l=1}^s \log \frac{G(ps/r + (l+k/2)/s) G(ps/r + (l-k/2)/s)}{(G(l/s + ps/r))^2} + o(1) \tag{3.23}$$

which implies

$$\tilde{P}\left(\frac{k}{(1/\beta \log N)^{1/2}}\right) \sim e^{-k^2/2} \tag{3.24}$$

in precise agreement with the Fourier transform of the expected asymptotic behaviour (3.6).

Let us now turn our attention to the potential statistic. In this case, from (2.1), (1.7) and (3.17) we have

$$\begin{aligned} \tilde{P}(k) &= \frac{1}{C} \prod_{l=1}^N \int_0^{2\pi} d\theta_l |1 - e^{i\theta_l}|^{-ik} \prod_{1 \leq j < k \leq N} |e^{i\theta_k} - e^{i\theta_j}|^\beta \\ &= \frac{M_N(-ik/2, -ik/2, \beta/2)}{M_N(0, 0, \beta/2)} \\ &= \prod_{j=0}^{N-1} \frac{\Gamma(\beta j/2 - ik + 1) \Gamma(\beta j/2 + 1)}{(\Gamma(\beta j/2 - ik/2 + 1))^2} \end{aligned} \tag{3.25}$$

Proceeding as above, for rational β ($\beta/2 = s/r$), by using (3.19) and (3.20) we can rewrite (3.25) as

$$\begin{aligned} \tilde{P}(k) &= \prod_{p=0}^{r-1} \prod_{l=1}^s \frac{G(N + ps/r + (l - ik)/s) G(N + ps/r + l/s)}{(G(N + ps/r + (l - ik/2)/s))^2 G(ps/r + (l - ik)/s)} \\ &\quad \times \frac{(G(ps/r + (l - ik/2)/s))^2}{(G(ps/r + l/s))} \end{aligned} \tag{3.26}$$

The asymptotic formula (3.22) now gives that for $N \rightarrow \infty$

$$\log \tilde{P}(k) \sim -\frac{k^2}{2\beta} \log N + \sum_{p=0}^{r-1} \sum_{l=1}^s \log \frac{(G(ps/r + (l - ik/2)/s))^2}{G(ps/r + (l - ik)/s) G(ps/r + l/s)} + o(1) \tag{3.27}$$

which again implies the anticipated result (3.24).

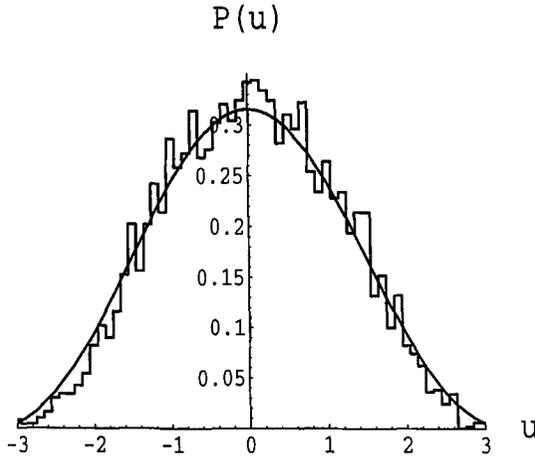


Fig. 3. Empirical and theoretical p.d.f. for the statistic $\frac{1}{2}\sum_{j=1}^N(\theta_j - \pi)$, where the θ_j ($0 \leq \theta_j < 2\pi$) are the phases for the eigenvalues of 3×3 CUE matrices.

In Fig. 3 we illustrate (3.18) by numerically constructing a histogram for the p.d.f. of $\frac{1}{2}\sum_{j=1}^N(\theta_j - \pi)$ for 5,000 3×3 matrices from the CUE (these are constructed according to the procedure specified in ref. 26), and comparing it against the inverse transform of (3.18) in the case $N=3$, $\beta=2$.

4. RELATIONSHIP BETWEEN THE VARIANCE OF THE DIPOLE MOMENT STATISTIC AND THE SUSCEPTIBILITY

In this final section we will show how the result (2.5) can be anticipated from its interpretation as a susceptibility. We recall that in macroscopic electrostatics the susceptibility tensor χ relates the electric polarization density of a Coulomb system, confined to a region A in a vacuum, to the applied electric field. The laws of macroscopic electrostatics allow the components of χ to be expressed in terms of the dielectric constant of the system. In particular, for a conducting ellipse, this theory gives⁽²⁷⁾

$$\chi_{xx} = \frac{1}{2\pi} \frac{L_x + L_y}{L_x} \quad (4.1)$$

where L_x is the length and L_y the width of the ellipse.

On the other hand, the susceptibility can be related to the microscopic quantities by linear response theory. With the electric field applied in the x -direction, this approach gives⁽²⁸⁾

$$\chi_{xx} = \frac{\beta}{|A|} (\langle P_x^2 \rangle - \langle P_x \rangle^2) \quad (4.2)$$

where P_x is the x -component of the instantaneous polarization (or equivalently, dipole moment) $\vec{P} := \sum_j q_j \vec{r}_j$.

Comparing (4.1) and (4.2), and noting that $|A| = (\pi/4) L_x L_y$, we obtain

$$\langle P_x^2 \rangle - \langle P_x \rangle^2 = \frac{1}{8\beta} L_x (L_x + L_y) \quad (4.3)$$

Now we construct an interval of length 2 in the x -direction from the ellipse by setting $L_x = 2$, $L_y = 0$. This gives

$$\langle P_x^2 \rangle - \langle P_x \rangle^2 = \frac{1}{2\beta} \quad (4.4)$$

which is in precise agreement with the result (2.5) for the variance of the dipole moment statistic.

ACKNOWLEDGMENTS

The authors thank the referee for a careful reading. This work was supported by the Australian Research Council.

REFERENCES

1. O. Costin and J. L. Lebowitz, *Phys. Rev. Lett.* **75**:69 (1995).
2. A. Soshnikov, Global level spacings distribution for large random matrices from classical compact groups, preprint.
3. F. J. Dyson, in *Chen Ning Yang*, ed. S.-T. Yau (Internat. Press, Cambridge, MA, 1995), p. 131.
4. M. M. Fogler and B. I. Shklovskii, *Phys. Rev. Lett.* **74**:3312 (1995).
5. K. Johansson, On fluctuations of random Hermitian matrices, preprint.
6. K. Johansson, On random matrices associated with compact classical groups, preprint.
7. E. L. Basor, Distribution functions for random variables for ensembles of positive Hermitian matrices, preprint.
8. T. H. Baker and P. J. Forrester, solv-int/9608004.
9. E. Brézin and Z. A. Zee, *Nucl. Phys. B* **402**:613 (1993).
10. C. W. J. Beenakker, *Nucl. Phys. B* **422**, 515 (1994).
11. A. Alastuey and B. Janovici, *J. Stat. Phys.* **34**:557 (1984).

12. M. E. Fisher and R. E. Hartwig, *Adv. Chem. Phys.* **15**:333 (1968).
13. A. Böttcher and B. Silbermann, *Analysis of Toeplitz operators* (Berlin, Akademie, 1989).
14. E. L. Basor and C. A. Tracy, *Physica A* **177**:167 (1991).
15. W. G. Morris, Constant term identities for finite and affine root systems, Ph.D. thesis, University of Wisconsin, Madison (1982).
16. A. Selberg, *Norsk Mat. Tids.* **26**:167 (1944).
17. E. C. Titchmarsh, *Theory of Functions*, p. 186 (Oxford University Press, London, 1939).
18. E. Brézin, C. Itzykson, G. Parisi, and J. B. Zuber, *Commun. Math. Phys.* **59**:35 (1978).
19. P. J. Forrester, *Nucl. Phys. B* **435**:421 (1995).
20. H. D. Politzer, *Phys. Rev. B* **40**:11917 (1989).
21. B. Jancovici, *J. Stat. Phys.* **80**:445 (1995).
22. G. Szegő, *Orthogonal Polynomials* (AMS, Providence RI, 1967).
23. P. J. Forrester, *Mod. Phys. Lett. B* **9**:359 (1995).
24. P. J. Forrester, *Physics Lett. A* **163**:121 (1992).
25. E. W. Barnes, *Quarterly J. Pure Appl. Math.* **31**:264 (1900).
26. K. Zyczkowski and M. Kus, *J. Phys. A* **27**:4235 (1994).
27. Ph. Choquard, B. Piller, and R. Rentsch, *J. Stat. Phys.* **46**:599 (1987).
28. Ph. Choquard, B. Piller, and R. Rentsch, *J. Stat. Phys.* **43**:197 (1986).

Communicated by J. L. Lebowitz