First-Order Spacings of Random Matrix Eigenvalues

Rebecca C. Lehman*

December 25, 2001

Mathematics Department
Princeton University
Princeton, NJ 08544

Abstract

The eigenvalues of large random matrices are useful in many contexts, particularly statistical physics. For the Gaussian Orthogonal Ensemble, we present the known distribution of their local spacings, an analogue of the Central Limit Theorem for eigenvalues. We then investigate the local spacings of eigenvalues from other distributions: in particular the Uniform, Cauchy and Poisson, and show evidence that the distribution from the Gaussian may in fact be universal.

*E-mail: `rclehman@math.princeton.edu`

1 A Little Motivation

1.1 Classical Theory

In statistical mechanics, large and complicated systems can often be modeled as ensembles of random numbers. But random numbers are only meaningful if their probability distribution is known, and in physics it is sometimes difficult to guess the appropriate distribution.

Classical probability theory gives us the Weak Law of Large Numbers and the Central Limit Theorem, which essentially state that large sums of random numbers seem to behave in a certain way regardless of the particular distribution. However, not everything is independent of the distribution. Some things can only be proved for certain nice distributions. For instance, the first-order spacings of an ordered set are exponential for the uniform distribution, but the proof does not work for other distributions.

Theorem 1.1 (Weak Law of Large Numbers) *If $S_n = \frac{1}{n}\sum_1^n x_i$, where the x_i are independently randomly distributed over any distribution with mean 0 and variance normalized to 1, then $E(S_n) = 0$, and $\lim_{n\to\infty} E(S_n^2) = 0$. So for all $\epsilon \geq 0$ the probability that $|S_n| \geq \epsilon$ goes to 0.*

Proof:

$$E(S_n) = \frac{1}{n}\sum_1^n E(x_i) = E(x_i) = 0 \tag{1.1}$$

$$\begin{aligned}
E(S_n^2) &= \frac{1}{n^2}\sum_{i,j} E(x_i * x_j) \\
\frac{1}{n^2}\sum E(x_i^2) &= \frac{E(x_i^2)}{n} \\
\frac{1}{n^2}\sum_{i\neq j} E(x_i x_j) &= \frac{(n^2-n)E(x_i)^2}{n^2} \qquad (1.2)\\
& \qquad (1.3)
\end{aligned}$$

But $E(x_i^2) = 1$ and $E(x_i)^2 = 0$. So

$$E(S_n^2) = \frac{1}{n} + 0 = \frac{1}{n} \tag{1.4}$$

The probability of $|S_n| \geq \epsilon$ is the probability of $|S_n|^2 \geq \epsilon^2$. By Chebyshev's inequality,

$$\int_{|x|\geq\lambda} dP \quad \leq \quad \int_{|x|\geq\lambda} \left|\frac{x}{\lambda}\right| dP,$$

$$
\begin{aligned}
\int_{|x|\geq\lambda} \left|\frac{x}{\lambda}\right| dP &\leq \frac{1}{\lambda}\int |x| dP \\
Prob(S_n^2 \geq \epsilon^2) &\leq \frac{1}{\epsilon^2}\int \left|S_n^2\right| dP \\
&= \frac{1}{n*\epsilon^2} && (1.5) \\
& && (1.6)
\end{aligned}
$$

If we hold ϵ constant and allow n to get large, $Prob(S_n^2 \geq \epsilon^2)$ approaches 0.

Theorem 1.2 (Central Limit Theorem) *Let $\Omega_n = \frac{1}{\sqrt{n}}\sum_1^n x_j$, where the x_j are independently randomly distributed according to some probability distribution with mean 0 and variance 1, and suitably integrable. Then as $n \to \infty, \Omega_n$ has a Gaussian limiting probability distribution $e^{-2\pi\xi^2}$, appropriately normalized.*

Proof: We use Fourier analysis. If x and y have probability distributions μ and ν, then $x+y$ has distribution $\mu \star \nu$. So Ω_n has distribution $\mu \star \ldots \star \mu$, where μ is the distribution evaluated at $x\sqrt{n}$, and normalized appropriately by multiplying by $\sqrt{n}$. The Fourier transform converts convolution to multiplication, and $\hat{f}(cx) = \frac{1}{c}\hat{f}(\frac{x}{c})$, so the Fourier transform of the distribution is $(\hat{\mu})^n$ evaluated at $\frac{\xi}{\sqrt{n}}$, with the normalizations canceling out.

$$
\begin{aligned}
\hat{\mu}(\xi) &= \int_{-\infty}^{\infty} e^{2\pi i x*\xi} d\mu && (1.7) \\
\hat{\mu}(0) &= 1 && (1.8) \\
\frac{d\hat{\mu}}{d\xi}|_{\xi=0} &= 2i\pi E(x) = 0 && (1.9) \\
\frac{d^2\hat{\mu}}{d\xi^2}|_{\xi=0} &= 4\pi^2 E(x^2) = 4\pi^2 && (1.10) \\
& && (1.11)
\end{aligned}
$$

We do a Taylor expansion of $\hat{\mu}(t)$ about the origin, then substitute $t = \frac{\xi}{\sqrt{n}}$.

$$
\begin{aligned}
(\hat{\mu})^n(t) &= (1 - 2\pi^2 t^2 + \cdots)^n \\
(\hat{\mu})^n\left(\frac{\xi}{\sqrt{n}}\right) &= \left(1 - \frac{2\pi^2\xi^2}{n}\right)^n && (1.12) \\
& && (1.13)
\end{aligned}
$$

As $n \to \infty$, $(\hat{\mu})^n \to e^{-2\pi^2\xi^2}$, a Gaussian. Since Gaussians are only renormalized by Fourier transform, the original distribution $\mu \star \cdots \star \mu$ must also be Gaussian.

Theorem 1.3 (First Order Spacings) *Let $x_0, \cdots, x_n$ be random numbers drawn from the uniform distribution on [0,1), ordered by size. Then the probability distribution of the first-order spacings $x_j - x_{j-1}$, normalized by dividing by the average spacing $\frac{1}{n}$, approaches e^{-x}.*

Proof: First we note that the uniform distribution on [0,1) can be considered as the uniform distribution on $S^1 = \frac{R}{Z}$. Since the distribution is uniform and therefore invariant under translations, without loss of generality we can relabel the zero so that x_j is x_1 and $x_{j-1} = x_0 = 0$. Since x_1 is the first non-zero value, the probability that the first value x_1 is greater than a is the probability that all the x_j except $x_0 = 0$ are greater than a, which is $(1-a)^n$. The probability that nx_1 is between t and $t + \Delta$ is thus

$$\left(1 - \frac{t}{n}\right)^n - \left(1 - \frac{t+\Delta}{n}\right)^n \tag{1.14}$$

In the limit, as $n \to \infty$, this goes to $e^{-t} - e^{-(t+\Delta)}$

Dividing by Δ and taking the limit as Δ shrinks to 0, the probability distribution of x_1 is e^{-x}.

The proof as given applies only to the uniform distribution. However, it can be extended to other suitably integrable distributions by scaling local density to 1, using the cumulative distribution function.

1.2 Quantum mechanics and Random Matrix Theory

In quantum mechanics, many properties of a system in a given state can be represented by the eigenvalues of a symmetric or Hermitian linear operator; for instance the energy levels of a system in state Ψ are the eigenvalues solving the equation $H\Psi = E\Psi$. Unfortunately, in practice, for systems of any reasonable complexity the size of the matrix is usually impracticably large, if it is even known to be finite. Computing the actual entries of such a matrix is usually impossible. Hence it is often useful to treat most systems as random matrices of size N approaching infinity. Just as classical statistical mechanics treats positions and velocities as random variables in order to study their aggregate properties (for instance, the frequency of their collisions or their total energy), so in the quantum framework the analogous assumption would be to treat the linear operators defining the system as random matrices, and the individual properties as their eigenvalues.

It would therefore be useful to have analogues of the Law of Large Numbers and the Central Limit Theorem, revealing universal properties of the eigenvalues of unknown matrix distributions. Wigner's Semicircle is one such property. The local eigenvalue spacings, unlike the spacings of a classical random number distribution, are conjectured to be another, but as yet this is unproved.

2 Facts about Random Matrix Eigenvalues

2.1 Joint Probability Density

Theorem 2.1 (Joint Probability Density) *If an $n \times n$ random symmetric matrix X is drawn from a probability density $g(\lambda_1, \cdots, \lambda_n)$ with respect to the Lebesgue measure $\prod_{i \leq j} dX_{ij}$ on $R^{\frac{(n)(n-1)}{2}}$, where g is expressed in terms of the increasingly ordered eigenvalues $\lambda_1, \cdots, \lambda_n$, then the joint probability density function of the eigenvalues on the eigenvalue space $R^n|_{\lambda_1 < \cdots < \lambda_n}$ is*

$$C_n g(\lambda_1, \cdots, \lambda_n) \prod_{i<j} |\lambda_i - \lambda_i| \tag{2.1}$$

.

Proof: We construct the measure on eigenvalue space corresponding to the standard Lebesgue measure. With probability 1, $X = ODO^T$, where D is a diagonal matrix with distinct eigenvalues, and O is an orthogonal matrix: $O * O^T = I$, so $dO^T * O = -O^T * dO$. Differentiating X, and substituting $dM = dO^T * O$, we get

$$\begin{aligned}
dX &= dO * D * O^T + \\
&\quad O * dD * O^T + O * D * dO^T \\
dO * D * O^T &= O * O^T * dO * D * O^T \\
&= -O * dO^T * O * D * O^T \\
&= -O * dM * D * O^T \\
O * D * dO^T &= O * D * dO^T * O * O^T \\
&= O * D * dM * O^T
\end{aligned} \tag{2.2}$$

Therefore we find

$$dX = O(dD - dM * D + D * dM)O^T \tag{2.3}$$

Since the measure dX is invariant under orthogonal transformations, we get

$$\begin{aligned}\prod_{i\leq j} dX_{ij} &= \prod_{i\leq j}(dD - dM * D + D * dM)_{ij} \\ &= \prod_{i<j}(\lambda_j - \lambda_i)dM_{ij}\prod_{i=1}^{n} d\lambda_i \\ &= \prod_{i<j}|\lambda_i - \lambda_j|\prod_{i=1}^{n} d\lambda_i \prod_{i<j} dM_{ij} \qquad (2.4)\end{aligned}$$

We can integrate out the dM_{ij} to find that the induced measure is

$$C_n \prod_{i<j}|\lambda_i - \lambda_j|\prod_{i=1}^{n} d\lambda_i$$

.

It follows that if X has a density $g(\lambda_1, \cdots, \lambda_n)\prod_{i\leq j} dX_{ij}$, then the probability density in eigenvalue space is $C_n g(\lambda_1, \cdots, \lambda_n)\prod_{i<j}|\lambda_i - \lambda_j|\prod_{i=1}^{n} d\lambda_i$.

2.2 Semicircle Law

Theorem 2.2 (Wigner's Semicircle Law) *If X is an $n\times n$ symmetric matrix from some probability distribution μ such that the elements ξ_{ij}, up to the symmetric condition, are independently randomly distributed with mean 0, variance 1 and as $n \to \infty$, $C_k(n) = sup_{1\leq i\leq j\leq n} E(|\xi_{ij}|^k) = O(1)$, then the mean eigenvalue distribution of the matrix $\frac{X}{\sqrt{n}}$ tends to the semicircle distribution $\frac{1}{2\pi}\sqrt{4 - x^2}$ as $n \to \infty$.*

Mehta [5] proves this in his discussion of the Gaussian ensembles, relying on the joint probability distribution. Hiai and Petz [2] prove the theorem by a more conceptual combinatorial argument citing Voiculescu, which does not rely on the the messy joint probability distribution, and we will follow their argument here. The proof is by the method of moments: we can write the moments of the mean eigenvalue distribution in a combinatorial form, and show that the same form characterizes the sequence of moments of the semicircle.

The k^{th} moment of the mean eigenvalue distribution is

$$E(Tr(X^k)) = \frac{1}{n^{\frac{k}{2}+1}} \sum_{1\leq m_1, m_2, \cdots, m_k \leq n} E(\xi_{m_1 m_2} \cdots \xi_{m_k m_1}) \qquad (2.5)$$

Definition 2.3 *Non-Crossing Partitions We define a* non-crossing partition *P of a set S to be a partition into pairs $S_j = \{s_{j_1}, s_{j_2}\}$ such that $s_{j_1} < s_{k_1} < s_{j_2}$ iff $s_{j_1} < s_{k_2} < s_{j_2}$.*

Lemma 2.4 *The k^{th} moment $E(Tr(X^k))$ approaches 0 if k is odd, and the number of non-crossing partitions of $[k]$ if k is even, as $n \to \infty$.*

Proof: If any $\xi_{m_i m_{i+1}}$ appears without repetition, its expectation value is 0, so by independence the expectation value of the product containing it is 0. In particular, any term containing more than $\frac{k}{2}+1$ distinct terms contributes nothing to the sum.

There are at most $\binom{n}{l} l^k$ possible terms where l of the m_j are distinct, since each of the l distinct m_j can take one of n values, and each of the k factors is chosen from the l distinct values. Since

$$|E(\xi_{m_1 m_2} \cdots \xi_{m_k m_1}))| \leq E(|\xi_{m_1 m_2}|^k)^{(\frac{1}{k})} \cdots E(|\xi_{m_k m_1}|^k)^{(\frac{1}{k})} \leq C_k(n), \quad (2.6)$$

the sum over all such terms is $\binom{n}{l} \frac{l^k C_k(n)}{n^{\frac{k}{2}+1}}$ which vanishes as $n \to \infty$ if $l \leq \frac{k}{2}+1$.

So the only possible sum that doesn't go to zero is over the terms with $l = \frac{k}{2}+1$ distinct factors. If k is odd, the moment is 0. If k is even, we replace k by $2k'$.

Then we are interested in

$$\frac{1}{n^{k'+1}} \sum (E(\xi_{m_1 m_2} \cdots \xi_{m_{k'} m_1})) \quad (2.7)$$

where the sum is over sequences $\{m_j\}$ such that exactly $k'+1$ of the m_j are distinct, and every consecutive pair $\{m_j, m_{j+1}\}$ (considering j modulo $2k'$) appears at least twice.

By induction, to any non-crossing partition of $[2k']$ we can associate $n(n-1)\cdots(n-k')$ terms in this sum: if $k'=1$ there is a single partition $\{1,2\}$ to which we assign the $n(n-1)$ terms defined by $\xi_{m_1 m_2}\xi_{m_2 m_1}$. Any non-crossing partition of $[2k'+2]$ must contain some pair of form $s_{j_1}, s_{j_1}+1$. Removing that pair from the partition, we associate a partition of $[2k']$ by shifting downward. To each of the $n(n-1)\cdots(n-k')$ terms that correspond to this partition, we associate $(n-k')$ $2k'+2$ terms by inserting one of the $(n-k'-1)$ terms not yet used, in the s_{j_1} and $s_{j_1}+2$ positions.

Conversely, any sequence satisfying the conditions has each pair $\{m_j, m_{j+1}\}$ appearing exactly twice, and defines a non-crossing partition by $\{i,j\} \in P$ iff $\{m_i, m_{i+1}\} = \{m_j, m_{j+1}\}$. We prove this inductively: for $k'=1$ it is trivial. Assume it holds for $k'-1$. There must be some r such that m_r appears only once in the m sequence. Then $m_{r-1} = m_{r+1} \neq m_r$. Removing m_{r-1} and m_{r+1} we get a sequence with k' distinct elements. It defines some partition. Combine this non-crossing partition with the additional pair $\{r-1, r\}$. The result is still non-crossing, so the result holds.

So the sum

$$\frac{1}{n^{k'+1}} \sum (E(\xi_{m_1 m_2} \cdots \xi_{m_{k'} m_1}) = \frac{n(n-1)\cdots(n-k')}{n^{k'+1}} s_{k'} \tag{2.8}$$

where $s_{k'}$ is the number of non-crossing partitions of $[2k']$. As $n \to \infty$ the coefficient goes to 1, and our lemma is proved.

Lemma 2.5 *The number s_k of non-crossing partitions of $[2k]$ is the k^{th} Catalan number $c_k = \frac{1}{k+1}\binom{2k}{k}$.*

Any non-crossing partition pairs 1 with some even element $2m$, since any element s_{j_1} between 1 and its pair partner must also have s_{j_2} between 1 and its pair partner. The number of pair partitions containing $\{1, 2m\}$ is $s_{m-1}s_{k-m}$: it is determined by a non-crossing partition of the numbers inside $(1, 2m)$ and one of those outside $(1, 2m)$. This gives us the recursion relation $s_k = \sum_{i=0}^{k-1} s_i s_{k-1-i}$ for $k \geq 2$.

The function

$$g(x) = \frac{1}{2}(1 - \sqrt{1-4x}) = \sum_0^\infty \frac{x^{k+1}}{k+1}\binom{2k}{k} \tag{2.9}$$

is a generator function of the Catalan numbers. Since $g(x)$ satisfies the functional equation $g(x)^2 = g(x) - x$, its Taylor coefficients satisfy $c_n = \sum_0^{n-1} c_i c_{n-1-i}$ for $n \geq 2$, the same recursion as the s_n, with the same initialization: $c_0 = s_0 = c_1 = s_1 = 1$. So $c_n = s_n$ by induction.

Lemma 2.6 *The $(2k+1)^{th}$ moment of the semicircle distribution is 0, and the $2k^{th}$ is c_k.*

The odd moments are 0 because the semicircle distribution is an even function, so the integral vanishes by symmetry. For the evens, we integrate by parts to get

$$\begin{aligned} m_{2k} = \frac{1}{2\pi}\int_{-2}^{2} \sqrt{4-x^2}(x^{2k-1}(4-x^2))' dx &= 4(2k-1)(m_{2k-2}) - (2k-1)m_{2k} \\ &= \frac{2(2k-1)}{k-1} m_{2k-2} \end{aligned} \tag{2.10}$$

$$\tag{2.11}$$

The Catalan sequence clearly satisfies this recursion relation, and $m_0 = c_0 = 1$, so the lemma is proved.

Thus the moments of the semicircle $\frac{1}{2\pi}\sqrt{4-x^2}$ and the moments of the mean eigenvalue distribution are both equal to 0 if k is odd and the Catalan numbers if k is even, so since a function is completely determined by its moments, the eigenvalue distribution must be semicircular.

3 Gaussian

The Gaussian Orthogonal Ensemble, the probability distribution defined on real symmetric matrices by choosing x_{ij} from a Gaussian distribution $C * e^{-ax_{ij}^2}$, where C and a are appropriate normalization constants chosen such that the variance of the trace $E(Tr(X^2)) = 1$, is a particularly nice distribution both physically and mathematically. It is invariant under the orthogonal group, which makes it suitable for modeling physical spaces, and also makes the critical properties of the eigenvalues relatively easy to compute.

The joint distribution for the GOE is $C'e^{-a(\sum_{i=1}^n \lambda_i^2)} \prod_{i<j} |\lambda_i - \lambda_j|$ (see Theorem 2.1).

Theorem 3.1 (Characterizing the Gaussian) *The probability distributions on real-symmetric matrices which are independent of choice of basis (i.e. $P(Q^T * X * Q) = P(X)$ for Q orthogonal) and have all entries independently randomly distributed are precisely those of form $e^{-aTr(X)^2+bTr(X)+c}$ for some constants $a, b, c, a \geq 0$.*

Proof: We follow Mehta [5]. Let $P = \prod_{i\leq j} f_{ij}(X_{ij})$: X has entries independently distributed, and suppose P is invariant under the Orthogonal group. In particular, if $X' = O^T * X * O$, where

$$O = \begin{pmatrix} \cos\theta & \sin\theta & 0 & 0 & \cdots & 0 \\ -\sin\theta & \cos\theta & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & \ddots & & \end{pmatrix}$$

then

$$\begin{aligned} \frac{\partial X}{\partial \theta} &= \frac{\partial O^T}{\partial \theta} X' O + O^T X' \frac{\partial O}{\partial \theta} \\ &= \frac{\partial O^T}{\partial \theta} OXO^T O + O^T OXO^T \frac{\partial O}{\partial \theta} \\ &= AX + XA^T \end{aligned} \tag{3.1}$$

where

$$A = \frac{\partial O^T}{\partial \theta} O = \begin{bmatrix} 0 & -1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & \ddots & \cdots & 0 \end{bmatrix}$$

Since P is invariant under orthogonal transformations, the logarithmic derivative $\frac{\partial}{\partial \theta} \sum \log(f_{ij}(X_{ij}))$ should vanish:

$$\sum \frac{1}{f_{ij}} \frac{\partial f_{ij}}{\partial X_{ij}} \frac{\partial X_{ij}}{\partial \theta} = 0 \tag{3.2}$$

Substituting for $\frac{\partial X_{ij}}{\partial \theta}$ and expanding, we get

$$\begin{aligned} &\left(\frac{1}{f_{11}} \frac{\partial f_{11}}{\partial X_{11}} + \frac{1}{f_{22}} \frac{\partial f_{22}}{\partial X_{22}}\right)(2X_{12}) + \left(\frac{1}{f_{12}} \frac{\partial f_{12}}{\partial X_{12}})(X_{11} - X_{22}\right) \\ + &\sum_{k=3}^{n} \left(-\frac{1}{f_{1k}} \frac{\partial f_{1k}}{\partial X_{1k}} X_{2k} + \frac{1}{f_{2k}} \frac{\partial f_{2k}}{\partial X_{2k}} X_{1k}\right) \end{aligned} \tag{3.3}$$

Since each term of the sum depends on independent variables, each individually must be constant. Dividing by $X_{1k}X_{2k}$ we get

$$-\frac{1}{X_{1k} f_{1k}} \frac{\partial f_{1k}}{\partial X_{1k}} + \frac{1}{X_{2k} f_{2k}} \frac{\partial f_{2k}}{\partial X_{2k}} = \frac{C_k}{X_{1k} X_{2k}} \tag{3.4}$$

This equation is of form $f(x_1) + g(x_2) = h(x_1 x_2)$, which can only be solved by functions of form $a + b \ln x$. So $C_k = 0$ and

$$\frac{1}{X_{1k} f_{1k}} \frac{\partial f_{1k}}{\partial X_{1k}} = \frac{1}{X_{2k} f_{2k}} \frac{\partial f_{2k}}{\partial X_{2k}} = c \tag{3.5}$$

Integrating, we get $f_{1k}(X_{1k}) = e^{\frac{a}{2} X_{1k}^2}$. We can do the same for f_{jk} for any $j \neq k$. Since all invariants can be expressed in terms of traces of powers of X, and the off-diagonal elements appear as squares in the exponent, $P(X)$ can be expressed as an exponential in $Tr(X)$ and $Tr(X^2)$.

3.1 Local Spacings

The Gaussian is also mathematically nice in that it is possible calculate its local eigenvalue spacings. Wigner (of the Wigner Semicircle Law) surmised that the local nearest neighbor distribution, on small intervals normalized to have density 1, would be Axe^{-Bx^2}, with constants chosen so as to set the integral and the mean to 1. Remarkably, Mehta [4]and Gaudin [1] have proved that the actual spacings are extremely close to Wigner's surmise.

Theorem 3.2 (First-Order Spacings) *As $n \to \infty$, the probability density $P(S)$ of the distance between two consecutive eigenvalues of an $n \times n$ matrix from the Gaussian distribution (in the region of constant density, normalized by their mean) approaches $\frac{1}{4}\pi^2 \frac{d^2\Psi}{dt^2}$ where $\Psi(t)$ is (up to several constants) the Fredholm determinant corresponding to the linear convolution operator $f \to \int_{-t}^{t} K * f$,for the kernel $K = \frac{1}{2\pi}\left(\frac{\sin(\xi-\eta)}{\xi-\eta} + \frac{\sin(\xi+\eta)}{\xi+\eta}\right)$*

The proof of this theorem is very technical. Essentially the idea is to fix n and rewrite $P_n(S)$ by repeatedly converting from product to determinant form and using determinant operations. Eventually $P_n(S)$ is written as a finite Fredholm determinant whose kernels, fortunately, have a known limit as $n \to \infty$. We give only the barest sketch of the key points:

Since we are interested in what happens as $n \to \infty$ it suffices to consider even $n = 2m$. Fixing m, and writing $S = 2\theta$, the spacing distribution $P_m(S)$ for a matrix of size $2m$ is derived from the 2-point correlation function

$$P_m(-\theta,\theta) = \frac{(2m-2)!}{\mu_0} \int P(-\theta,\theta,\theta_1 \cdots \theta_{2m-2}) d\theta_1 \cdots d\theta_{2m-2} \qquad (3.6)$$

where

$$P(\theta_1 \cdots \theta_n) = e^{-(\theta_1^2+\cdots+\theta_n^2)} \prod_{i<j} |\theta_i - \theta_j| \qquad (3.7)$$

and the integral is taken over θ_i ordered in increasing size and with no θ_i in the interval $(-\theta, \theta)$.

Expanding the absolute value product and factoring out a 2θ, we find

$$P_m(-\theta,\theta) = \frac{(2m-2)!}{\mu_0} 2\theta e^{-2\theta^2} R(\theta) \qquad (3.8)$$

where

$$\mu_0 = (2m)!\ 2^{-\frac{2m(2m-1)}{4}} \prod_1^{\infty} \Gamma\left(\frac{k}{2}\right) \qquad (3.9)$$

$$R(\theta) = \int e^{-(\theta_1^2+\cdots+\theta_{2m-2}^2)} \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ -\theta & \theta & \theta_1 & \cdots & \theta_{2m-2} \\ \theta^2 & \theta^2 & \theta_1^2 & \cdots & \theta_{2m-2}^2 \\ \vdots & & \cdots & & \vdots \end{vmatrix} d\theta_1 \cdots d\theta_{2m-2} \qquad (3.10)$$

By integrating over the odd variables and applying column operations to the determinant, we get $R(\theta)$ as a symmetric integral over the even variables. This allows us to integrate independently over all the variables and divide by $(m-1)!$

Expanding the determinant by minors and changing variables several times, it is possible to derive

$$R(\theta) = -\left\{R_1(\theta) - \frac{1}{4\theta}\frac{d}{d\theta}R_1(\theta)\right\} \tag{3.11}$$

where if $\xi_{2i} = 2\int_\theta^\infty e^{-2y^2} y^{2i} dy$,

$$R(1) = \begin{vmatrix} 0 & 1 & \theta^2 & \cdots & \theta^{2m-2} \\ 1 & \xi_0 & \xi_2 & \cdots & \xi_{2m} \\ & & & \ddots & \\ \theta^{2m-2} & \xi_{2m-2} & \xi_{2m} & \cdots & \xi_{4m-4} \end{vmatrix}$$

Then $P(-\theta, \theta) = \frac{2m-2}{2\mu_0}\frac{d}{d\theta}\left(e^{-2\theta^2}R_1(\theta)\right)$ so the probability of an arbitrary spacing being at $(-\theta, \theta)$ is $2m(2m-1)P_m(-\theta, \theta)$.

R_1 can also be written out explicitly as an integral θ and dy, and if we define

$$\phi_m(\theta) = \int_\theta^\infty \cdots \int_\theta^\infty e^{-2(y_1^2+\cdots+y_m^2)} \prod_{i<j}(y_i^2 - y_j^2)^2 dy_1 \cdots dy_m, \tag{3.12}$$

then it is possible to substitute in for R_1 to get

$$P(S) = \frac{2m!2^{m-1}}{m!\mu_0}\frac{d^2}{d\theta^2}\phi_m(\theta) \tag{3.13}$$

Let Ψ_m be the normalization $\frac{\phi_m(\theta)}{\phi_m(0)}$.

We want the limit of Ψ_m when m increases but θ is normalized to the magnitude of the mean spacing: set $t = 2\theta\sqrt{2m}$.

Let $\Psi(t) = \lim_{m\to\infty}\Psi_m(t)$; then $P(S)$ is a constant times $\frac{d^2\Psi}{dt^2}$. The constant can be found to be $\frac{\pi^2}{4}$ by calculating the moments. But we still need to find Ψ.

Changing variables one more time, we renormalize t to τ and y_i to z_i by dividing by $\sqrt{2}$:

$$\begin{aligned} \phi_m(\theta) &= \int_\theta^\infty \cdots \int_\theta^\infty e^{-2(y_1^2+\cdots+y_m^2)} \prod_{i<j}(y_i^2 - y_j^2)^2 dy_1 \cdots dy_m && (3.14) \\ &= C\int_\tau^\infty \cdots \int_\tau^\infty e^{-2(z_1^2+\cdots+z_m^2)} \begin{vmatrix} 1 & z_1^2 & \cdots & z_1^{2m-2} \\ 1 & z_2^2 & \cdots & z_2^{2m-2} \\ \vdots & & \cdots & \vdots \\ 1 & z_m^2 & \cdots & z_m^{2m-2} \end{vmatrix} && (3.15) \end{aligned}$$

Rewriting the determinant with column operations, one can write it in terms of the Hermite polynomials and then the harmonic oscillator functions; finally the integration is brought inside the determinant to get

$$\Psi_m(t) = det\left[\delta_{ij} - \int_{-\tau}^{\tau} u_{2i}(z)u_{2j}(z)dz\right] \tag{3.16}$$

which is the Fredholm determinant for the kernel $K_m = \sum_0^{m-1} u_{2k}(x)u_{2k}(y)$. This kernel can be rewritten more suggestively as

$$\sqrt{m}\frac{1}{2}\left(\frac{u_{2m}(x)u_{2m-1}(y) - u_{2m}(y)u_{2m-1}(x)}{x-y} + \frac{u_{2m}(x)u_{2m-1}(y) + u_{2m}(y)u_{2m-1}(x)}{x+y}\right)$$

Finally changing variables and letting $m \to \infty$ we get $K = \frac{1}{2\pi}(\frac{sin(\xi-\eta)}{\xi-\eta}) + \frac{sin(\xi+\eta)}{\xi+\eta})$.

Despite its ugly infinite determinant form, Ψ is actually rapidly convergent and $P(S)$ can be computed to reasonable accuracy. It is approximated within 5% [1] by the Wigner surmise, which we will use in our plots because it is much faster to compute.

Note that the proof of the spacings distribution relied crucially on the normal distribution: with any other distribution it would not be possible to write $P(-\theta, \theta)$ as a determinant and integrate out the odd terms. Most distributions are thought to be completely intractable at this time. Nevertheless, the level spacing distribution is conjectured to be a robust universal property: in fact, our data suggests that it may be even more universal than the semicircle law. This is the conjecture which we investigate numerically.

4 Computations

It is conjectured that the known first-order distributions for the Gaussian are in fact universal, like the Central Limit Theorem and the Semicircle Law, so that they hold for any suitably normalized, sufficiently integrable function, or and even beyond. We investigate this conjecture by constructing large matrices using the software package Matlab [3] and plotting their first-order spacings against the Wigner surmise.

To construct matrix entries according to an arbitrary distribution μ, one can draw random numbers from the uniform distribution on [0, 1], and then invert the cumulative distribution function $\int_0^t d\mu$. In some cases this integral and its inverse can be computed in closed form. If not, it is necessary to construct a lookup table of the CDF values for small increments. Matlab contains efficient built-in random matrix generators for the uniform, Gaussian and various other distributions. We found that to generate a symmetric matrix, it was significantly more efficient to call the random matrix generator (generating n^2 entries) then and replace the lower half by symmetry, than to make $\frac{n(n-1)}{2}$ calls to the random number generator.

We look at matrix elements identically distributed according to the following continuous and discrete distributions:

Continuous:

Gaussian (to test the programs)

Uniform on [-1,1]: $P(t) = \frac{1}{2}$,

Symmetric Cauchy on $(-\infty, \infty)$: $P(t) = \frac{C}{1+t^2}$,

Discrete:

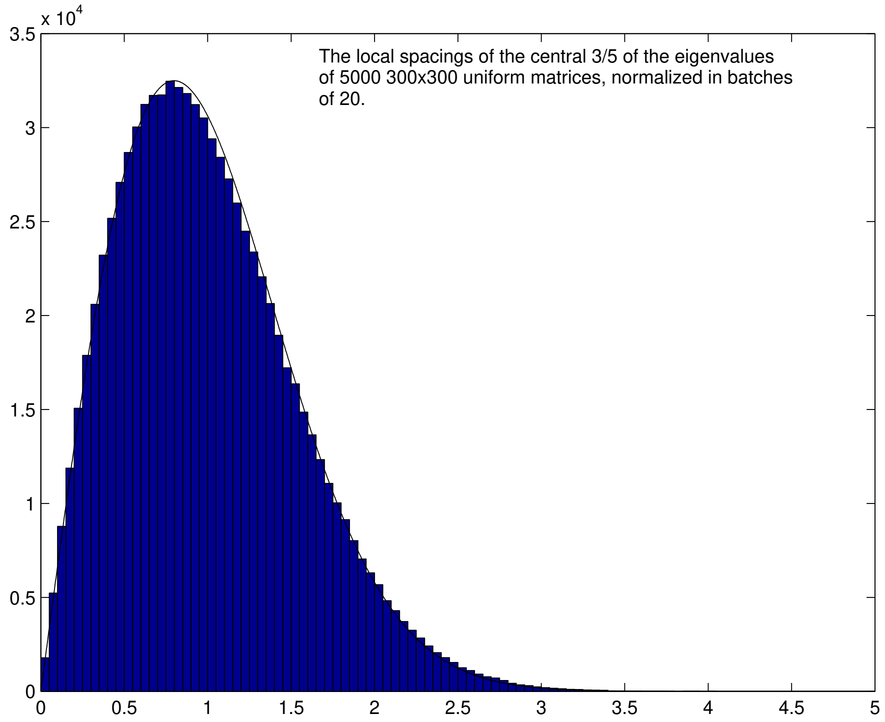
Sign: $P(1) = \frac{1}{2}, P(-1) = \frac{1}{2}$,

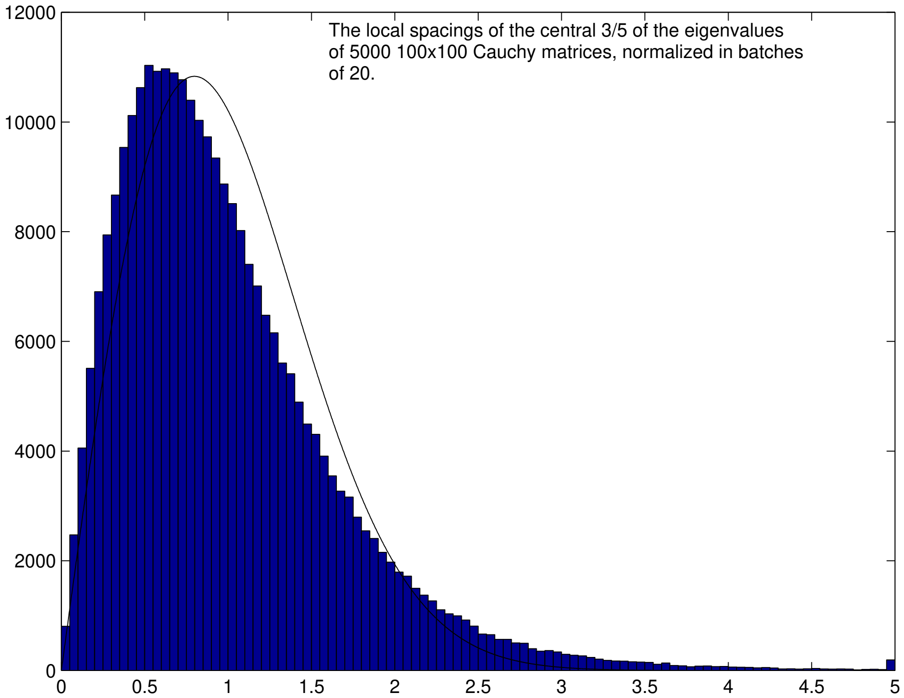
Poisson: $P(n) = e^{-\lambda}\frac{\lambda^n}{n!}$.

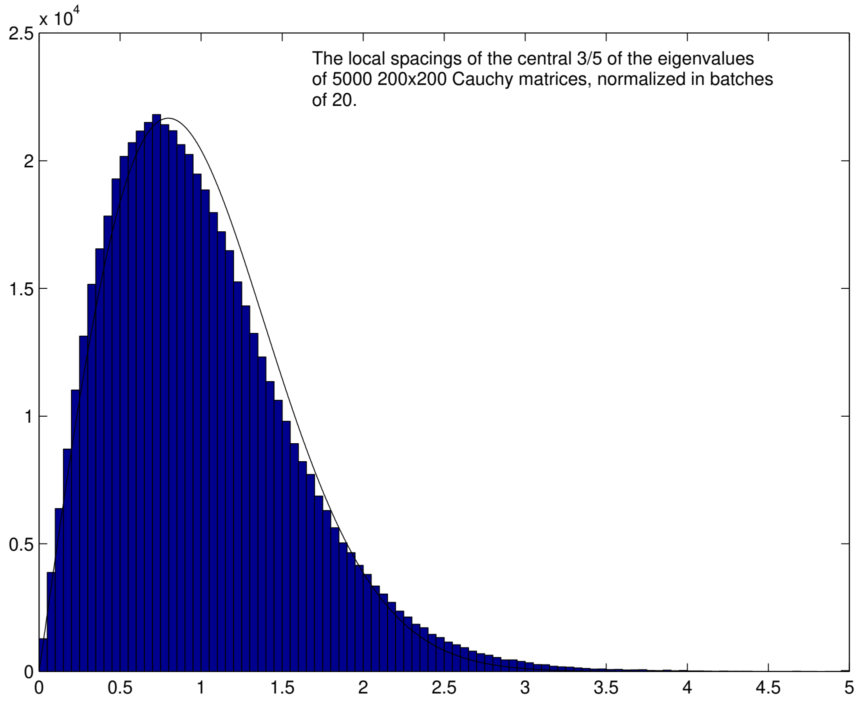
Note that the Poisson and Cauchy distributions do not have mean zero. The Poisson distribution has variable mean λ. The Cauchy distribution does not even have finite moments, and therefore does not obey a semicircle law at all.

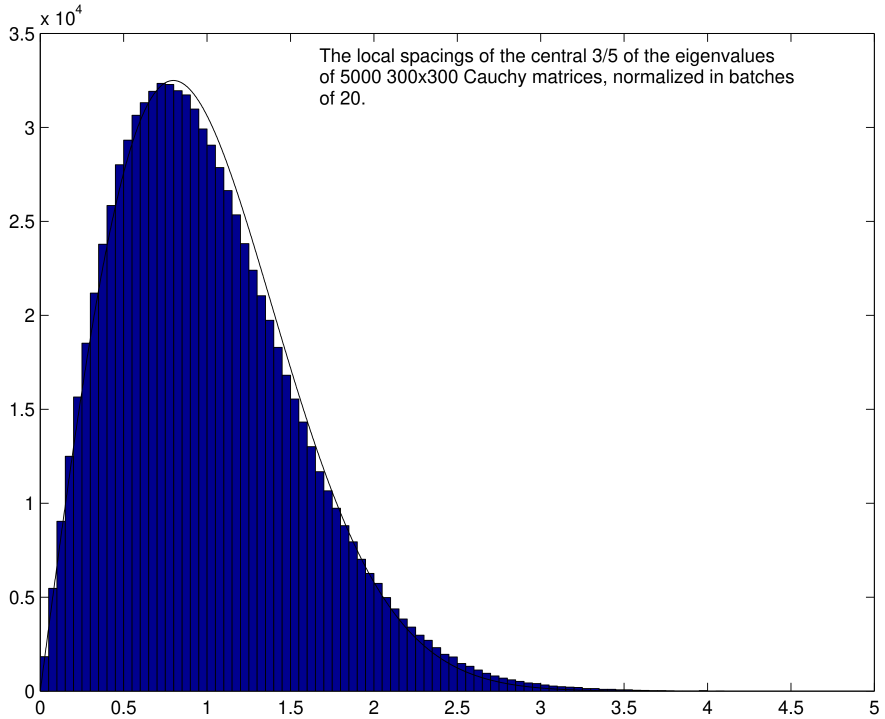
5 Results

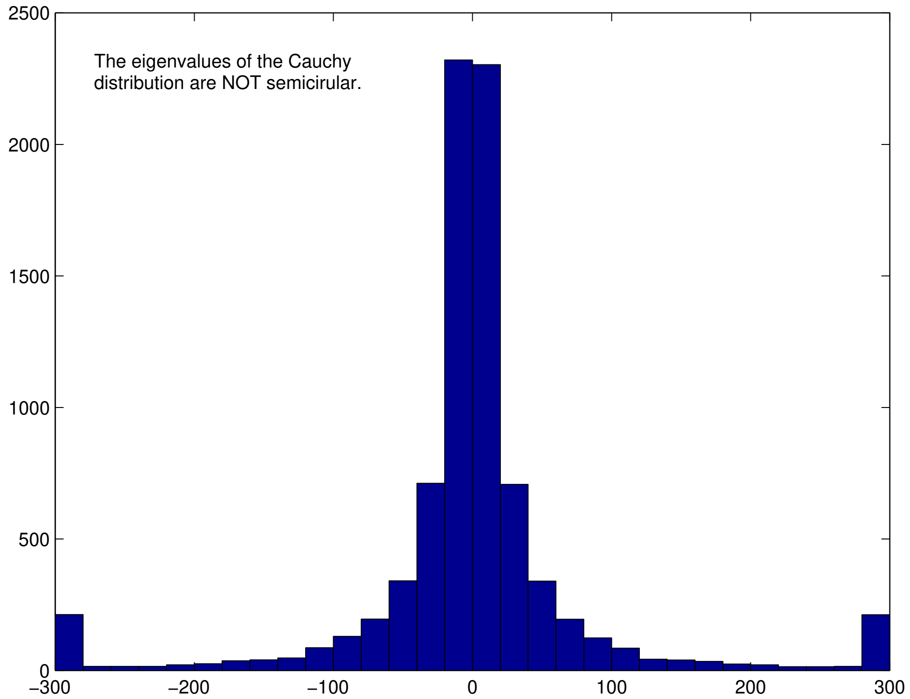
The program could process about 700 300x300 matrices per hour, varying slightly according to the distribution. Our data seems consistent with the conjecture for all the test distributions, for sets of up to 5000 matrices of size up to 300x300. The Cauchy distribution (which is not semicircular and does not provide a bound to the entries) converges more slowly to the surmise than the others, but as the size of the matrix grows, all do seem to be approaching the surmise.

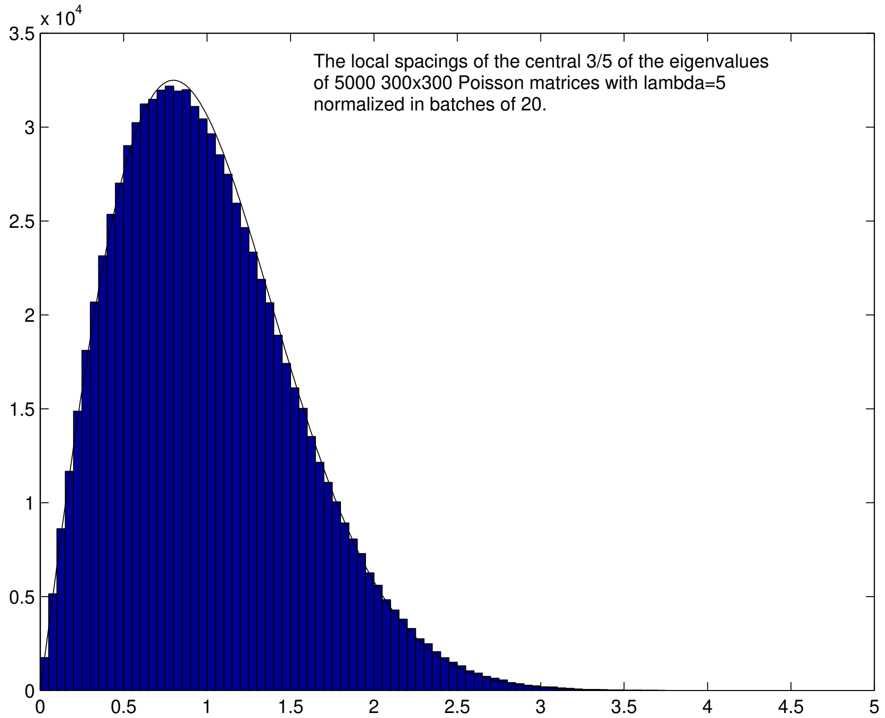

x 10^4
The local spacings of the central 3/5 of the eigenvalues of 5000 300x300 uniform matrices, normalized in batches of 20.

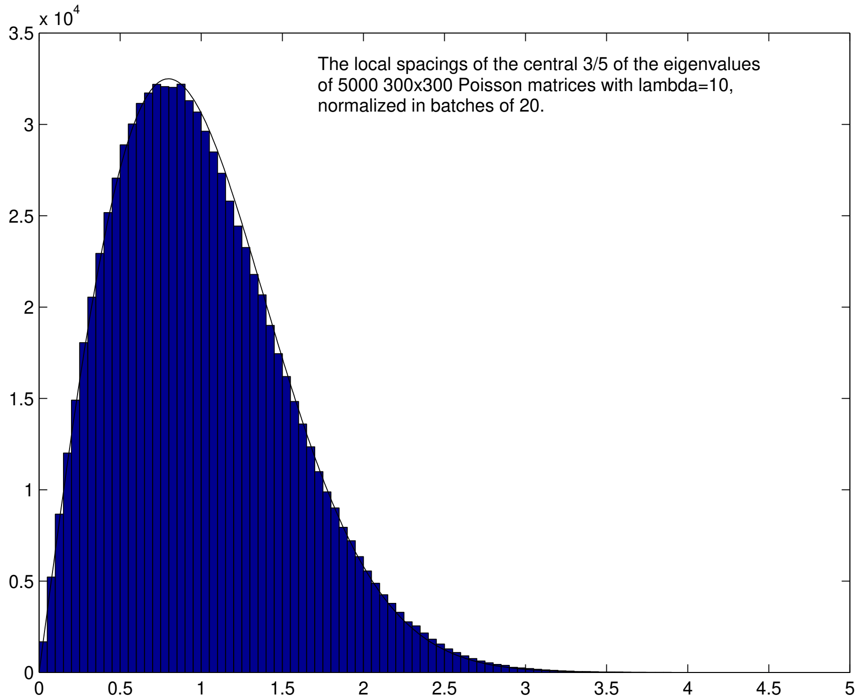

The local spacings of the central 3/5 of the eigenvalues of 5000 100x100 Cauchy matrices, normalized in batches of 20.

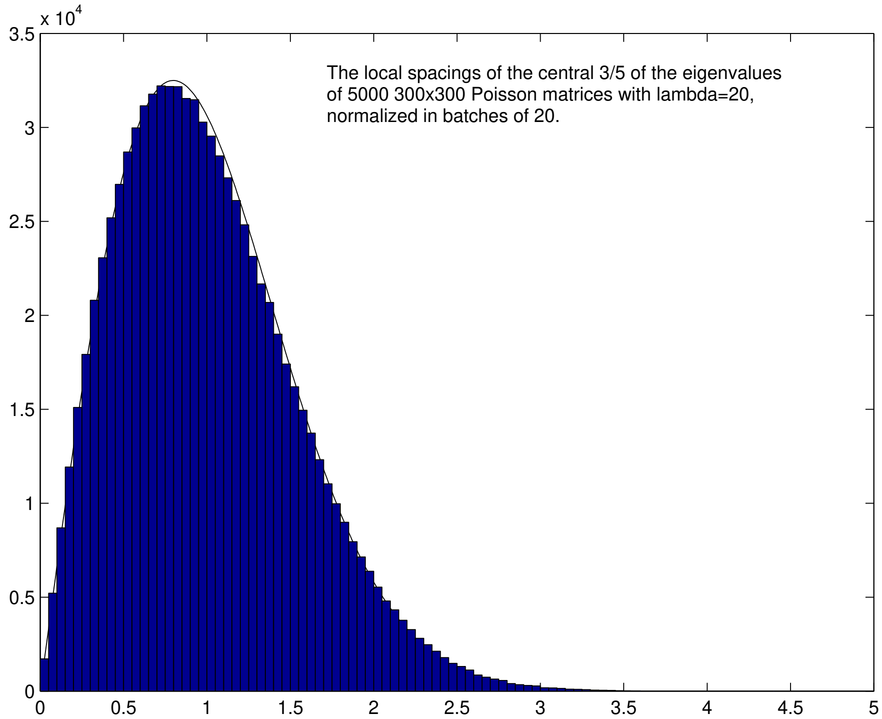

x 10^4
2.5
2
1.5
1
0.5
0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
The local spacings of the central 3/5 of the eigenvalues of 5000 200x200 Cauchy matrices, normalized in batches of 20.

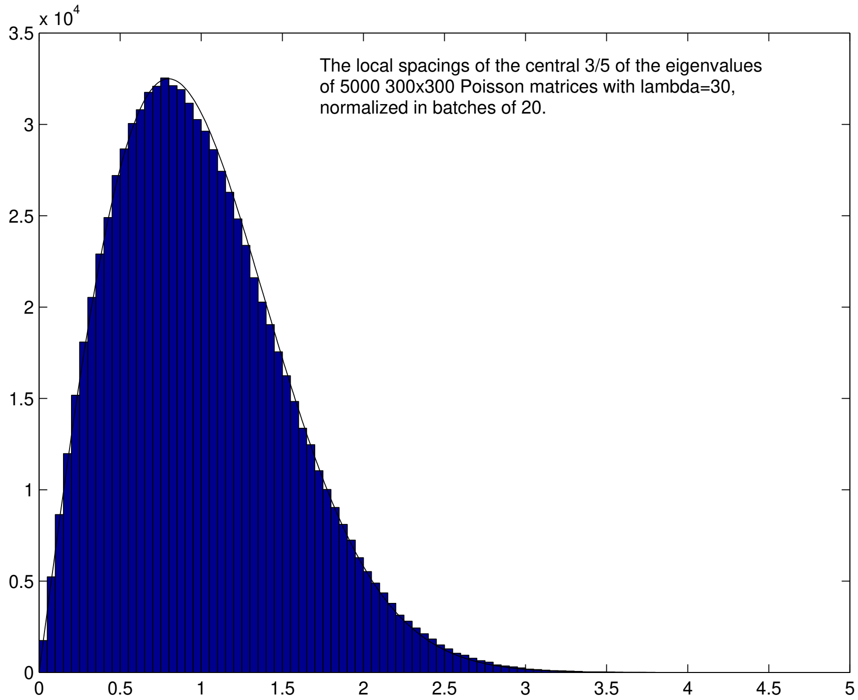

x 10^4
3.5
3
2.5
2
1.5
1
0.5
0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
The local spacings of the central 3/5 of the eigenvalues of 5000 300x300 Cauchy matrices, normalized in batches of 20.

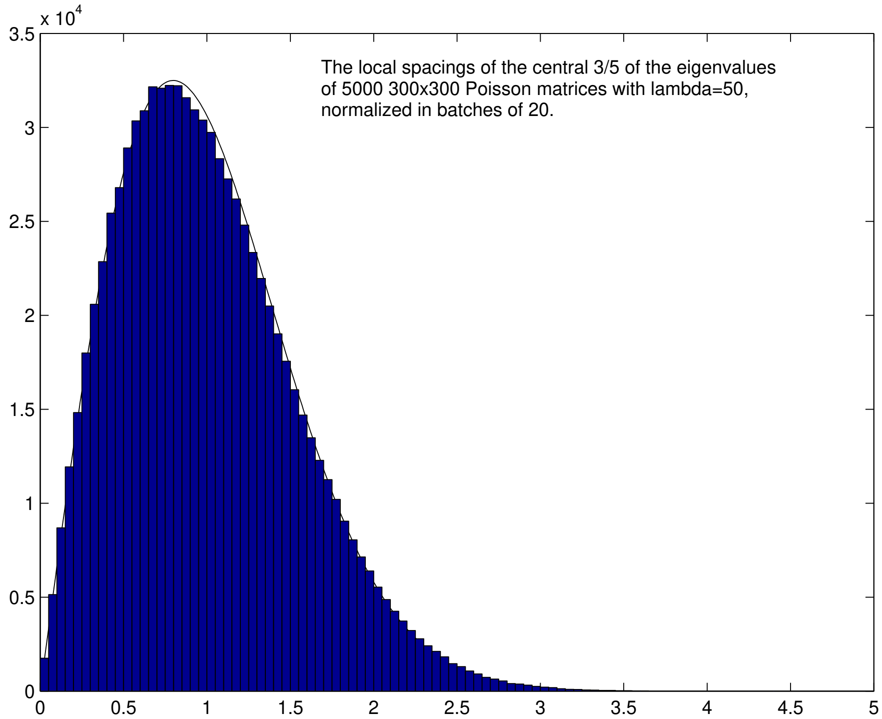

The eigenvalues of the Cauchy
distribution are NOT semicirular.
2500
2000
1500
1000
500
0
–300
–200
–100
0
100
200
300

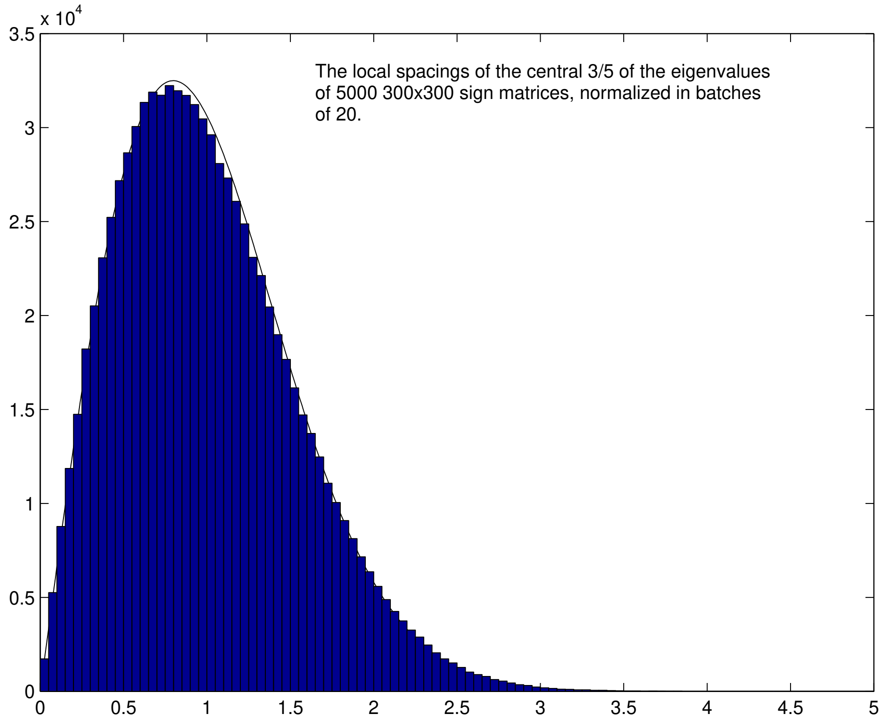

x 10^4
The local spacings of the central 3/5 of the eigenvalues
of 5000 300x300 Poisson matrices with lambda=5
normalized in batches of 20.
3.5
3
2.5
2
1.5
1
0.5
0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5


x 10^4
The local spacings of the central 3/5 of the eigenvalues
of 5000 300x300 Poisson matrices with lambda=10,
normalized in batches of 20.
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

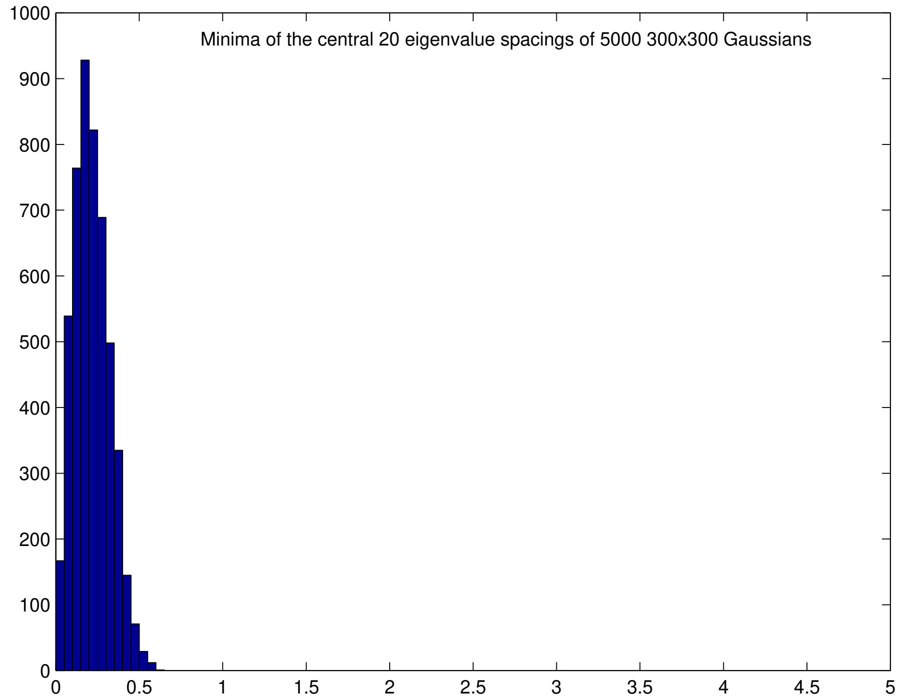

x 10^4
The local spacings of the central 3/5 of the eigenvalues
of 5000 300x300 Poisson matrices with lambda=20,
normalized in batches of 20.
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

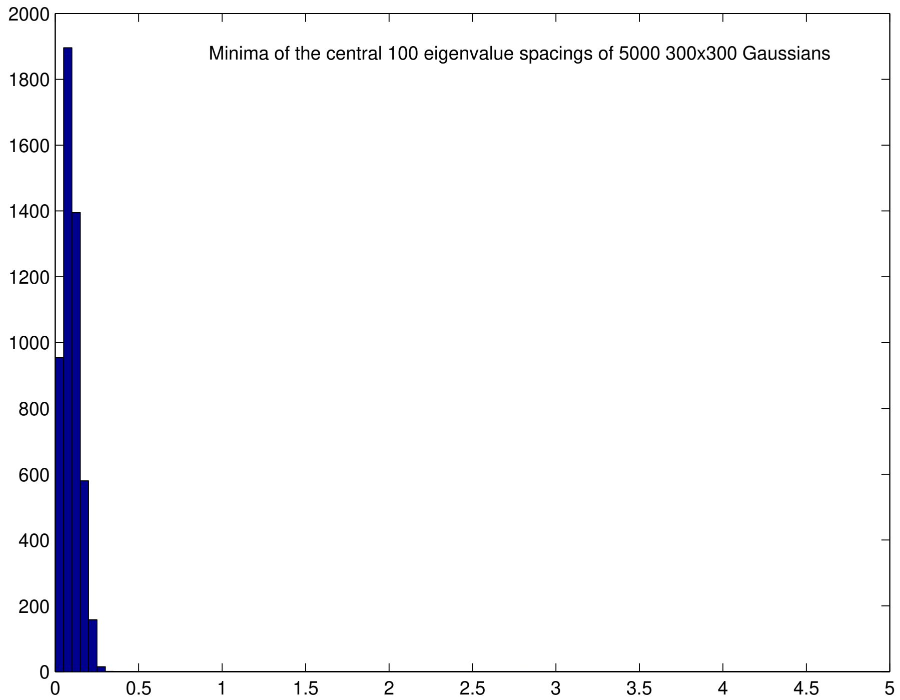

x 10^4
The local spacings of the central 3/5 of the eigenvalues
of 5000 300x300 Poisson matrices with lambda=30,
normalized in batches of 20.
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

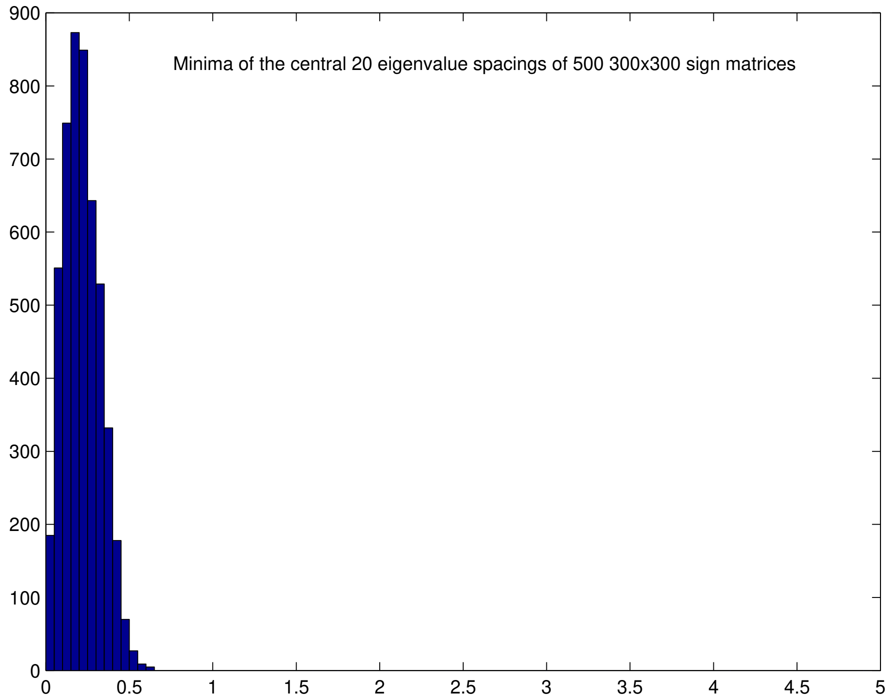

x 10^4
The local spacings of the central 3/5 of the eigenvalues
of 5000 300x300 Poisson matrices with lambda=50,
normalized in batches of 20.
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

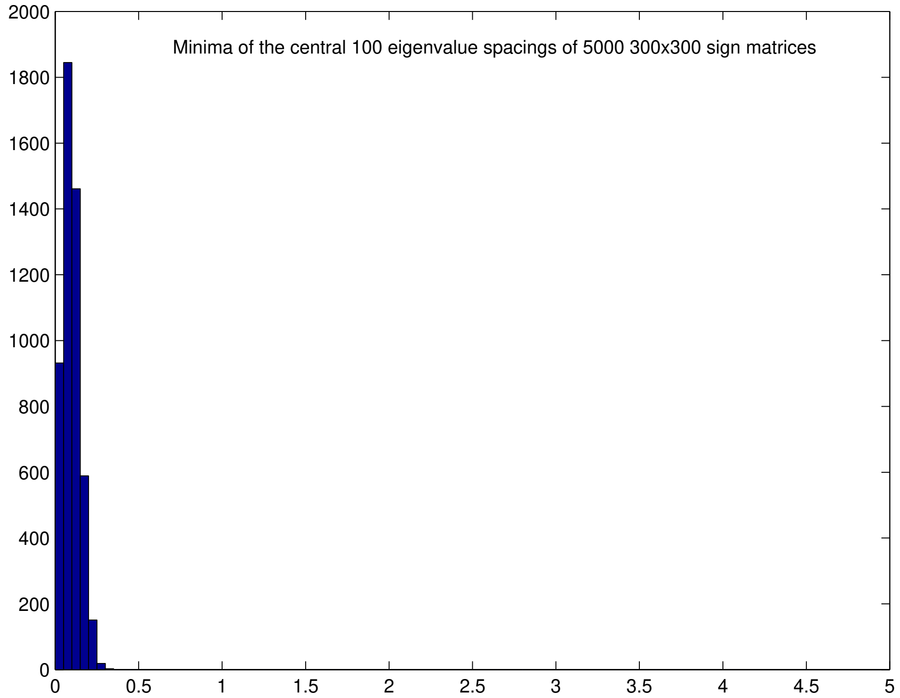


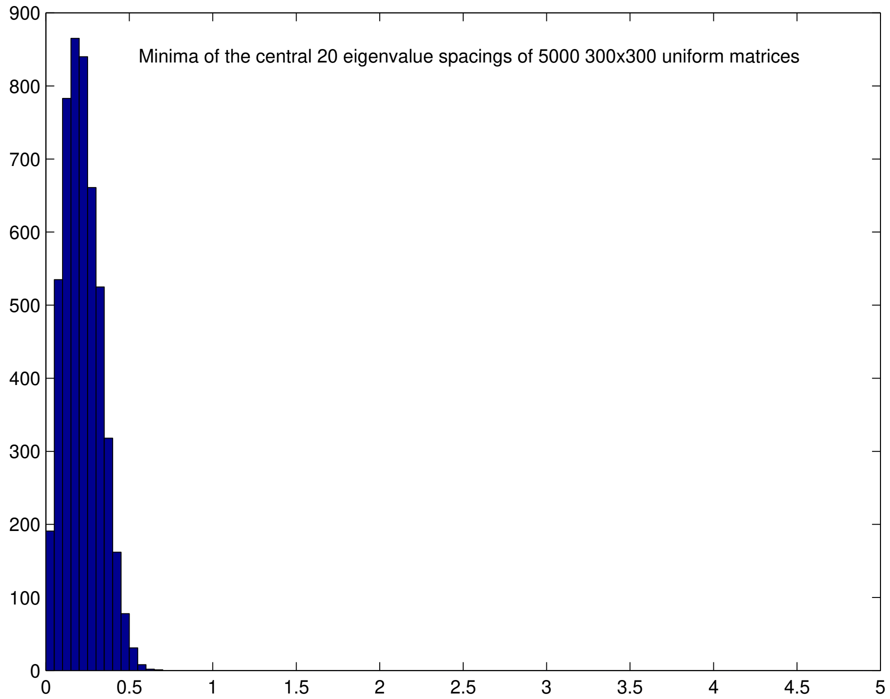
We then examined the minima from sets of normalized spacings. The minima from sets of 20 and 100 were predicted to look like e^{-x}, but our data appear like a compressed front end of the Wigner curve. More work must be done in this area.

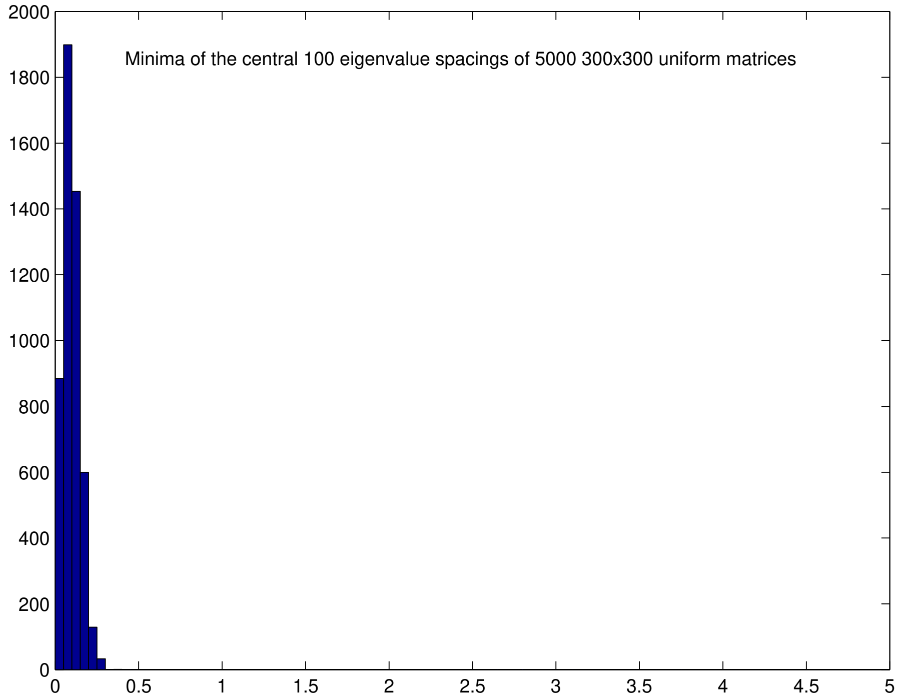

Minima of the central 20 eigenvalue spacings of 5000 300x300 Gaussians

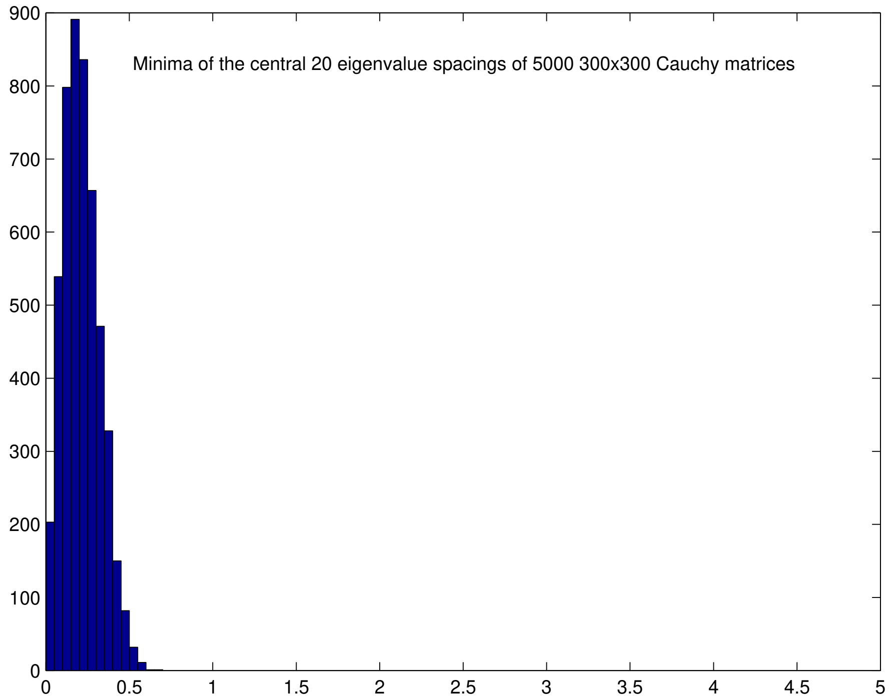

Minima of the central 100 eigenvalue spacings of 5000 300x300 Gaussians

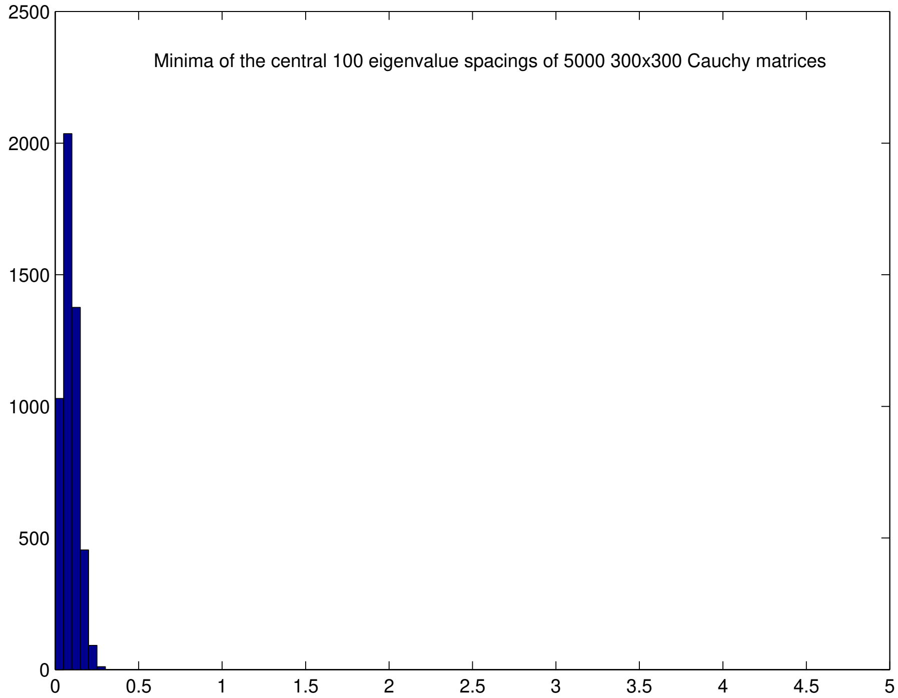

Minima of the central 20 eigenvalue spacings of 500 300x300 sign matrices


Minima of the central 100 eigenvalue spacings of 5000 300x300 sign matrices


Minima of the central 20 eigenvalue spacings of 5000 300x300 uniform matrices
900
800
700
600
500
400
300
200
100
0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5


Minima of the central 100 eigenvalue spacings of 5000 300x300 uniform matrices
2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5


Minima of the central 20 eigenvalue spacings of 5000 300x300 Cauchy matrices


Minima of the central 100 eigenvalue spacings of 5000 300x300 Cauchy matrices

6 Bibliography

[1] Michael Gaudin. "Sur la Loi Limite de L' Espacement des Valeurs Propres d'une Matrice Aleatoire", *Nuclear Physics*, vol 25 (1961), reprinted in [6].

[2] Furmio Hiai and Denes Petz. *The Semicircle Law, Free Random Variables, and Entropy.* American Mathematical Society, Providence, RI, 2000. Mathematical Surveys and Monographs, 77.

[3] Matlab. The Mathworks, Natick, Mass.

[4] M. L. Mehta. "On the Statistical Properties of the Level-Spacing", *Nuclear Physics*, vol 18, 1960. Reprinted in [6].

[5] M. L. Mehta. *Random Matrices.* Academic Press, San Deigo, CA, 1991.

[6] Charles Porter. *Statistical Theories of Spectra: Fluctuations.* Academic Press, New York and London, 1965.

[7] Graig A. Tracy and Harold Widom. "Introduction to Random Matrices", *Geometric and Quantum Aspects of Integral Systems.* Springer, Berlin, 1993. Lecture Notes in Physics, vol 424.