# MOCK-GAUSSIAN BEHAVIOUR

C.P. HUGHES

ABSTRACT. We show that the moments of the smooth counting function of a set of points encodes the same information as the $n$-level density of these points. If the points are the eigenvalues of matrices taken from the classical compact groups with Haar measure, then we show that the first few moments of the smooth counting function are Gaussian, while the distribution is not. The same phenomenon occurs for smooth counting functions of the zeros of $L$–functions, and we give examples relating to each classical compact group. The advantage of calculating moments of the counting function is that combinatorially, they are far easier to handle than the $n$-level densities.

One of the connections between random matrix theory and number theory is that the correlations and densities between eigenangles of matrices chosen at random from the classical compact groups appear to be the same as correlations and densities between zeros of $L$–functions taken from certain families. This connection was first suggested by Katz and Sarnak [7], and is discussed in more detail elsewhere in this Proceedings.

Rather than study densities of zeros, the purpose of this note is to argue that the same results can be obtained more easily by calculating the moments of smooth counting functions. The mock-Gaussian behaviour of the title refers to the fact that in all the cases examined here, the first few moments are Gaussian, while the overall distribution is not.

In the first section we will explicitly demonstrate the connections between the moments of smooth counting functions and the $n$-level densities. In the middle section we will sketch the proofs of mock-Gaussian behaviour for the classical compact groups. At the end of the paper we will give examples from number theory where mock-Gaussian behaviour holds, and therefore re-prove, in a manner that does not require a lot of combinatorial sieving, that the $n$-level densities of these examples agree with what one obtains from random matrix theory.

## 1. General connections between moments and $n$-level densities

Let $(x_1, \dots, x_N)$ be chosen from some probability distribution on $\mathbb{R}^N$. The $n$-level density function for this distribution is

$$D_n(g_1, \dots, g_n) = \mathbb{E}\left[ \sideset{}{'}\sum_{1 \le j_1, \dots, j_n \le N} g_1(x_{j_1}) \dots g_n(x_{j_n}) \right]. \tag{1}$$

Here $\sum'$ denotes the sum over *distinct* indices, that is $j_i \neq j_l$ for $i \neq l$, and $\mathbb{E}$ denotes expectation with respect to the density function.

*Remark.* Sometimes the numbers $x_j$ have a symmetry condition. An example would be if for each $j$ there exists an $j'$ such that $x_j = -x_{j'}$. In that case sometimes the $n$-level density is defined with the further condition that $j_i \neq j_{l'}$ imposed, as well as the current distinctness condition that $j_i \neq j_l$. We assume the $x_j$ are desymmetrized.

*Date*: 8th April, 2004.

Another statistic is the moments of the smooth counting function,

$$M_n(f) = \mathbb{E}\left[\left(\sum_{j=1}^N f(x_j)\right)^n\right] = \mathbb{E}\left[\sum_{1\le j_1,\dots,j_n\le N} f(x_{j_1})\cdots f(x_{j_n})\right] \tag{2}$$

which is also just the $n$th moment of the one level density.

We will show that the two statistics provide the same information. This might seem a little surprising since the $n$-level density appears to have a more general test function, being a product of $n$ different functions.

1.1. **The $n$-level density implies the moments of the counting function.** Note that the sums in (2) range unrestrictedly over all variables (they include both diagonal and off-diagonal terms), whereas the sum in (1) is over distinct variables (off-diagonals). This problem can be overcome by summing over the diagonals separately.

**Definition 1.** *We say $\sigma$ is a set partition of $m$ elements into $r$ non-empty blocks if*

$$\sigma : \{1,\dots,m\} \longrightarrow \{1,\dots,r\} \tag{3}$$

*satisfies*

1. *For every $q \in \{1,\dots,r\}$ there exists at least one $j$ such that $\sigma(j) = q$ (this is the non-emptiness of the blocks).*
2. *For all $j$, either $\sigma(j) = 1$ or there exists a $k < j$ such that $\sigma(j) = \sigma(k) + 1$.*

*The collection of all set partitions of $m$ elements into $r$ blocks is denoted $P(m,r)$.*

Roughly speaking, if we think of $\{1,\dots,r\}$ as denoting ordered pigeonholes, then $\sigma(j)$ either goes into a non-empty pigeonhole, or into the next empty hole.

*Remark.* The number of $\sigma \in P(m,r)$ is equal to $S(m,r)$, a Stirling number of the second kind. The number of set partitions of $m$ elements into any number of non-empty blocks is $\sum_{r=1}^m S(m,r) = B_m$, a Bell number.

Therefore,

$$\sum_{1\le j_1,\dots,j_m\le N} g_1(x_{j_1})\dots g_m(x_{j_m}) = \sum_{r=1}^m \sum_{\sigma\in P(m,r)} \sideset{}{'}\sum_{\substack{1\le i_1,\dots,i_r\le N\\ i_j \text{ all distinct}}} g_1(x_{i_{\sigma(1)}})\dots g_m(x_{i_{\sigma(m)}}) \tag{4}$$

(think of this a summing over the diagonals separately).

From this we may conclude that

$$\left(\sum_{j=1}^N f(x_j)\right)^n = \sum_{r=1}^m \sum_{\sigma\in P(m,r)} D_r(f^{\lambda_1},\dots,f^{\lambda_r}) \tag{5}$$

where $\lambda_q = \#\{j \ : \ 1 \le j \le m \ , \ \sigma(j) = q\}$.

Therefore, from knowing the $n$-level densities, one can immediately deduce the moments of the smooth counting function.

1.2. **The moments of the counting function imply the $n$-level densities.** Let us create an inductive hypothesis that for all $1 \le r < n$, $\mathbb{E}[D_r(g_1,\dots,g_r)]$ can be written in terms of $\mathbb{E}[M_m(f)]$ for various $f$ with $1 \le m \le r$.

An inclusion / exclusion type formula gives

$$\sum_{S\subseteq\{1,\dots,n\}} (-1)^{n-|S|}\left(\sum_{i\in S} a_i\right)^n = n!a_1\dots a_n \tag{6}$$

where the sum is over *all* subsets of $\{1,\dots,n\}$, and so removing the subset $\{1,\dots,n\}$ we get

$$M_n(g_1+\cdots+g_n) = (M_1(g_1)+\cdots+M_1(g_n))^n \tag{7}$$

$$= n!M_1(g_1)\dots M_1(g_n) - \sum_{S\subsetneq\{1,\dots,n\}} (-1)^{n-|S|} M_n\left(\sum_{i\in S} g_i\right) \tag{8}$$

where the sum is over all *proper* subsets of $\{1,\dots,n\}$.

By (4) the term $n!M_1(g_1)\dots M_n(g_n)$ equals $n!D_n(g_1,\dots,g_n)$ plus terms involving $D_r$ for $r<n$. By the inductive hypothesis those terms can be written in terms of the moments of the counting function, and we are done.

This is more easily seen in terms of an example. Consider the 2-level density. We have

$$M_2(g_1+g_2) = [M_1(g_1)+M_1(g_2)]^2 \tag{9}$$

$$= 2M_1(g_1)M_1(g_2) + M_1(g_1)^2 + M_1(g_2)^2 \tag{10}$$

$$= 2M_1(g_1)M_1(g_2) + M_2(g_1) + M_2(g_2). \tag{11}$$

Now,

$$M_1(g_1)M_1(g_2) = D_2(g_1,g_2) + M_1(g_1g_2) \tag{12}$$

and so we see that from knowing $\mathbb{E}[M_2(f)]$ and $\mathbb{E}[M_1(f)]$, we have recovered $\mathbb{E}[D_2(g_1,g_2)]$, since

$$\mathbb{E}[D_2(g_1,g_2)] = \tfrac{1}{2}\,\mathbb{E}[M_2(g_1+g_2)] - \mathbb{E}[M_1(g_1g_2)] - \tfrac{1}{2}\,\mathbb{E}[M_2(g_1)] - \tfrac{1}{2}\,\mathbb{E}[M_2(g_2)]. \tag{13}$$

1.3. **Restricted range.** Often in number theory, it is only possible to prove the $n$-level density or the moments of the counting function for test functions whose Fourier transforms are supported in a restricted range. However, the above arguments go through without change, and if we know $M_m(f)$ for all $f$ with $\operatorname{supp}\widehat{f}\in[-\alpha/m,\alpha/m]$, for all $1\le m\le n$, then we know $D_n(g_1,\dots,g_n)$ for all $g_i$ with $\operatorname{supp}\widehat{g_i}\in[-\alpha/n,\alpha/n]$. We should remark that this is a little bit weaker than what is often proved within number theory, where the support restriction is frequently of the form $\sum_{j=1}^n \alpha_i = \alpha$, where $\operatorname{supp}\widehat{g_i}\in[-\alpha_i,\alpha_i]$. Clearly the result above, where $\alpha_i=\alpha/n$, fits this.

## 2. Mock-Gaussian behaviour in the classical compact groups

The classical compact groups are:

- $\mathrm{U}(N)$, the group of all $N\times N$ unitary matrices.
- $\mathrm{SO}(2N)$, the subgroup of $\mathrm{U}(2N)$ containing the even orthogonal matrices with determinant one. If $e^{\mathrm{i}\theta}$ is an eigenvalue, then so is $e^{-\mathrm{i}\theta}$.
- $\mathrm{SO}(2N+1)$, the subgroup of $\mathrm{U}(2N+1)$ containing the odd orthogonal matrices with determinant one. If $e^{\mathrm{i}\theta}$ is an eigenvalue, then so is $e^{-\mathrm{i}\theta}$, and there is an additional eigenvalue at 1.
- $\mathrm{USp}(2N)$, the subgroup of $\mathrm{U}(2N)$ containing the symplectic unitary matrices. That is, $UU^\dagger = I_{2N}$ and $U^tJU=J$ where $J=\begin{pmatrix}0 & I_N\\ -I_N & 0\end{pmatrix}$. Again, if $e^{\mathrm{i}\theta}$ is an eigenvalue then $e^{-\mathrm{i}\theta}$ is also an eigenvalue.

Averages with respect to Haar measure over all of these compact classical groups can be written in the form

$$\mathbb{E}_{\mathrm{G}(N)}\left[\prod_{n=1}^N f(\theta_n)\right] = \frac{1}{N!}\int_{\mathbb{T}^N} \det_{N\times N}\{Q^{\mathrm{G}(N)}(x_i,x_j)\}\prod_{n=1}^N f(x_n)\,\mathrm{d}x_n \tag{14}$$

where G($N$) denotes one of U($N$), SO($2N$), SO($2N+1$) or USp($2N$), ($N$ being the number of independent eigenvalues). We call $Q^{\mathrm{G}(N)}$ the kernel of the group, and $\mathbb{T}$ the range. Let

$$S_N(z) = \frac{1}{2\pi}\frac{\sin(Nz/2)}{\sin(z/2)}. \tag{15}$$

then, the kernels and ranges are given by

| Group G($N$) | Kernel $Q^{\mathrm{G}(N)}(x,y)$ | Range $\mathbb{T}$ |
|---|---|---|
| U($N$) | $S_N(x-y)$ | $(-\pi,\pi]$ |
| SO($2N$) | $S_{2N-1}(x-y)+S_{2N-1}(x+y)$ | $[0,\pi]$ |
| SO($2N+1$) | $S_{2N}(x-y)-S_{2N}(x+y)$ | $[0,\pi]$ |
| USp($2N$) | $S_{2N+1}(x-y)-S_{2N+1}(x+y)$ | $[0,\pi]$ |

Choose a function $\phi$ from the set of all real functions whose Fourier transform,

$$\widehat{\phi}(u) := \int_{-\infty}^{\infty} \phi(x) e^{-2\pi i x u}\,\mathrm{d}x, \tag{16}$$

is smooth and compactly supported. Note that for any $A > 1$, this means $\phi(x) \ll (1+|x|)^{-A}$ for all $x \in \mathbb{R}$. From such a $\phi$ we create a $2\pi$-periodic function

$$F_M(\theta) := \sum_{n=-\infty}^{\infty} \phi\left(\frac{M}{2\pi}(\theta + 2\pi n)\right). \tag{17}$$

Given an $M \times M$ unitary matrix $U$ with eigenangles $\theta_n$, the smooth counting function, or one-level density, or linear statistic, of the eigenangles of the matrix $U$ is

$$Z_\phi(U) := \operatorname{Tr} F_M(U) := \sum_{n=1}^{M} F_M(\theta_n). \tag{18}$$

*Remark.* Note that the matrix is chosen to be size $M \times M$, and it has $N$ independent eigenangles. The counting function sums over all $M$ of the eigenangles, but Haar measure integrates only over the $N$ independent terms.

Note that due to the rapid decay on $\phi$, $Z_\phi(U)$ has the largest contribution from the eigenvalues of $U$ close to 1. We will study moments of $Z_\phi(U)$ when $U$ is averaged over one of the classical compact groups, and show that the first few moments are Gaussian, but the higher ones are not.

**Theorem 1** (Hughes and Rudnick [5, 6]). *If $\phi$ is chosen so that $\widehat{\phi}$ is smooth and has compact support, then:*

i) *If* $\operatorname{supp}\widehat{\phi} \subseteq [-2/m, 2/m]$ *then the first $m$ moments of $Z_\phi(U)$ over the unitary group* U($N$) *converge as $N \to \infty$ to the moments of a Gaussian random variable with mean*

$$\mu_\phi^{\mathrm{U}} = \int_{-\infty}^{\infty} \phi(x)\,\mathrm{d}x \tag{19}$$

*and variance*

$$(\sigma_\phi^{\mathrm{U}})^2 = \int_{-1}^{1} |u||\widehat{\phi}(u)|^2\,\mathrm{d}u. \tag{20}$$

ii) *If $\phi$ is even, and* $\operatorname{supp}\widehat{\phi} \subseteq [-1/m, 1/m]$, *then the first $m$ moments of $Z_\phi(U)$ when averaged over the symplectic group* USp($2N$) *converge to the moments of a Gaussian with mean*

$$\mu_\phi^{\mathrm{USp}} = \int_{-\infty}^{\infty} \phi(x)\,\mathrm{d}x - \int_0^1 \widehat{\phi}(u)\,\mathrm{d}u \tag{21}$$

*and variance*

$$(\sigma_\phi^{\mathrm{USp}})^2 = 2\int_{-1/2}^{1/2} |u|\widehat{\phi}(u)^2 \, \mathrm{d}u. \tag{22}$$

iii) *If* $\phi$ *is even, and* $\operatorname{supp}\widehat{\phi} \subseteq [-1/m, 1/m]$*, then the first* $m$ *moments of* $Z_\phi(U)$ *when averaged over either* $\mathrm{SO}(2N)$ *or* $\mathrm{SO}(2N+1)$ *converge to the moments of a Gaussian with mean*

$$\mu_\phi^{\mathrm{SO}} = \int_{-\infty}^{\infty} \phi(x) \, \mathrm{d}x + \int_0^1 \widehat{\phi}(u) \, \mathrm{d}u \tag{23}$$

*and variance*

$$(\sigma_\phi^{\mathrm{SO}})^2 = 2\int_{-1/2}^{1/2} |u|\widehat{\phi}(u)^2 \, \mathrm{d}u. \tag{24}$$

To re-phrase this theorem, part (i) says that if $\operatorname{supp}\widehat{\phi} \subseteq [-\frac{2}{m}, \frac{2}{m}]$, then

$$\lim_{N\to\infty} \mathbb{E}_{\mathrm{U}(N)}\left[\left(Z_\phi(U) - \mu_\phi^{\mathrm{U}}\right)^m\right] = \begin{cases} \frac{(2k)!}{2^k k!}(\sigma_\phi^{\mathrm{U}})^{2k} & \text{if } m = 2k \text{ is even} \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

*Remark.* This theorem is sharp, in the sense that if the support of $\widehat{\phi}$ was increased beyond $[-1/m, 1/m]$ (or $[-2/m, 2/m]$ in the unitary case), then the $m$th moment ceases to be Gaussian for $m \geq 3$.

This theorem can be proven via a study of the cumulants (though in [6] a different approach is taken), since if one knows the first $\ell$ cumulants then one knows the first $\ell$ moments. If $\theta_1, \ldots, \theta_N$ are the independent eigenangles of a matrix $U \in \mathrm{G}(N)$, then for a $2\pi$-periodic function $g$, the cumulants of $\sum_{n=1}^N g(\theta_n)$ are defined as

$$\log \mathbb{E}_{\mathrm{G}(N)}\left[\exp\left(t\sum_{n=1}^N g(\theta_n)\right)\right] = \sum_{\ell=1}^{\infty} \frac{t^\ell}{\ell!} C_\ell^{\mathrm{G}(N)}(g) \tag{26}$$

and for the classical compact groups they can be written in terms of the kernel as follows (this is non-obvious: See, for example, [9])

$$C_\ell^{\mathrm{G}(N)}(g) = \sum_{m=1}^{\ell} \sum_{\sigma\in P(\ell,m)} (-1)^{m+1}(m-1)! \int_{\mathbb{T}^m} \prod_{j=1}^m g^{\lambda_j}(x_j) Q^{\mathrm{G}(N)}(x_j, x_{j+1}) \, \mathrm{d}x_j \tag{27}$$

where we identify $x_{m+1}$ with $x_1$. Here $P(\ell, m)$ is the set of all partitions of $\ell$ objects into $m$ non-empty blocks, as in Definition 1, where the $j$th block has $\lambda_j = \lambda_j(\sigma)$ elements.

Put

$$C_{\ell,N}^{\mathrm{even}}(g) = \frac{1}{2}\sum_{m=1}^{\ell} \sum_{\sigma\in P(\ell,m)} (-1)^{m+1}(m-1)! \int_{[-\pi,\pi]^m} \prod_{j=1}^m g^{\lambda_j}(x_j) S_N(x_j - x_{j+1}) \, \mathrm{d}x_j \tag{28}$$

and

$$C_{\ell,N}^{\mathrm{odd}}(g) = \frac{1}{2}\sum_{m=1}^{\ell} \sum_{\sigma\in P(\ell,m)} (-1)^{m+1}(m-1)! \int_{[-\pi,\pi]^m} \prod_{j=1}^m g^{\lambda_j}(x_j) S_N(x_j - \epsilon_j x_{j+1}) \, \mathrm{d}x_j \tag{29}$$

where $\epsilon_j = +1$ for $j = 1, \ldots, m-1$ and $\epsilon_m = -1$, and where $S_N(z)$ is defined in (15).

Expanding out the kernels $Q^{\mathrm{G}(N)}(x, y)$ for the various groups, we find

$$C_\ell^{U(N)}(g) = 2C_{\ell,N}^{\mathrm{even}}(g) \tag{30}$$

$$C_\ell^{\mathrm{USp}(2N)}(g) = C_{\ell,2N+1}^{\mathrm{even}}(g) - C_{\ell,2N+1}^{\mathrm{odd}}(g) \tag{31}$$

$$C_\ell^{\mathrm{SO}(2N)}(g) = C_{\ell,2N-1}^{\mathrm{even}}(g) + C_{\ell,2N-1}^{\mathrm{odd}}(g) \tag{32}$$

$$C_\ell^{\mathrm{SO}(2N+1)}(g) = C_{\ell,2N}^{\mathrm{even}}(g) - C_{\ell,2N}^{\mathrm{odd}}(g) \tag{33}$$

Extending the combinatorics introduced by Soshnikov [9], we deduced in [5] that if $\ell \geq 2$,

$$\left|C^{\mathrm{odd}}_{\ell,L}(g)\right| \leq \mathrm{const}_\ell \sum_{\substack{\mathbf{k}\in\mathbb{Z}^\ell \\ |k_1|+\cdots+|k_\ell|>L}} |g_{k_1}|\ldots|g_{k_\ell}| \tag{34}$$

and if $\ell \geq 3$

$$\left|C^{\mathrm{even}}_{\ell,L}(g)\right| \leq \mathrm{const}_\ell \sum_{\substack{k_1+\cdots+k_\ell=0 \\ |k_1|+\cdots+|k_\ell|>2L}} |g_{k_1}|\ldots|g_{k_\ell}| \tag{35}$$

where $g_k$ is the $k$th Fourier coefficient of $g$, so $g(\theta) = \sum_{k=-\infty}^{\infty} g_k e^{\mathrm{i}\theta}$.

In order to prove Theorem 1 we must show that for $\ell \geq 3$, the $\ell$th cumulant of

$$Z_\phi(U) := \sum_{n=1}^{M} F_M(\theta_n) \tag{36}$$

tends to zero when averaged over the unitary group if $\operatorname{supp}\widehat{\phi} \subseteq [-2/\ell, 2/\ell]$, and tends to zero when averaged over the symplectic or orthogonal groups if $\operatorname{supp}\widehat{\phi} \subseteq [-1/\ell, 1/\ell]$. Recall that $M$ is the total number of eigenangles of the matrix $U$, while $N$ is the number of independent ones. Therefore, we choose $g$ as follows:

- If $\mathrm{G}(N) = \mathrm{U}(N)$, we choose $g(\theta) = F_N(\theta)$.
- If $\mathrm{G}(N) = \mathrm{USp}(2N)$, $\phi$ must be even, and we choose $g(\theta) = 2F_{2N}(\theta)$
- If $\mathrm{G}(N) = \mathrm{SO}(2N)$, $\phi$ must be even, and we choose $g(\theta) = 2F_{2N}(\theta)$
- If $\mathrm{G}(N) = \mathrm{SO}(2N+1)$, $\phi$ must be even, and we choose $g(\theta) = 2F_{2N+1}(\theta) + F_{2N+1}(0)$

Note that from the definition of $F_M(\theta)$, (17), the Fourier coefficients of $g$ can be computed, since

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F_M(\theta) e^{\mathrm{i}k\theta}\,\mathrm{d}\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sum_{j=-\infty}^{\infty} \phi\left(\frac{M}{2\pi}(\theta - 2\pi j)\right)\,\mathrm{d}\theta \tag{37}$$

$$= \frac{1}{M}\widehat{\phi}\left(\frac{k}{M}\right). \tag{38}$$

Therefore part (i) of Theorem 1 follows from (30) and (35), since if $\operatorname{supp}\widehat{\phi} \in (-2/\ell, 2/\ell)$ the $\ell$th cumulant of $Z_\phi(U)$ is zero (for $\ell \geq 3$). The mean and variance of $Z_\phi(U)$ equals $C_1^{\mathrm{U}(N)}(g)$ and $C_2^{\mathrm{U}(N)}(g)$ can be calculated from (28).

Similarly, from (32)–(31) and (34)–(35), we have that if $\operatorname{supp}\widehat{\phi} \in [-1/\ell, 1/\ell]$ the first $\ell$ cumulants of $Z_\phi(U)$ are Gaussian for $\mathrm{USp}(2N)$, $\mathrm{SO}(2N)$, and $\mathrm{SO}(2N+1)$, and this proves parts (ii) and (iii) of Theorem 1.

### 2.1. **Connections to moments of traces of matrices.**

From the cumulants one can obtain moments of traces of powers of $U$, first investigated by Diaconis and Shahshahani [1].

Expanding $Z_\phi(U)$ out as a Fourier series, we obtain

$$\mathbb{E}_{\mathrm{G}(N)}\left\{(Z_\phi)^m\right\} = \frac{1}{N^m}\sum_{n_1=-\infty}^{\infty}\cdots\sum_{n_m-\infty}^{\infty}\widehat{\phi}\left(\frac{n_1}{N}\right)\ldots\widehat{\phi}\left(\frac{n_m}{N}\right)\mathbb{E}_{\mathrm{G}(N)}\left\{\operatorname{Tr}U^{n_1}\ldots\operatorname{Tr}U^{n_m}\right\}. \tag{39}$$

Writing the moments in terms of cumulants, using (34) and (35), and comparing Fourier coefficients, we find that

**Corollary 1.1.** *Let $Z_j$ be independent standard normal random variables, and let*

$$\eta_j = \begin{cases} 1 & \text{if } j \text{ is even} \\ 0 & \text{if } j \text{ is odd} \end{cases} \tag{40}$$

*Let $a_j \in \{0,1,2,\dots\}$ for $j = 1,2,\dots$.*

- *If $\sum ja_j \le M-1$, then*

$$\mathbb{E}_{\mathrm{SO}(M)}\left\{\prod(\operatorname{Tr}U^j)^{a_j}\right\} = E\left\{\prod(\sqrt{j}Z_j + \eta_j)^{a_j}\right\}. \tag{41}$$

- *If $\sum ja_j \le M+1$, where $M$ is even, then*

$$\mathbb{E}_{\mathrm{USp}(M)}\left\{\prod(\operatorname{Tr}U^j)^{a_j}\right\} = E\left\{\prod(\sqrt{j}Z_j - \eta_j)^{a_j}\right\}. \tag{42}$$

These results were found by Diaconis and Shahshahani [1], though only for half the range of the parameters (and they dealt with the full orthogonal group, not the special orthogonal group). Recently Michael Stolz [10] has provided a further proof of this theorem, for the full range, using invariant theory (though again, he does not deal with the special orthogonal group).

Analogously, for the unitary group, one obtains

**Corollary 1.2.** *For $a_j, b_j \in \{0,1,2,\dots\}$, if*

$$\max\left(\sum_{j\ge1} ja_j \ , \ \sum_{j\ge1} jb_j\right) \le N \tag{43}$$

*then*

$$\mathbb{E}_{\mathrm{U}(N)}\left\{\prod_{j\ge1}\left(\operatorname{Tr}U^j\right)^{a_j}\left(\operatorname{Tr}U^{-j}\right)^{b_j}\right\} = \delta_{a,b}\prod_{j\ge1} j^{a_j}a_j! \tag{44}$$

*where $\delta_{a,b} = 1$ if $a_j = b_j$ for all $j$, and zero otherwise.*

This is exactly the result of Diaconis and Shahshahani, [1]. Indeed, in [6] the mock-Gaussian result for the unitary case was proved via this result, rather than evaluating the cumulants, as this was a more direct approach.

## 3. Number theory examples

3.1. **The Riemann zeta function: A unitary example.** Consider the non-trivial zeros of the Riemann zeta function, $\frac{1}{2} + \mathrm{i}\gamma$. The Riemann Hypothesis (which we do not assume) is the statement that $\gamma \in \mathbb{R}$ for all $\gamma$.

The counting function of Riemann zeros is

$$N(T) = \#\{\gamma \ : \ 0 \le \mathfrak{Re}(\gamma) \le T\} \tag{45}$$

$$= \overline{N}(T) + S(T) \tag{46}$$

where

$$\overline{N}(T) = 1 + \frac{1}{\pi}\mathfrak{Im}\log\left(\pi^{-\mathrm{i}T/2}\Gamma(\tfrac{1}{4} + \tfrac{1}{2}\mathrm{i}T)\right) \tag{47}$$

$$= \frac{T}{2\pi}\log\frac{T}{2\pi e} + \frac{7}{8} + \mathcal{O}(\frac{1}{T}) \tag{48}$$

and the error term is

$$S(T) = \frac{1}{\pi}\mathfrak{Im}\log\zeta(\tfrac{1}{2} + \mathrm{i}T) \tag{49}$$

$$= \mathcal{O}(\log T). \tag{50}$$

To motivate the study of the smooth counting function, we ask the question: What is the distribution of the number of zeros lying in an interval of size $h$ around height $T$? That is: What is the distribution of $N(t+h) - N(t)$ averaged over $T \le t \le 2T$?

Clearly, the mean is

$$\frac{1}{T}\int_T^{2T} \overline{N}(t+h)-\overline{N}(t)\ \mathrm{d}t \sim \frac{h}{2\pi}\log T \tag{51}$$

and Fujii [2] (among others) has shown that the centered moments are

$$\frac{1}{T}\int_T^{2T}\left(\frac{S(t+h)-S(t)}{\sigma}\right)^{2k}\mathrm{d}t = \frac{(2k)!}{2^k k!}+\mathcal{O}(\frac{1}{\sigma}) \tag{52}$$

where

$$\sigma^2 = \begin{cases} \frac{1}{\pi^2}\int_0^{h\log T}\frac{1-\cos t}{t}\ \mathrm{d}t & 0<h\ll 1\\ \frac{1}{\pi^2}(\log\log T - \log|\zeta(1+\mathrm{i}h)|) & 1\ll h\ll T\end{cases} \tag{53}$$

Thus if $h\log T\to\infty$, the moments converge to the Gaussian moments, and so the distribution is normal.

However, when $h\log T = \mathcal{O}(1)$, the main term, $\frac{(2k)!}{2^k k!}$, is of the same order as the error term $\mathcal{O}(1/\sigma)$. Therefore, we cannot conclude from (52) that the distribution in normal. In fact, the distribution is not normal, as it is discrete.

This motivates the study of the smooth counting function when the zeros are critically scaled,

$$N_\phi(t) = \sum_\gamma \phi\left(\frac{\log T}{2\pi}(\gamma - t)\right). \tag{54}$$

In [4] the moments of $N_\phi(t)$ were calculated, and the first few were found to be Gaussian.

For technical reasons we change the average. Instead of integrating over $t\in[T,2T]$ we define the average to be

$$\langle N_\phi\rangle_T = \int_{-\infty}^{\infty} N_\phi(t)\omega(\frac{t-T}{T})\frac{\mathrm{d}t}{T} \tag{55}$$

where $\int_{-\infty}^{\infty}\omega(x)\ \mathrm{d}x = 1$ and $\widehat{\omega}$ is compactly supported. The previous average would come from setting $\omega$ to be the indicator function of the interval $[0,1]$, but this is not allowed.

**Theorem 2** (Hughes and Rudnick [4])**.** *If* $\operatorname{supp}\widehat{\phi}\subset(-2/m,2/m)$ *then the first* $m$ *moments of* $N_\phi$ *converge as* $T\to\infty$ *to those of a Gaussian random variable with mean*

$$\int_{-\infty}^{\infty}\phi(x)\ \mathrm{d}x \tag{56}$$

*and variance*

$$\sigma_\phi^2 = \int_{-\infty}^{\infty}\min(|u|,1)\widehat{\phi}(u)^2\ \mathrm{d}u. \tag{57}$$

*Sketch of proof.* From a smooth version of Riemann's explicit formula we have that $N_\phi(\tau) = \overline{N_\phi}(\tau)+S_\phi(\tau)$, where

$$\overline{N_\phi}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi\left(\frac{\log T}{2\pi}(r-\tau)\right)\Omega(r)\mathrm{d}r + \phi\left(\frac{\log T}{2\pi}(\frac{\mathrm{i}}{2}-\tau)\right)+\phi\left(\frac{\log T}{2\pi}(-\frac{\mathrm{i}}{2}-\tau)\right) \tag{58}$$

with

$$\Omega(r) = \frac{1}{2}\Psi(\frac{1}{4}+\frac{1}{2}\mathrm{i}r)+\frac{1}{2}\Psi(\frac{1}{4}-\frac{1}{2}\mathrm{i}r) - \log\pi \tag{59}$$

and where

$$S_\phi(\tau) = -\frac{1}{\log T}\sum_{n\geq 2}\frac{\Lambda(n)}{\sqrt{n}}\widehat{\phi}(\frac{\log n}{\log T})2\cos(\tau\log n). \tag{60}$$

Asymptotic analysis gives that if $\widehat{\phi}\in C_c^\infty(\mathbb{R})$, then the mean value of $N_\phi$ is given by

$$\langle N_\phi\rangle_T = \langle\overline{N_\phi}\rangle_T \tag{61}$$

$$= \int_{-\infty}^{\infty}\phi(x)\mathrm{d}x + \mathcal{O}(\frac{1}{\log T}), \qquad T\to\infty\ . \tag{62}$$

Since

$$\lim_{T\to\infty} \left\langle \left(N_\phi - \langle N_\phi\rangle_T\right)^m \right\rangle_T = \lim_{T\to\infty} \langle (S_\phi)^m\rangle_T \tag{63}$$

it is therefore sufficient to show that the $m$th moment of $S_\phi$ is the same as that of a centered normal random variable with variance $\sigma_\phi^2$.

Multiplying out and integrating

$$\begin{aligned}\langle (S_\phi)^m\rangle_T = (\frac{-1}{\log T})^m \sum_{\epsilon_1,\dots,\epsilon_m=\pm1} \sum_{n_1,\dots,n_m} \prod_{j=1}^m \frac{\Lambda(n_j)}{\sqrt{n_j}} \widehat{\phi}(\frac{\log n_j}{\log T}) \\ \times \widehat{w}(\frac{T}{2\pi}\sum_{j=1}^m \epsilon_j \log n_j) e^{-\mathrm{i}T\sum_{j=1}^m \epsilon_j \log n_j}. \end{aligned}\tag{64}$$

Note $n_j \leq T^{2/m-\epsilon}$ since $\operatorname{supp}\widehat{\phi} \in (-2/m, 2/m)$.

Since $\widehat{w}$ has compact support, in order to get a nonzero contribution we need

$$|\sum_{j=1}^m \epsilon_j \log n_j| \ll \frac{1}{T}, \tag{65}$$

and thus $\sum \epsilon_j \log n_j = 0$.

Thus for $T \gg 1$, we find (taking into account that $\widehat{w}(0) = \int_{-\infty}^\infty w(x)\mathrm{d}x = 1$)

$$\langle (S_\phi)^m\rangle_T = (\frac{-1}{\log T})^m \sum_{\substack{n_1,\dots,n_m\geq 2\\ \epsilon_1,\dots,\epsilon_m=\pm1\\ \sum_{j=1}^m \epsilon_j\log n_j=0}} \prod_{j=1}^m \frac{\Lambda(n_j)}{\sqrt{n_j}} \widehat{\phi}(\frac{\log n_j}{\log T}). \tag{66}$$

The only terms which do not vanish as $T \to \infty$ are those where $m = 2k$ is even, and there is a partition $\{1, \dots, 2k\} = S \cup S'$ into disjoint subsets and a bijection $\sigma : S \to S'$ such that $n_j = n_{\sigma(j)}$ and $\epsilon_j = -\epsilon_{\sigma(j)}$. There are $k!\binom{2k}{k}$ such terms, and so

$$\left\langle (S_\phi)^{2k}\right\rangle_T \to \frac{(2k)!}{k!}\left(\frac{1}{\log^2 T}\sum_n \frac{\Lambda(n)^2}{n}\widehat{\phi}(\frac{\log n}{\log T})^2\right)^k \tag{67}$$

$$\to \frac{(2k)!}{k!}\left(\int_0^\infty u\widehat{\phi}(u)^2\mathrm{d}u\right)^k \tag{68}$$

by the Prime Number Theorem. □

This theorem compares perfectly with part (i) of Theorem 1.

3.2. **Real Dirichlet $L$–functions: A symplectic family.** Consider the zeros of quadratic $L$-functions, that is of $L$-functions of the form

$$L(s,\chi_d) = \sum_{n=1}^\infty \frac{\chi_d(n)}{n^s} \tag{69}$$

where $\chi_d(n) = \left(\frac{d}{n}\right)$ is the Kronecker symbol.

Rather than averaging over $t$, we will average the low-lying zeros of the $L$-function over characters, that is over $d \in D(X) := \{d\ :|d| \leq X\ ,\ \chi_d \text{ primitive}\}$.

From an explicit formula we can show that the smooth counting function equals

$$N_\phi(\chi_d) := \sum_{\gamma_d} \phi\left(\frac{\log X}{2\pi}\gamma_d\right) \tag{70}$$

$$= \int_{-\infty}^{\infty} \phi(x)\,\mathrm{d}x - \int_0^\infty \widehat{\phi}(u)\,\mathrm{d}u - \frac{2}{\log X}\sum_{\substack{n=1\\ n\neq\square}}^{\infty} \frac{\Lambda(n)}{n^{1/2}}\chi_d(n)\widehat{\phi}\left(\frac{\log n}{\log X}\right). \tag{71}$$

The mean of $N_\phi(\chi_d)$ is $\mu_\phi^{\mathrm{USp}} := \int_{-\infty}^\infty \phi(x)\,\mathrm{d}x - \int_0^1 \widehat{\phi}(u)\,\mathrm{d}u$, and so the centered moments are

$$\frac{1}{|D(X)|}\sum_{d\in D(X)}\left(N_\phi(\chi_d) - \mu_\phi^{\mathrm{USp}}\right)^m = \frac{1}{|D(X)|}\sum_{d\in D(X)}\left(-\frac{2}{\log X}\sum_{\substack{n=1\\ n\neq\square}}^{\infty}\frac{\Lambda(n)}{n^{1/2}}\chi_d(n)\widehat{\phi}\left(\frac{\log n}{\log X}\right)\right)^m \tag{72}$$

Expanding out the bracket, one can show that if $\operatorname{supp}\widehat{\phi} \in (-1/m, 1/m)$, then the only contribution comes from the terms where $n_1 \dots n_m = \square$ (in which case $\chi_d(n_1\dots n_m) = 1$ for all $d$). That is, we have

**Theorem 3.** *Let $D(X)$ be the set of primitive quadratic characters $\chi_d$ with $|d| \le X$. If* $\operatorname{supp}\widehat{\phi} \in [-1/m, 1/m]$ *then*

$$\frac{1}{|D(X)|}\sum_{d\in D(X)}\left(N_\phi(\chi_d) - \mu_\phi^{\mathrm{USp}}\right)^m \to \begin{cases} \frac{(2k)!}{2^k k!}\left(4\int_0^{1/2} u\widehat{\phi}(u)^2\,\mathrm{d}u\right)^k & \text{if } m = 2k \text{ is even} \\ 0 & \text{if } m \text{ is odd}\end{cases} \tag{73}$$

*where*

$$\mu_\phi^{\mathrm{USp}} = \int_{-\infty}^{\infty}\phi(x)\,\mathrm{d}x - \int_0^1 \widehat{\phi}(u)\,\mathrm{d}u. \tag{74}$$

This theorem agrees perfectly with part (ii) of Theorem 1. By the work in Section 1, this theorem implies the $n$-level densities of the zeros $L(s,\chi_d)$ are the same as the $n$-level densities of the symplectic group (within a restricted range), a result first derived by Mike Rubinstein [8], though this approach avoids the combinatorial sieving necessary there. Indeed, also by the work in Section 1, one can derive this theorem immediately from Rubinstein's result.

### 3.3. $L$–functions arising from cuspidal newforms: An orthogonal example.

Let $H_k^\star(N)$ be the set of all holomorphic cusp forms which are newforms of weight $k$ and level $N$.

Let the Fourier coefficients of an $f \in H_k^\star(N)$ be $a_f(n)$, and let $\lambda_f(n) = a_f(n)n^{-(k-1)/2}$. The $L$-function associated with $f$ is

$$L(s,f) = \sum_{n=1}^\infty \lambda_f(n)n^{-s}. \tag{75}$$

It satisfies the functional equation mapping $s \longrightarrow 1-s$ which has sign $\epsilon_f = \pm 1$.

Therefore $H_k^\star(N)$ splits into two disjoint subsets, $H_k^+(N) = \{f \in H_k^\star(N) : \epsilon_f = +1\}$ and $H_k^-(N) = \{f \in H_k^\star(N) : \epsilon_f = -1\}$.

For $\widehat{\phi} \in C_c^\infty(\mathbb{R})$, define the smooth counting function

$$N_\phi(f) = \sum_{\gamma_f}\phi\left(\frac{\log(k^2N)}{2\pi}\gamma_f\right). \tag{76}$$

Here, $\gamma_f$ runs through the non-trivial zeros of $L(s,f)$. We rescale the zeros by $\log(k^2N)$ as this is the order of the number of zeros with imaginary part less than a large absolute constant.

We define the average over either $H_k^+(N)$ or $H_k^-(N)$ by

$$\langle N_\phi(f)\rangle_\pm := \frac{1}{|H_k^\pm(N)|} \sum_{f\in H_k^\pm(N)} N_\phi(f). \tag{77}$$

We let $N \to \infty$ through the primes, with $k$ held fixed.

**Theorem 4** (Hughes and Miller [3])**.** *If* $\operatorname{supp}\widehat{\phi} \subseteq (-\frac{1}{m}, \frac{1}{m})$ *then the $m^{th}$ moment of $N_\phi(f)$, when averaged over the elements of either $H_k^+(N)$ or $H_k^-(N)$, converges to the $m^{th}$ moment of a normal distribution with mean*

$$\mu \;=\; \widehat{\phi}(0) + \frac{1}{2}\int_{-1}^{1} \widehat{\phi}(y)\,\mathrm{d}y \tag{78}$$

*and variance*

$$\sigma^2 \;=\; 2\int_{-1/2}^{1/2} |u|\widehat{\phi}(y)\,\mathrm{d}y. \tag{79}$$

This result is in complete agreement with part (iii) of Theorem 1, and also shows that the $n$-level densities do, as expected, agree with the $n$-level densities for the special orthogonal group. However, in this case we are able to go beyond the diagonal, and show that Gaussian behaviour ceases at the point predicted by random matrix theory.

**Theorem 5** (Hughes and Miller [3])**.** *Let $\mu_\pm = \langle N_\phi(f)\rangle_\pm$, and let $S(x) = \frac{\sin\pi x}{\pi x}$. For $n \geq 2$, let* $\operatorname{supp}(\widehat{\phi}) \subset (-\frac{2}{2n-1}, \frac{2}{2n-1})$*. Then as $N \to \infty$ through the primes, if $n = 2m$ is an even integer,*

$$\lim_{N\to\infty} \left\langle (N_\phi(f) - \mu_\pm)^{2m} \right\rangle_\pm$$
$$= \frac{(2m)!}{2^m m!}\left(2\int_{-1}^{1}\widehat{\phi}(x)^2|x|\,\mathrm{d}x\right)^m \mp 2^{2m-1}\left[\int_{-\infty}^{\infty}\phi(x)^{2m}S(2x)\,\mathrm{d}x - \frac{1}{2}\phi(0)^{2m}\right], \tag{80}$$

*and if $n = 2m+1$ is an odd integer, then*

$$\lim_{N\to\infty} \left\langle (N_\phi(f) - \mu_\pm)^{2m+1} \right\rangle_\pm = \pm 2^{2m}\left[\int_{-\infty}^{\infty}\phi(x)^{2m+1}S(2x)\mathrm{d}x - \frac{1}{2}\phi(0)^{2m+1}\right]. \tag{81}$$

In particular, as the Fourier Transform of $S(2x)$ is $\frac{1}{2}\mathbb{1}_{\{|x|\leq 1\}}$, the third centered moment is zero if $\operatorname{supp}\widehat{\phi} \subset (1/3, 1/3)$, but non-zero if the support exceeds this interval. These non-Gaussian results still agree with the random matrix results for $Z_\phi(U)$.

## Acknowledgements

Apart from the work with Steven Miller at the end, all the work presented in this talk is joint with Zeév Rudnick, and was carried out in 2001–2002 while funded by the EC TMR network "Mathematical Aspects of Quantum Chaos", EC-contract no HPRN-CT-2000-00103. This talk was given on 8th April 2004 at the *School on Recent Perspectives in Random Matrix Theory and Number Theory*, where my stay was funded by EPSRC Grant N09176

## References

[1] P. Diaconis and M. Shahshahani, "On the eigenvalues of random matrices", *J. Appl. Probab.* **31A** (1994) 49–62

[2] 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

[3] C.P. Hughes and S.J. Miller, "Low-lying zeros of $L$–functions with orthogonal symmetry" (in preparation)

[4] C.P. Hughes and Z. Rudnick, "Linear statistics for zeros of Riemann's zeta function", *C. R. Acad. Sci. Paris* Ser I **335** (2002) 667–670

[5] C.P. Hughes and Z. Rudnick, "Mock Gaussian behaviour for linear statistics of classical compact groups", *J. Phys. A.* **36** (2003) 2919–2932

[6] C.P. Hughes and Z. Rudnick, "Linear statistics for low-lying zeros of $L$–functions", *Quart. J. Math. Oxford* **54** (2003) 309–333
[7] N.M. Katz and P. Sarnak, *Random Matrices, Frobenius Eigenvalues, and Monodromy*, (AMS Colloquium Publications, 1999)
[8] M. Rubinstein, *Low-lying zeros of L–functions and random matrix theory*, Duke Math. J. **109** (2001) 147–181.
[9] A. Soshnikov, *Central limit theorem for local linear statistics in classical compact groups and related combinatorial identities*, Ann. Probab. **28** (2000) 1353–1370.
[10] M. Stolz, "On the Diaconis-Shahshahani method in random matrix theory", PhD Thesis, (University of Tuebingen, 2004)

*E-mail address*: hughes@aimath.org