Distribution of Eigenvalue Spacings for Band-Diagonal Matrices

Nathan Miller[1]
Mathematics Department
Princeton University
Princeton, NJ 08544

May 15, 2003

[1]E-mail: njmiller@math.princeton.edu

Abstract

The purpose of this paper is to look at the nearest neighbor spacings of eigenvalues for band-diagonal matrices. We study the cases when the matrix elements are determined using a Cauchy distribution, a Laplace (double exponential) distribution, and a Uniform distribution. When a band of radius 1, i.e. a diagonal matrix, is used, the spacings of eigenvalues should look very differently than when the spacings of eigenvalues for a band of radius N (of an $N \times N$ matrix), i.e. just a random symmetric $N \times N$ matrix, is used.

Chapter 1

Theory

1.1 Introduction

To show what the theory of random matrices can do, it is best to see what it is used for. In nuclear and quantum physics a system can be described by the eigenvalue problem of

$$H\psi_n = E_n\psi_n \tag{1.1}$$

where H is the Hamiltonian operator, the ψ_n's are the eigenfunctions, and the E_n's are the eigenvalues. E_n usually corresponds to the energy of state n, and ψ_n^2 is the probability distribution function of state n.

In the case of the nucleus of an atom, there are two problems. First, the Hamiltonian is not fully known, and second, it would be too complicated to describe even if it was known. The solution to these problems is that we make statistical hypotheses on H, compatible with the general symmetry principles. Taking a complete set of functions as a basis, H can be represented as a matrix. The elements of this matrix are random variables whose distribution is restricted only by the symmetry properties imposed on the ensemble of operators. The problem is to get information of the behavior of its eigenvalues.

1.2 The Problem

The experimental content of this paper is to look at the distribution of eigenvalues for symmetric band-diagonal matrices. An $N \times N$ matrix A has a band of radius k if $A_{ij} = 0$ when $i - j \geq k$. If $i - j < k$ then A_{ij} is randomly pulled from a probability distribution. In all cases $A_{ij} = A_{ij}$. A matrix with a band of radius (size) 1 is just a diagonal matrix, whereas a matrix with a band of radius (size) N is a random symmetric matrix.

For a diagonal matrix, the eigenvalues are just the entries along the diagonal. They are the only non-zero entries of the matrix. The probability distrubution that the non-zero entries are pulled from will have a direct impact on the distribution of eigenvalues. A Cauchy probability distribution will have a different distribution of eigenvalues than a Gaussian distribution.

For a random symmetric matrix, the distribution of eigenvalues does not depend on the distribution used to generate the elements.

Conjecture 1.2.1. *Let H be an $N \times N$ real symmetric matrix, where off-diagonal elements H_{ij}, for $i \leq j$, are independent identically distributed (iid) random variables with mean μ and variance $\sigma^2 > 0$. When $i > j$, $H_{ij} = H_{ji}$. Then in the limit of large N the statistical properties of n eigenvalues of H become independent of the probability density of the H_{ij}. The joint probability density of arbitrarily chosen n eigenvalues of H tends, for every finite n, to the n-point correlation function of the Gaussian Orthogonal Ensemble (GOE).*

Rebecca Lehman and Yi-Kai Liu have experimentally investigated this conjecture in [Le] and [Liu], and their results seem to agree with what is being proposed.

The problem being investigated in this paper is how the distribution of the spacings of the eigenvalues depends on the band radius as the radius is increased. As the band radius is increased the spacings should tend to the GOE. It has been conjectured that for an $N \times N$ matrix, a band radius of $\sqrt{N}$ is when the distribution starts looking like the GOE and less like the distribution associated with the band radius of one. Liu has experimentally investiagted this in [Liu] using a Gaussian distribution. This paper will attempt to see whether this is true using the Cauchy, Laplace (Double Exponential), and Uniform distributions as probability distributions for the elements of the matrix.

The rest of this paper will prove some known results about the GOE and the distribution of eigenvalues of random symmetric matrices.

1.3 The Semicircle Rule

Lets look at a family of symmetric $N \times N$ random matrices, with its elements chosen from a probability distribution D. If the distribution of eigenvalues is normalized so all the eigenvalues lie in the interval $[-1, 1]$, then for certain distributions, D, the distribution of eigenvalues, $P(x)$, converges to the "semicircle". The "semicircle" is

$$P(x) = \frac{2}{\pi}\sqrt{1-x^2} \tag{1.2}$$

That gives the following theorem

Theorem 1.3.1 (Semicircle Law). *If D is a probability distribution satisfying the following:*

1. $E(x) = 0$*, i.e. mean is* 0
2. $E(x^2) = 1$*, i.e. variance is* 1
3. *for all* $k \geq 3$, $E(x^k)$ *is finite*

then the following is true

Construct a family of real $N \times N$ symmetric matrices. If A_{ij} represents the element occupying the i^{th} row and j^{th} column, then choose A_{ij} with $i \leq j$ independently from D and $A_{ji} = A_{ij}$.

Let $\frac{\lambda_i}{2\sqrt{N}}$ be the normalized eigenvalues of A, and define the distribution as

$$\mu_{A,N}(x) = \frac{1}{N}\sum_{i=1}^{N}\delta\left(x - \frac{\lambda_i}{2\sqrt{N}}\right) \tag{1.3}$$

As $N \to \infty$

$$\mu_{A,N}(x) \to P(x) = \frac{2}{\pi}\sqrt{1-x^2} \tag{1.4}$$

Proof: The proof of this theorem will come from the "method of moments". We can define the k^{th} moment of the distribution as

$$M_{A,N}(k) = \int_{-\infty}^{\infty} x^k \mu_{A,N}(x) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\lambda_i(A)}{2\sqrt{N}}\right)^k \tag{1.5}$$

We know that $Tr(A^k) = \sum \lambda_i(A)^k$ so we get

$$M_{A,N}(k) = \frac{1}{2^k N^{1+\frac{k}{2}}} Tr(A^K) \tag{1.6}$$

We can use this result to calculate the expected values of the moments for $\mu_{A,N}$. The moments determine the distribution so we can compare the expected moments of $\mu_{A,N}$ with the moments of the semicirlce and if they match then they are the same distribution.

The first several expected moments of $\mu_{A.N}$ are

$$E(M_{A,N}(0)) = E(\frac{1}{N}\sum_{i=1}^{N} 1) = E(1) = 1 \tag{1.7}$$

$$\begin{aligned} E(M_{A,N}(1)) &= \frac{1}{2N^{3/2}} E(Tr(A)) \\ &= \frac{1}{2N^{3/2}} E(\sum_{i=1}^{N} A_{ii}) \\ &= \frac{1}{2N^{3/2}} \sum_{i=1}^{N} E(A_{ii}) = 0 \end{aligned} \tag{1.8}$$

Because the mean of the distribution is 0, $E(A_{ii}) = 0$. The expected value of the zeroth moment is 1, and the first moment is 0. For the second moment we will need to use the following

$$(A^2)_{ii} = \sum_{j=1}^{N} A_{ij}A_{ji} = \sum_{j=1}^{N} (A_{ij})^2 \tag{1.9}$$

In that we used the fact that the matrix is symmetric, i.e. $A_{ij} = A_{ji}$.

$$
\begin{aligned}
E(M_{A,N}(2)) &= E\left(\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\lambda_i}{2\sqrt{N}}\right)^2\right) = \frac{1}{4N^2}E(Tr(A^2)) \\
&= \frac{1}{4N^2}E\left(\sum_{i=1}^{N}(A^2)_{ii}\right) \\
&= \frac{1}{4N^2}E\left(\sum_{i=1}^{N}\sum_{j=1}^{N}(A_{ij})^2\right) \\
&= \frac{1}{4N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}E((A_{ij})^2) \\
&= \frac{1}{4N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}1 = \frac{N^2}{4N^2} = \frac{1}{4} \qquad (1.10)
\end{aligned}
$$

$E(A_{ij}^2) = 1$ because the variance of the distribution used to choose the elements is 1. We have only calculated the first three moments of the distribution, but this can be extended to include all other higher moments. In [Le], Lehman shows that all the odd moments are 0 and that the even moments $E(M'_{A,N}(2k)) = c_k$, where c_k are the catalan numbers. $c_k = 2^k\frac{(2k-1)!!}{(k+1)!}$. However, she only normalized her eigenvalues by $\sqrt{N}$, not $2\sqrt{N}$ as done here and in [Liu]. Therefore the expected moments of our distribution will be

$$
\begin{aligned}
E(M_{A,N}(2k)) &= E(\frac{1}{2^{2k}}M'_{A,N}(2k)) = \frac{1}{2^{2k}}E(M'_{A,N}(2k) \\
&= \frac{1}{2^{2k}}c_k \qquad (1.11)
\end{aligned}
$$

Now we will calculate the moments of the semicircle and compare them to the moments of the calculated distribution. We will let $C(n)$ be the n^{th} moment of the semicircle

$$
C(n) = \int_{-1}^{1}\frac{2}{\pi}x^n\sqrt{1-x^2}dx \qquad (1.12)
$$

If we set $x = \sin\theta$, we then get

$$C(n) = \int_{-\pi/2}^{\pi/2} \frac{2}{\pi} \sin^n \theta \cos^2 \theta d\theta \tag{1.13}$$

If n is odd the integrand is an odd function and the integral vanishes, so $C(n) = 0$ for n odd. For n even we can write n as $2k$ and we get

$$\begin{aligned} C(2k) &= \int_{-\pi/2}^{\pi/2} \frac{2}{\pi} \sin^{2k} \theta (1 - \sin^2 \theta) d\theta \\ &= \int_{\pi/2}^{\pi/2} \frac{2}{\pi} \sin^{2k} \theta d\theta - \int_{-pi/2}^{\pi/2} \frac{2}{\pi} \sin^{2k+2} \theta d\theta \end{aligned} \tag{1.14}$$

From a table of integrals, in [CRC] page 399, we see

$$\int_0^{\pi/2} \sin^{2k} \theta d\theta = \frac{(2k-1)!!}{2k!!} \frac{\pi}{2} \tag{1.15}$$

Since $\sin^{2k} \theta$ is an even function $\int_{-\pi/2}^{\pi/2} = 2 * \int_0^{\pi/2}$ and we get

$$\begin{aligned} C(2k) &= 2\frac{(2k-1)!!}{2k!!} - 2\frac{(2k+1)!!}{(2k+2)!!} \\ &= \frac{2*(2k-1)!!}{2k!!} \left(1 - \frac{2k+1}{2k+1}\right) \\ &= \frac{2*(2k-1)!!}{2*(k+1)(2k)!!} \\ &= \frac{(2k-1)!!}{(k+1)2^k(k!)} \\ &= \frac{1}{2^{2k}} 2^k \frac{(2k-1)!!}{(k+1)!} \\ &= \frac{1}{2^{2k}} c_k \end{aligned} \tag{1.16}$$

Therefore the moments of the semicircle agree with what we calculated and with Lehman's calculation in [Le]. We have $E(M_{A,N}(0)) = C(0) = 1$, $E(M_{A,N}(1) = C(1)0$, $E(M_{A,N}(2) = C(2) = \frac{1}{4}$, and so on. Our distribution of eigenvalues must therefore be the semicircle.

As said in the theorem, it only works if the moments of the distribution used for the elements have mean 0, variance 1, and finite higher moments. In the experimental part of this paper two distributions used will satisfy this, the Laplace distribution and the Uniform Distribution, and the other, the Cauchy distribution, will not satisfy these requirements and will not converge to the semicircle. In fact all the moments of the Cauchy distribution are infinite. The calculation of the moments is as follows

$$\int_{-\infty}^{\infty} \frac{1}{\pi} \frac{x^k}{1+x^2} dx = \infty \tag{1.17}$$

for all even k. For odd k, and especially $k = 1$, if the limits are taken as $-y$ to y, then the integral is 0 and we can say that the previous integral is just in the limit $y \to \infty$. That is the reason we say that this distribution has mean 0, though we still say that the higher odd moments are not defined. Even though the three distributions used in this paper are completely different, it is theorized that the eigenvalue spacings will all converge to the GOE as the band radius is increased.

1.4 Gaussian Orthogonal Ensemble

Definition 1.4.1 (GOE). *The GOE is defined in the space of real symmetric matrices by two requirements*

1. *(Orthogonal Invariance) The ensemble is invariant under every transformation*

$$H \to W^T H W \tag{1.18}$$

   where W is any real orthogonal matrix

2. *(Independance of Matrix Elements) The various elements H_{kj}, $k \leq j$ are statistically independant*

The first requirement can also be written as

$$P(H')dH' = P(H)dH \tag{1.19}$$

where $P(H)$ is the probability of the system belonging to the volume element $dH = \prod_{k \leq j} dH_{kj}$, $H' = W^T H W$, and $W^T W = W W^T = 1$.

The second requirement says that the probability density function $P(H)$ is a product of functions, each of which depends on a single variable:

$$P(H) = \prod_{k \leq j} f_{kj}(H_{kj}) \tag{1.20}$$

f_{kj} is the probability distribution for the H_{kj}.

Throughout all of this, a probability distribution for the matrix elements, H_{ij}, has not been chosen. The following theorem proves that the only distribution that satisfies these two requirements is the Gaussian distribution.

Theorem 1.4.2. *Let H be a real symmetric matrix, with dimension N. If the matrix elements H_{ij} are independantly and identically distributed, as previously described. If there exists a probability distribution $P(H)$ satusfying both requirements given above, then the matrix elements H_{ij} must be Gaussian distributed.*

Proof: This proof follows both chapter 2, section 6, of [Me1] and the one given in [Liu], which he notes follows [Me1]. The proof in [Liu] is easier to follow as Mehta solves one equation for the unitary case and not the orthogonal case as needed here. To begin, let us look at the matrix of the two dimensional rotation through an angle θ:

$$U = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 & \ldots & 0 \\ -\sin(\theta) & \cos(\theta) & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix} \tag{1.21}$$

For this rotation we have

$$H = U^T H' U \tag{1.22}$$

It is fairly simple to calculate the relations between the matrix elements of H'_{ij} and H_{ij}. We get:

$$\begin{aligned}
H_{11} &= \frac{H'_{11}+H'_{22}}{2}+\frac{H'_{11}-H'_{22}}{2}\cos 2\theta - H'_{12}\sin 2\theta \\
H_{12} &= \frac{H'_{11}-H'_{22}}{2}\sin 2\theta + H'_{12}\cos 2\theta \\
H_{22} &= \frac{H'_{11}+H'_{22}}{2}-\frac{H'_{11}-H'_{22}}{2}\cos 2\theta + H'_{12}\sin 2\theta \\
H_{ij} &= H'_{ij} \text{ for all other } i,j
\end{aligned} \tag{1.23}$$

Remember that $H_{ij} = H_{ji}$ for all i, j. When writing out $P(H)$, the only factors which depend on θ are f_{11}, f_{12}, and f_{22}. In general if we have

$$\begin{aligned}
Q_{ii} &= Q_{jj} = \cos\theta \\
Q_{ij} &= -Q_{ji} = \sin\theta \\
Q_{kk} &= 1 \\
Q_{kl} &= 0 \text{ for the other elements of the matrix}
\end{aligned} \tag{1.24}$$

where i and j are specific numbers less than or equal to N and $i \neq j$.

Doing that we can let $f_{11} = f_{ii}$, $f_{12} = f_{ij}$, and $f_{22} = f_{jj}$. We now get:

$$P(H) = f_{ii}(H_{ii})f_{ij}(H_{ij})f_{jj}(H_{jj})\prod f_{kl}(H_{kl}) \tag{1.25}$$

From the orthogonal invariance condition imposed on $P(H)$ we see that $\frac{dP}{d\theta} = 0$. Therefore we get

$$\frac{dP}{d\theta} = \frac{f'_{ii}}{f_{ii}}\frac{dH_{ii}}{d\theta}P + \frac{f'_{ij}}{f_{ij}}\frac{dH_{ij}}{d\theta}P + \frac{f'_{jj}}{f_{jj}}\frac{dH_{jj}}{d\theta}P = 0 \tag{1.26}$$

Using our previous results for H_{ii}, H_{ij}, and H_{jj}, we can see that

$$\begin{aligned}
\frac{dH_{ii}}{d\theta} &= -2H_{ij} \\
\frac{dH_{ij}}{d\theta} &= H_{ii} - H_{jj} \\
\frac{dH_{jj}}{d\theta} &= 2H_{ij}
\end{aligned} \tag{1.27}$$

Substituting these back in to $\frac{dP}{d\theta}$ and dividing through by P, we get

$$\frac{f'_{ii}}{f_{ii}}(-2H_{ij}) + \frac{f'_{ij}}{f_{ij}}(H_{ii} - H_{jj}) + \frac{f'_{jj}}{f_{jj}}(2H_{ij}) = 0 \tag{1.28}$$

We can simplify this equation to get

$$\frac{f'_{ii}}{f_{ii}}\frac{2}{H_{ii} - H_{jj}} - \frac{f'_{jj}}{f_{jj}}\frac{2}{H_{ii} - H_{jj}} = \frac{f'_{ij}}{f_{ij}}\frac{1}{H_{ij}} = -C \tag{1.29}$$

The right hand side of the equation depends on H_{ij}, wherease the left hand side depends on H_{ii} and H_{ij}. which are different variables, so we introduce the constant C. Using this we can solve for f_{ij}.

$$\begin{aligned} f'_{ij} &= -CH_{ij}f_{ij} \\ &\Rightarrow \; f_{ij} = B_{ij}\exp(-CH_{ij}^2/2) \end{aligned} \tag{1.30}$$

Therefore the probability distribution on H_{ij} must be Gaussian. Now using the constant and the left hand side of the previous equation, we can get

$$\begin{aligned} \frac{f'_{ii}}{f_{ii}} - \frac{f'_{jj}}{f_{jj}} &= -\frac{C}{2}(H_{ii} - H_{jj}) \\ \frac{f'_{ii}}{f_{ii}} + \frac{C}{2}H_{ii} &= \frac{f'_{jj}}{f_{jj}} + \frac{C}{2}H_{jj} = K \end{aligned} \tag{1.31}$$

The constant K was introduced for the same reason C was previously introduced. The two sides of the last equation depend on different variables. We can now solve for both f_{ii} and fjj. They are

$$\begin{aligned} f_{ii} &= B_{ii}\exp(-CH_{ii}^2/4 + KH_{ii}) \\ f_{jj} &= B_{jj}\exp(-CH_{jj}^2/4 + KH_{jj}) \end{aligned} \tag{1.32}$$

In order for H_{ii} and H_{jj} to have mean 0, $K = 0$. Therefore we have shown that H_{ii}, H_{ij}, and H_{jj} must follow a Gaussian distribution for any i, j. Therefore all the elements must follow a Gaussian distribution and the theorem is proved.

We can now write $P(H)$, the probability density, as

$$
\begin{aligned}
P(H) &= C\exp\left(-\frac{1}{4\sigma^2}\left(\sum_j H_{jj}^2 + 2\sum_{i<j} H_{ij}^2\right)\right) \\
&= C\exp\left(-\frac{1}{4\sigma^2}\left(\sum_j H_{jj}^2 + \sum_{i<j} H_{ij}^2 + \sum_{i>j} H_{ij}^2\right)\right) \\
&= C\exp\left(-\frac{1}{4\sigma^2}\sum_{i,j} H_{ij}^2\right) \qquad (1.33)
\end{aligned}
$$

As said before $(H^2)_{ii} = \sum_j (H_{ij})^2$ for a symmetric matrix, so

$$
\begin{aligned}
P(H) &= C\exp\left(-\frac{1}{4\sigma^2}\sum_i (H^2)_{ii}\right) \\
&= C\exp\left(-\frac{1}{4\sigma^2}Tr(H^2)\right) \qquad (1.34)
\end{aligned}
$$

If we let $E_1, \ldots, E_N$ be the N eigenvalues of H then we can create the matrix A whose j^{th} column is the eigenvector corresponding to E_j. Then we have $HA = AE$, where E is the diagonal matrix of eigenvalues. The matrix A is determined by N normalization constraints and $\frac{N(N-1)}{2}$ orthogonal constraints, which will be called $\alpha_1, \ldots, \alpha_{N(N-1)/2}$. We can then write $P(H)$ as a function of the N eigenvalues of H and the $\frac{N(N-1)}{2}$ αs. We get

$$
P(E_1, \ldots, E_n, \alpha_1, \ldots, \alpha_{N(N-1)/2}) = P(H)|J| \qquad (1.35)
$$

where J is the Jacobian of the variable change.

$$
J = \begin{vmatrix}
\partial H_{11}/\partial E_1 & \ldots & \partial H_{11}/\partial E_N & \partial H_{11}/\partial \alpha_1 & \ldots & \partial H_{11}/\partial \alpha_{N(N-1)/2} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\partial H_{NN}/\partial E_1 & \ldots & \partial H_{NN}/\partial E_N & \partial H_{NN}/\partial \alpha_1 & \ldots & \partial H_{NN}/\partial \alpha_{N(N-1)/2}
\end{vmatrix} \qquad (1.36)
$$

where H_{ij} occuring in the Jacobian has $i \leq j$. From $HA = AE$, we can get $H = AEA^T$. We can write each H_{ij} in terms of the elements of A and the eigenvalues. We get

$$H_{ij} = \sum_k E_k A_{ik} A_{jk} \tag{1.37}$$

Therefore H_{ij} is linear in the eigenvalues. Therefore $\partial H_{ij}/\partial\alpha_k$ is linear in the eigenvalues. Therefore J is a polynomial of degree $N(N-1)/2$ in the eigenvalues. If $E_i = E_j$, then the corresponding eigenvectors cannot be uniquely determined and therefore J must vanish. This gives the result that J must contain the factor $(E_i - E_j)^\beta$ for all i, j. Since there are exactly $N(N-1)/2$ combinations for $(E_i - E_j)$, when we consider $(E_i - E_j)$ the same as $(E_j - E_i)$, in order to to get J as a polynomial of degree $N(N-1)/2$ in the eigenvalues, $\beta = 1$. We therefore get

$$J \prod_{i<j} (E_i - E_j) \cdot g(\alpha_1, \ldots, \alpha_{N(N-1)/2}) \tag{1.38}$$

where g is a function of the αs.We can can write out P in the new variables.

$$\begin{aligned} & P(E_1, \ldots, E_n, \alpha_1, \ldots, \alpha_{N(N-1)/2}) = && (1.39) \\ = \quad & C \exp\left(-\frac{1}{4\sigma^2} Tr(H^2)\right) \cdot \left| \prod_{i<j} (E_i - E_j) \cdots g(\alpha_1, \ldots, \alpha_{N(N-1)/2}) \right| \end{aligned}$$

If we integrate out over the αs and see that $Tr(H^2) = \sum_k E_k^2$, we get

$$P(E_1, \ldots, E_N) = K \cdot \exp\left(-\frac{1}{4\sigma^2} \sum_k E_k^2\right) \prod_{i<j} |E_i - E_j| \tag{1.40}$$

1.5 Spacings

Using the probability measure gotten above it would be possible to derive the eigenvalue spacings for the GOE. The Wigner surmise says that spacings between adjacent eigenvalues is

$$P(s) \approx \frac{\pi}{2} s \exp\left(-\frac{\pi}{4} s^2\right) \tag{1.41}$$

It turns out that the spacings do not actually follow that probability though they do follow a probability that is very close to Wigner's surmise. The derivation of the distribution is lengthy, technical, and complicated so only a brief summary will be given here. The full proof is in [Me2] and [Gau]

We start by integrating out, P, over all but 2 variables.

$$P(E_1, E_2) = \int \cdots \int P(E_1, \ldots, P_N) dE_3 \cdots dE_N \tag{1.42}$$

Since we want E_1 and E_2 to be adjacent eigenvalues, we take our integral for $E_3, \ldots, E_N$ over all values outside the interval between E_1 and E_2. If we set $E_1 = -\theta$ and $E_2 = \theta$, and let $N = 2m$, it can then be proved that

$$2m(2m-1)P(-\theta, \theta) = C \frac{2m! 2^{m-1}}{m!} \frac{d^2}{d\theta^2} \phi_m(\theta) \tag{1.43}$$

and

$$\phi_m(\theta) = \int_\theta^\infty \cdots \int_\theta^\infty \exp(-2(y_1^2 + \cdots + y_m^2)) \cdot \prod_{i<j} (y_i^2 - y_j^2)^2 dy_1 \cdots dy_m \tag{1.44}$$

If we define $\Psi_m(\theta) = \phi_m(\theta)/\phi_m(0)$. We then let $D = 2\sqrt{2m}/\pi$, and fix

$$\frac{s}{D} = \frac{2\theta}{D} = \frac{2t}{\pi} \tag{1.45}$$

If we let $\Psi(t) = \lim_{m \to \infty} \Psi_m(\theta)$, then it can be shown that

$$P(s) = \frac{\pi^2}{4} \frac{d^2 \Psi}{dt^2} \tag{1.46}$$

To evaluate the integral for ϕ_m, the ingerand is rewritten using a Vandermonde determinant, expressed as a determinant involving orthonormal wavefunctions of the harmonic oscillator, and finally Gram's Theorem is apllied. $\Psi_m(\theta)$ then becomes a Fredholm determinant of a certain integral operator. $\Psi(t)$ can then be expressed as the Fredholm determinant of the oparator

$$Tf(x) = \int_{-t}^{t} Q(x, y) f(y) dy \tag{1.47}$$

with

$$Q(x,y) = \frac{1}{2\pi}\left(\frac{\sin(x-y)}{x-y} + \frac{\sin(x+y)}{x+y}\right) \tag{1.48}$$

Using that we can then express $\Psi(t)$ as

$$\Psi(t) = \prod_{q=0}^{\infty}\left(1 - \frac{t}{2\pi}\gamma_{2q}^2\right) \tag{1.49}$$

where $\gamma_2 q$ are certain constants. This product converges quickly and it can be seen numerically that Wigner's Surmise is a very good approximation.

1.6 Diagonal Matrices

A derivation of the spacings between the eigenvalues of a diagonal matrix is not needed. What we want is the derivationg of the spacings between the middle $3/5$ of the eigenvalues of a diagonal matrix. The eigenvalues of a diagonal matrix are just the non-zero elements of the matrix, which happen to just be the numbers we pull from the distribution we are looking at. Therefore we would expect that the eigenvalue spacings should depend on the distribution used to generate the elements of the matrix. This is correct when all the spacings are used. However, in this paper, for every band radius only the middle $3/5$ of the eigenvalues were used. In turns out, that for every distribution, when a small enough interval is taken, the spacings of the generated elements are Poissonian. Note that this is only true if the spacings are normalized to mean 1, which they always are in this paper. The experimental part of this paper is now essentially determining how fast the eigenvalue spacings go from Poissonian to the GOE as the band radius is increased for different distributions generating the elements of the matrix.

Here is the proof that in a small enough interval, the spacings of the eigenvalues of a diagonal matrix will be Poissonian.

If $x_1, \ldots, x_N$ are the N eigenvalues of the diagonal matrix in increasing order, then we can fix x_1, and find the probability that the next spacing is a units away from x_1. Lets look at 2 intervals. The first interval is $[x_1, x_1 + a]$ and the second interval is $[x_1, x_1 + a + \Delta a]$

The probability of an element being in the first interval is:

$$\int_{x_1}^{x_1+a} p(u)du \tag{1.50}$$

where $p(u)$ is the probability distribution of the elements. The probability of an element being outside the interval is $1-$ the probability of being in the interval. If we let $a = \frac{t}{N}$, then $\Delta a = \frac{\Delta t}{N}$ and since the average spacing of N eigenvalues in an interval $[-b, b]$, which is what we are doing when taking the middle $3/5$ of the eigenvalues in the experiments, is $\approx \frac{2b}{N}$, the average spacing is now $2b$, when the variable is t instead of a. The probability of every other element (eigenvalue), being outside that interval is

$$\begin{aligned} \left(1 - \int_{x_1}^{x_1+a} p(u)du\right)^{N-1} &\approx \left(1 - \frac{p(x_1)t}{N}\right)^{N-1} \\ &\approx \exp(-p(x_1)t) \end{aligned} \tag{1.51}$$

Similarly, for the 2^{nd} interval, the probability of no element being in it is

$$\exp\left(-p(x_1)(t + \Delta t)\right) \tag{1.52}$$

Now we need to look at the difference between these two probabilities, which gives us the probability of the first eigenvalue being in the interval $[a, a + \Delta a]$ for a specific x_1.

$$\begin{aligned} \exp(-p(x_1)t) - \exp(-p(x_1)(t + \Delta t)) &= \exp(-p(x_1)t) * (1 - \exp(-p(x_1)\Delta t)) \\ &\approx \exp\left(-p(x_1)t\right) * p(x_1) * \Delta t \end{aligned} \tag{1.53}$$

We have to now integrate over all x_1 so the probability of having a spacing Δt is therefore

$$\Delta t \int_{x_1=-\infty}^{\infty} \exp(-p(x_1)t) * p(x_1)^2 dx_1 \tag{1.54}$$

When we restrict the values of the elements (eigenvalues) to a small enough interval, we can approximate the probability density of the elements as constant. Therefore we get the spacings as.

$$\Delta t \int \exp(-C * t) * C^2 dx_1 \tag{1.55}$$

where x_1 is defined on an interval $[-c, c]$. This integral results in a function of the form

$$C_1 * \exp(-C * t) \tag{1.56}$$

The probability of a spacing decreases exponentially as the spacing is increased. Therefore when the spacings are normalized to mean 1 and the eigenvalues are restricted to a small enough interval, the spacings will be Poissonian.

Chapter 2

Experiments

2.1 Background

As said previously, the experimental part of this paper looks at band-diagonal matrices and the spacing of their eigenvalues. The theory that a band of radius $\sqrt{N}$ behaves like the GOE is looked at. This is essentially that the same thing that part of [Liu] looks at, but using different probability distributions that are only conjectured to have eigenvalue spacings which converge to the GOE. The Cauchy distribution and the Uniform distribution of $[0.1)$, two of the distributions used here, were shown to have spacings in accordance with the GOE as $N \to \infty$ in [Le]. A third distribution, the Laplace distribution, was not looked at in either [Liu] or [Le].

The experiments consisted of the following steps in *Mathematica*:

1. Input variables such as *BandRadius*, N.

2. Generate a random $N \times N$ matrix using a built in random matrix generator. Set $A_{ij} = A_{ji}$ and set $A_{ij} = 0$ if $|i - j| <$ *BandRadius*.

3. Calculate the eigenvalues of the matrix

4. Calculate the nearest neighbor spacings of the middle $3/5$ of the eigenvalues and set the mean of these spacings to 1

5. Repeat many times until a good distribution can be seen from the eigenvalue spacings.

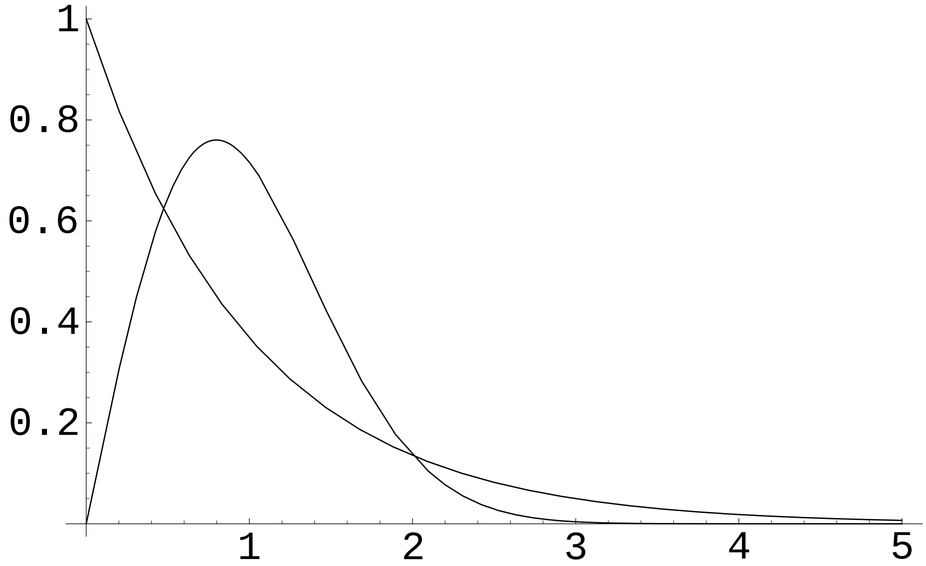
6. Plot the cumulative distribution functions of theoretical and experimental and calculate *Disc* and *Disc2*, defined below and in [Mi].

In both [Liu] and [Le] the probability density functions (PDF) were plotted instead of the cumulative distribution functions (CDF). The CDFs will be plotted in this paper because something that looks like an actual line, instead of a series of rectangles, can be seen. The discrepancy functions can also be applied easier to the CDFs.

Both *Disc* and *Disc2* were used previously in [Mi] to check the difference between theoretical and experimental CDFs. *Disc3* is the same as *Disc* except with a difference theoretical CDF. The same applies for *Disc4* with *Disc2*. In order to understand how the discrepance functions are applied it is necessary to know how the experimental CDF plot is generated. After getting all the K eigenvalue spacings, they are sorted. The CDF says the probability of a spacing being less than the value it is evaluated at. The experimental plot is generated the same way. A new $2 \times K$ matrix is generated. For the n^{th} column column, the first row contains the value of the spacing, v, where the spacings are sorted in ascending value as $n\ toK$ and the second row contains $\frac{n}{K}$, i.e. the probability that the spacing is less than or equal to v. There are n spacings less than are equal to v so the probability of being less than that in the experiment is n divided by the total number of spacings. The matrix is then plotted as K points, with the value in the first row being the x-coord and the value in the second row being the y-coord. When lots of points are used, the plot looks like a line rather than multiple points.

Disc and *Disc2* take each column of the matrix and evaluates the first row in the theorectical CDF and compares it with the second row. *Disc* returns the biggest absolute diffence between the theoretical and experimental distribtions, whereas *Disc2* returns the average absolute difference between them.

Some PDFs are also plotted for the Cauchy and Laplace distributions. The GOE and the Poisson distribtuion are plotted on this first plot so comparisons can be made between the experimental PDFs and the theoretical PDFs.



The Laplace distribution was normalized to have mean 0 and variance 1. Because the Cauchy distribution always has infinite variance, a mean of 0, i.e. symmetric about 0, was just used. The Uniform distribution is uniform on the interval $[0, 1)$. In equation form, for the Cauchy distribution, the probability of having a value between x and $x + dx$ is:

$$P_C(x)dx = \frac{1}{\pi}\frac{1}{1+x^2} \tag{2.1}$$

For the Laplce distribution, in equation form, the probability of having a value between x and $x + dx$ is:

$$P_L(x)dx = \frac{1}{\sqrt{2}}\exp(-\sqrt{2}|x|) \tag{2.2}$$

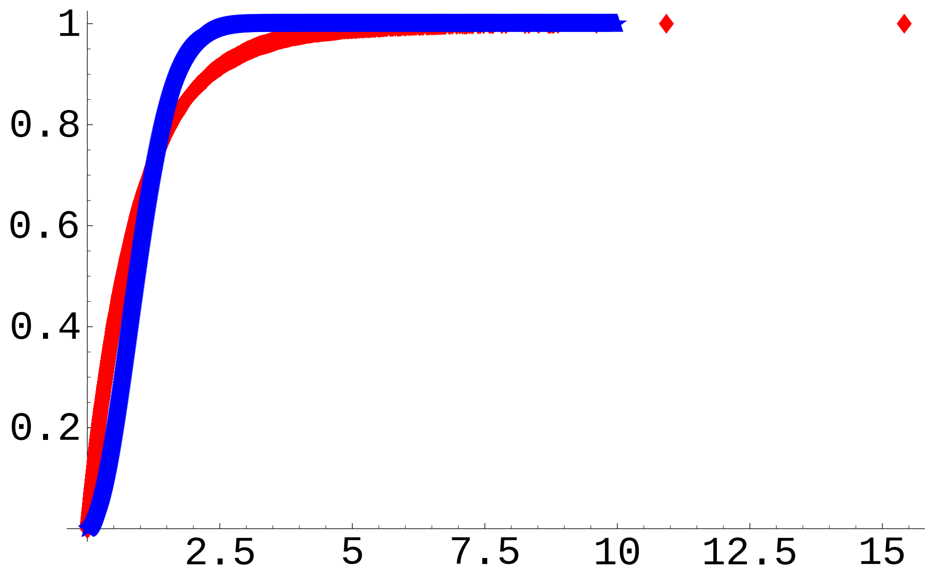
For the Uniform distribution, we have:

$$P_U(x)dx = 1 \tag{2.3}$$

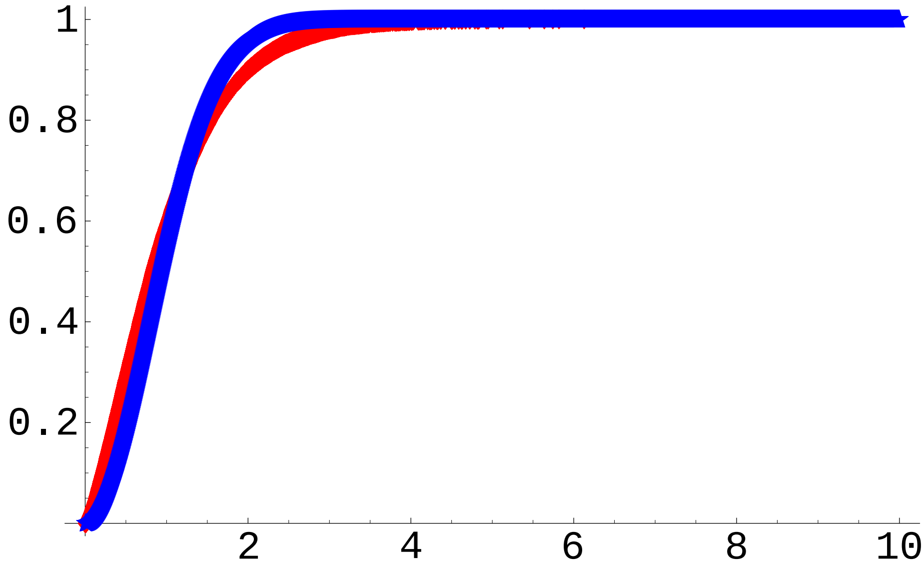
for $0 \leq x < 1$

2.2 Data

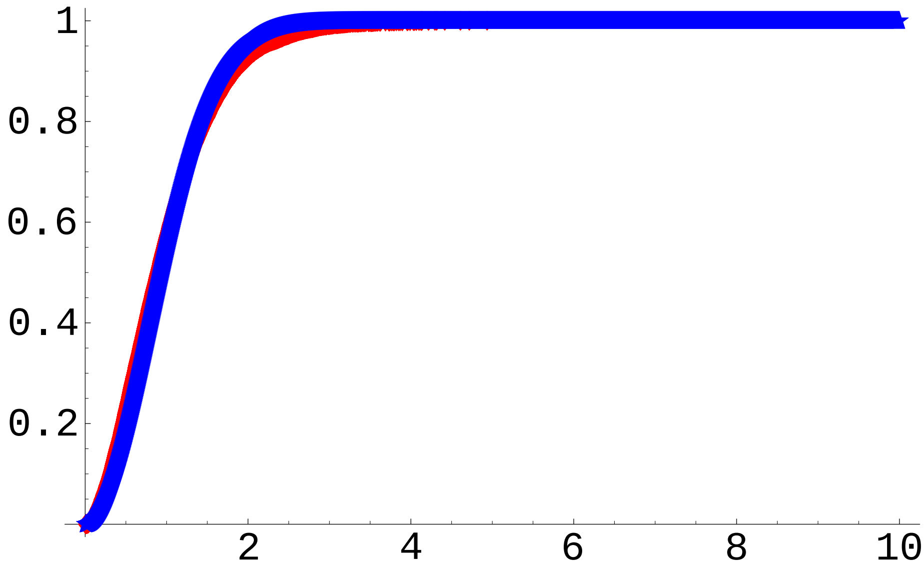
We will start off with several plots and their corresponding *Disc*s for the Laplace Distribution



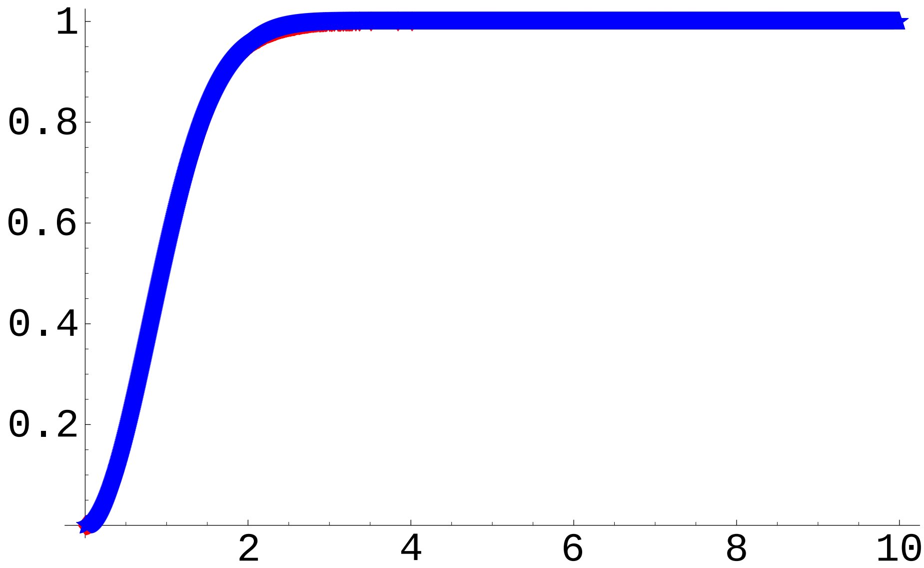
$N = 300$, 100 matrices, Band Radius (BR) $= 1$
Disc $= 0.229$, *Disc2* $= 0.123$



$N = 300$, 100 matrices, BR $= 10$
Disc $= 0.111$, *Disc2* $= 0.063$



$N = 300$, 100 matrices, BR $= 17$
Disc $= 0.055$, *Disc2* $= 0.031$

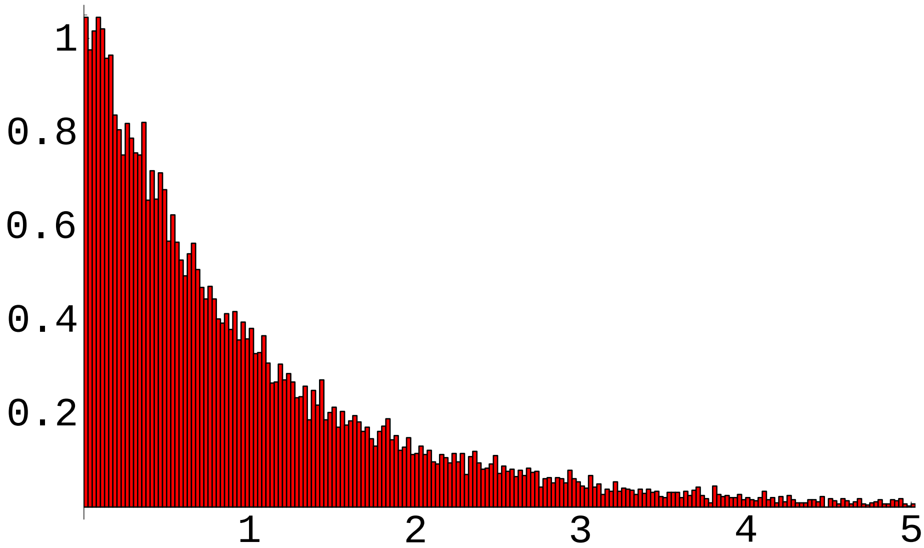


$N = 300$, 100 matrices, BR $= 300$
Disc ≈ 0.009, *Disc2* ≈ 0.005

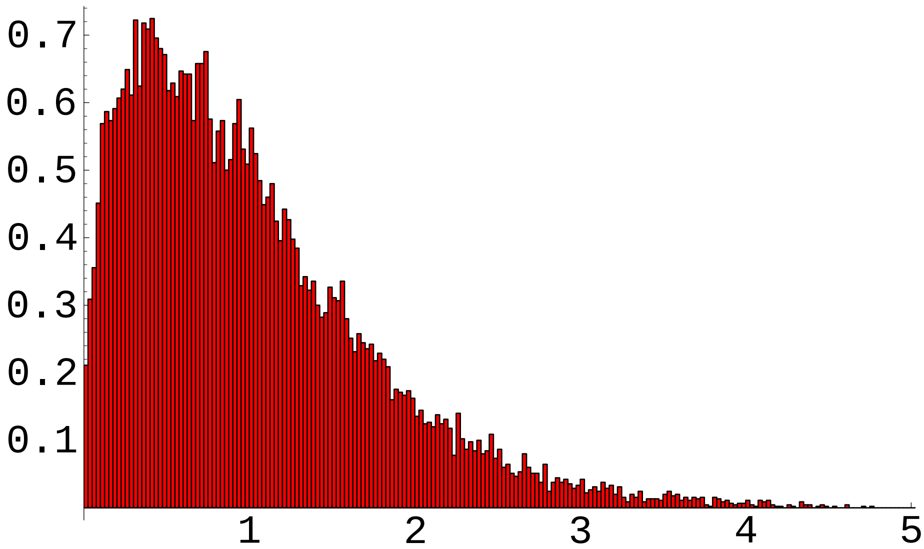
The following table lists the discrepancies for all these plus several other band radii. All of these were taken using a 100 300×300 matrices, where the middle $3/5$ of the eigenvalues were used in the calculation of the spacings

Band Radius	*Disc*	*Disc2*
1	0.229	0.123
5	0.177	0.096
8	0.139	0.076
10	0.111	0.063
13	0.082	0.046
16	0.058	0.034
17	0.055	0.031
18	0.049	0.028
20	0.044	0.024
300	0.009	0.005

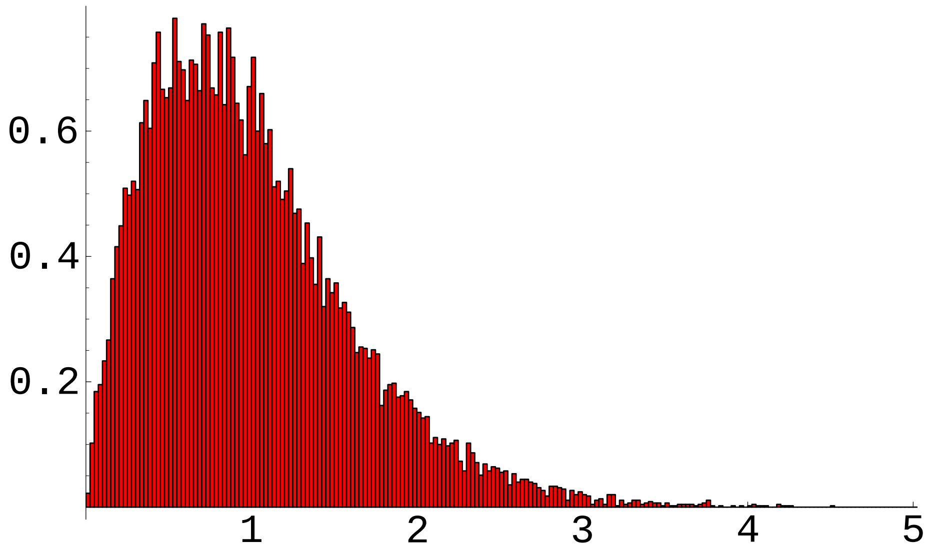
The following histograms are plots of the PDF for the Laplace Distribution.



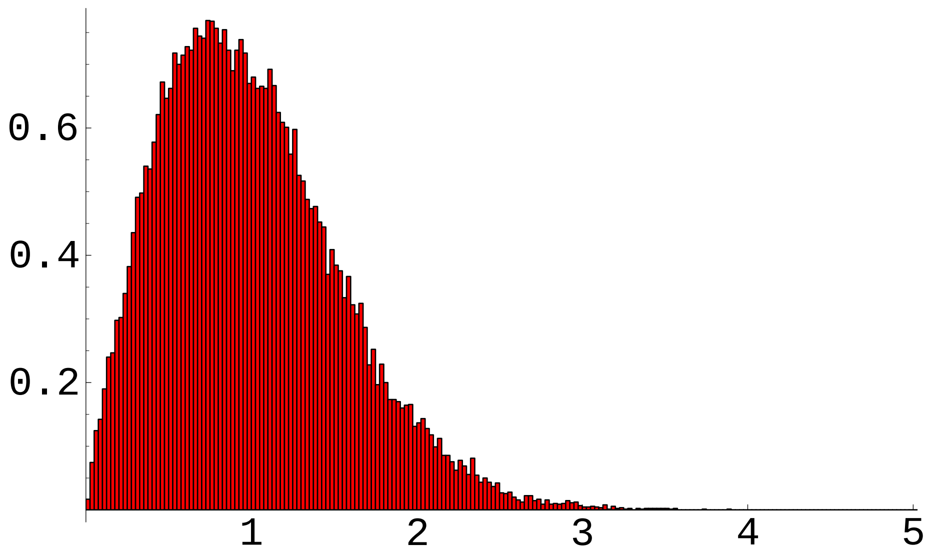
BR = 1



BR = 10



BR $= 18$



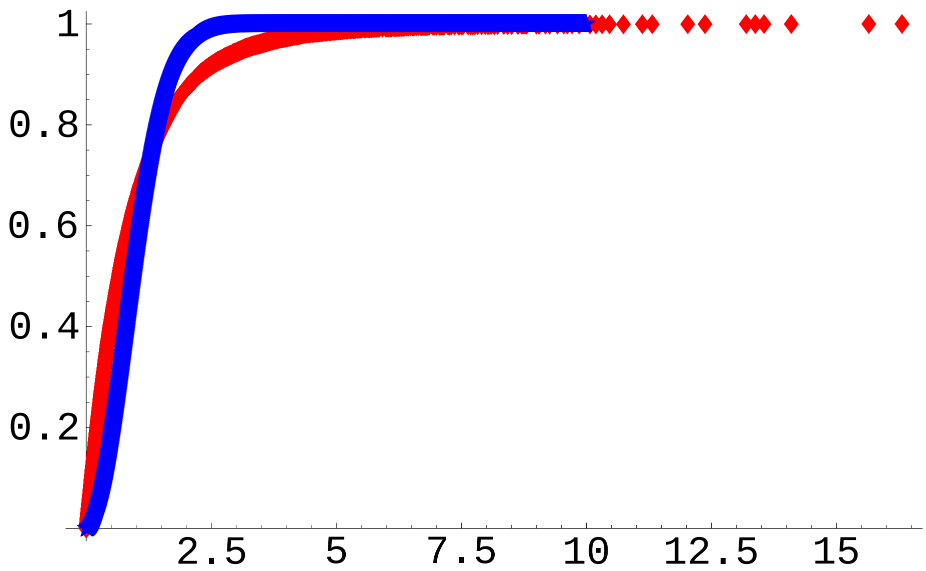
BR $= 300$

The spacings become essentially GOE at around a band radius of 16 to 18. 16 was chosen because it is the first to have *Disc* within 0.5 of *Disc* for a radius of 300. 18 was chosen because it is the first to have *Disc* ≤ 0.05. 0.05 is as good as any other number when testing convergence. Someone could say that within

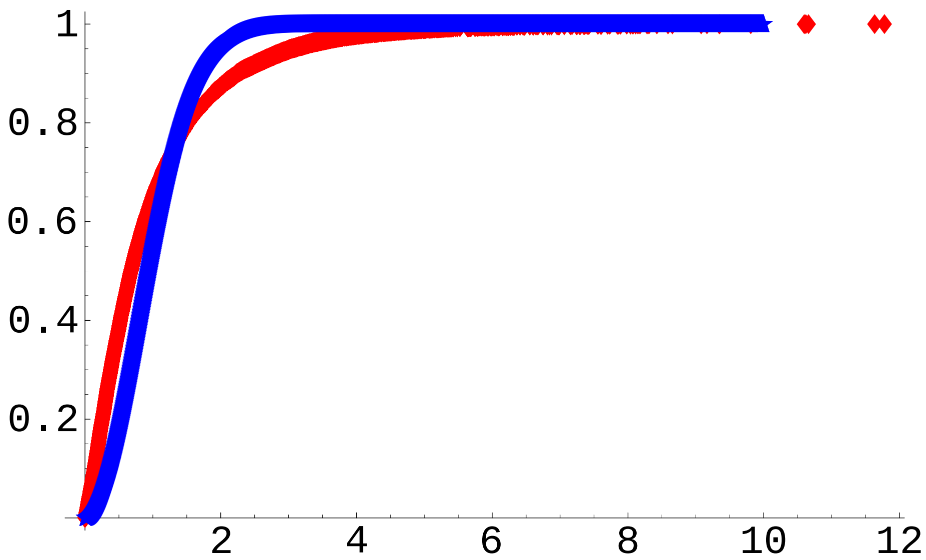
0.06 or 0.07 shows that the spacings have converged to the GOE and he/she would get different numbers on their estimate for the convergence to the GOE from the Poisson distribution. This estimate of 16 to 18 is around $\sqrt{300} = 17.32$, but we do not know for sure that the spacings will go to the GOE around $\sqrt{N}$ for a $N \times N$ matrix unless we look at much larger matrices, which are impossible at the current time due to speed.

The plots of the PDFs are using a much smaller number of matrices than the plots gotten in [Le] and [Liu], so there is a reason they do not look as smooth as in those papers.

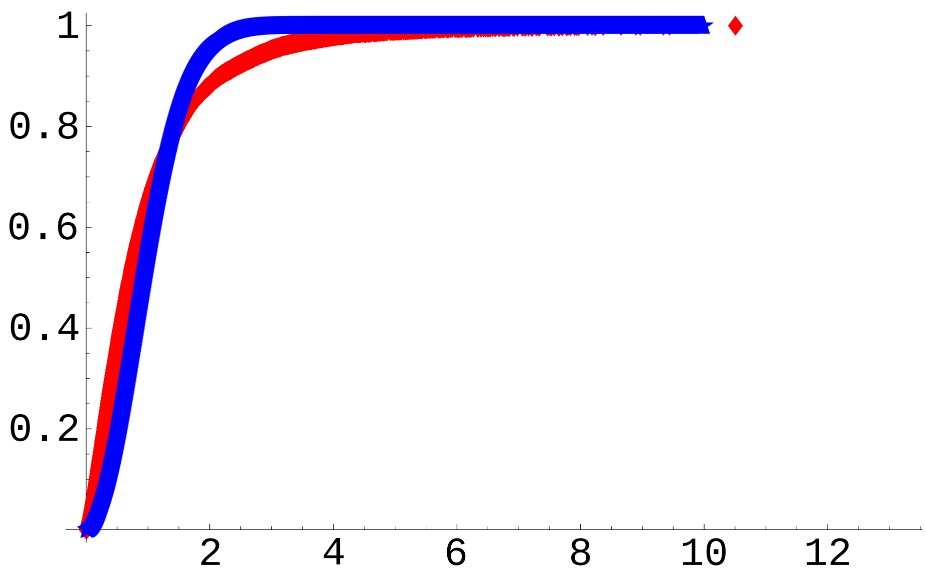
The following pictures illustrate the convergence for the Cauchy distribution.



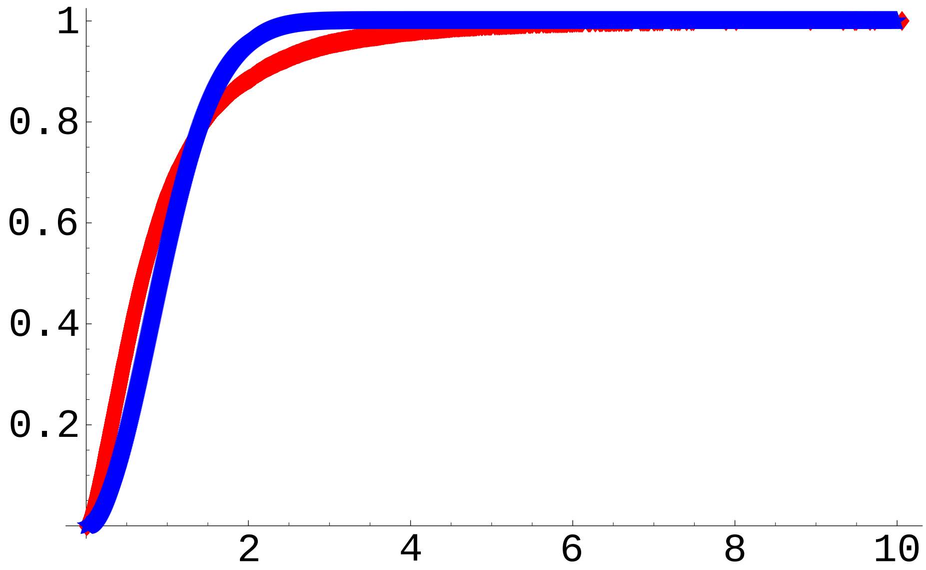
$N = 300$, 100 matrices, BR $= 5$
Disc $= 0.238$, *Disc2* $= 0.128$



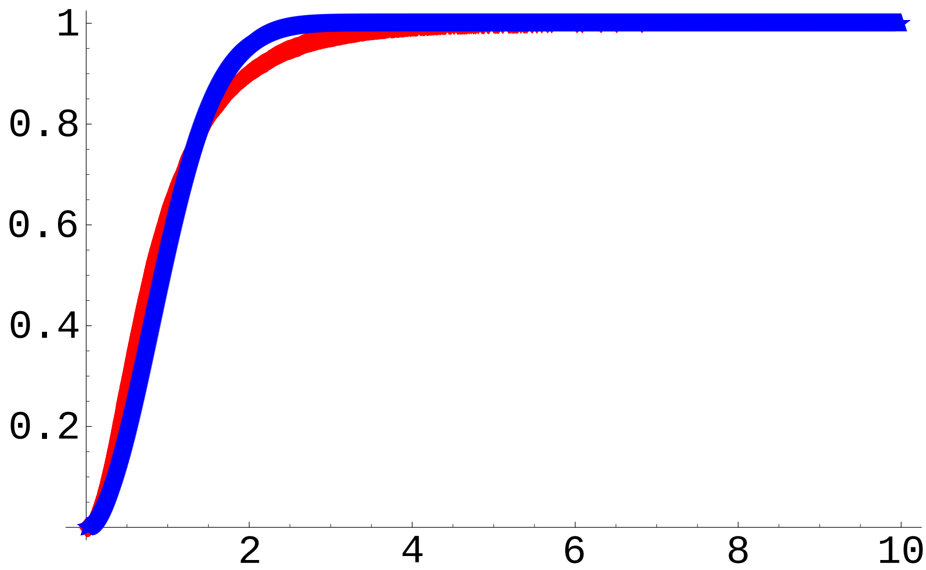
$N = 300$, 100 matrices, BR $= 10$
Disc $= 0.210$, *Disc2* $= 0.118$



$N = 300$, 100 matrices, BR $= 15$
Disc $= 0.189$, *Disc2* $= 0.108$



$N = 300$, 100 matrices, BR $= 18$
Disc $= 0.180$, *Disc2* $= 0.104$

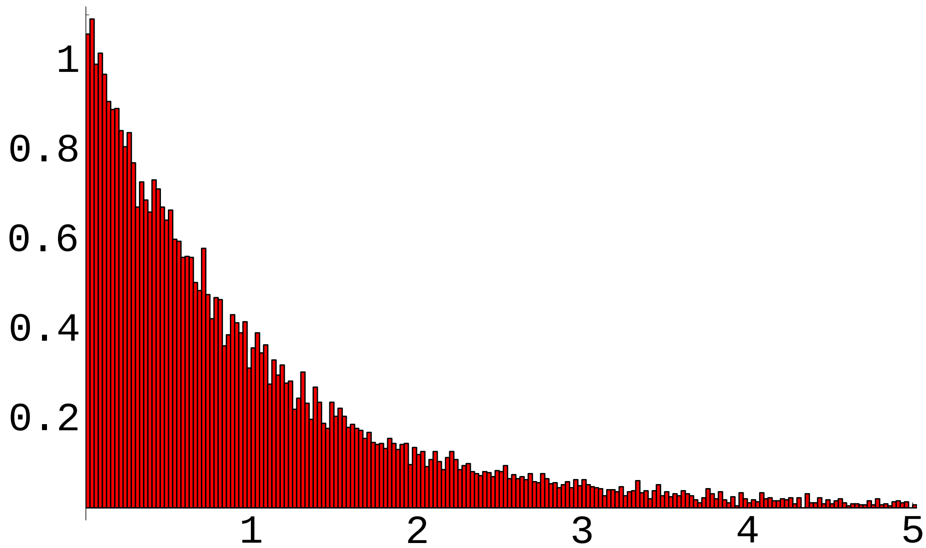


$N = 300$, 100 matrices, BR $= 300$
Disc $= 0.116$, *Disc2* $= 0.066$

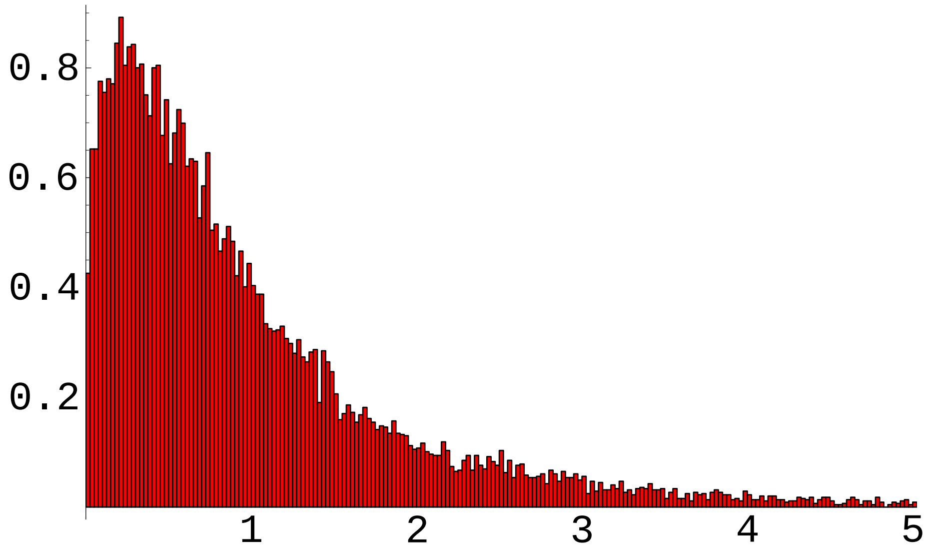
The following table lists discrepancies for several more band sizes.

Band Radius	*Disc*	*Disc2*
1	0.235	0.127
5	0.238	0.128
10	0.210	0.118
15	0.189	0.108
18	0.180	0.104
21	0.173	0.098
24	0.156	0.090
27	0.154	0.089
35	0.138	0.079
300	0.116	0.066

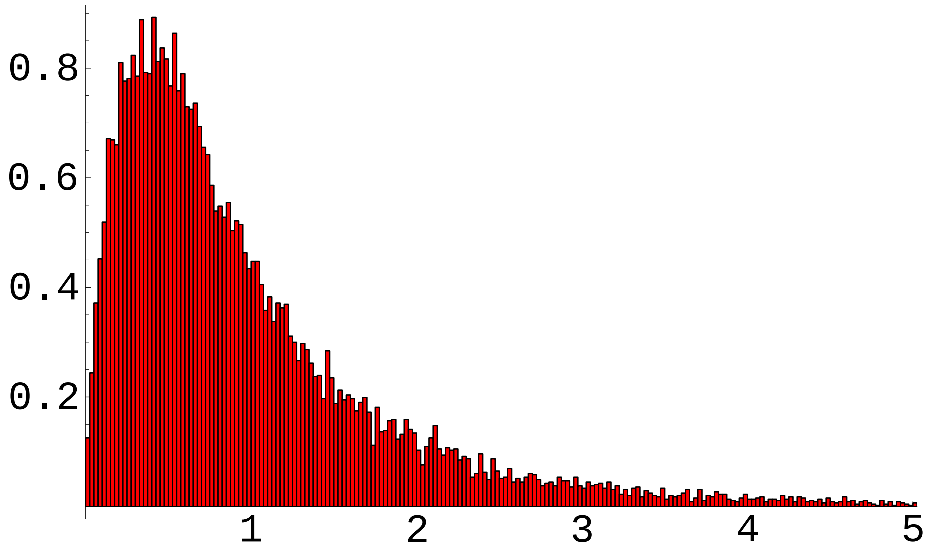
The following are the histogram plots for the PDFs.



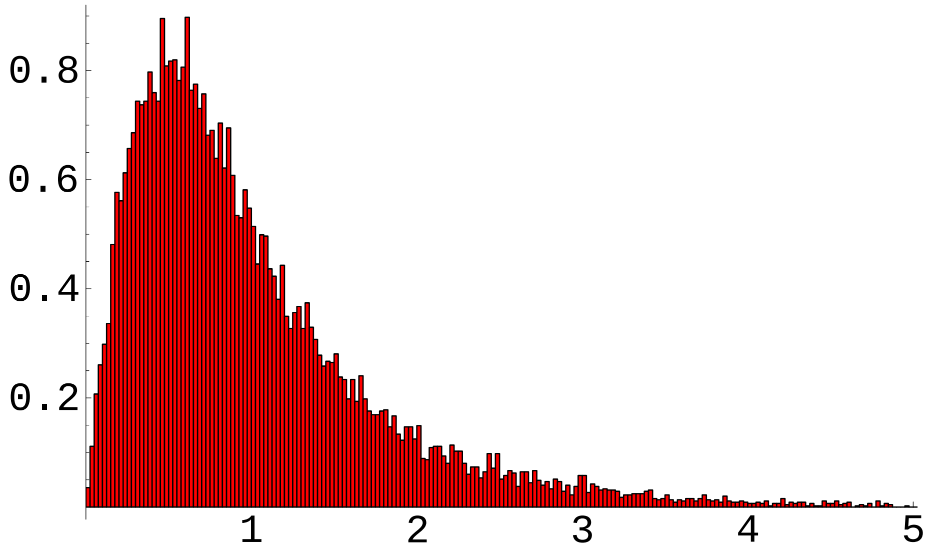
BR = 1



BR = 12



BR = 21



BR $= 300$

In this case, the discrepancies, and the spacings, converge more slowly to the spacings seem for the diagonal matrix. This could be due to the fact that the size of the matrix is small enough that the spacings of a random symmetric matrix with Cauchy distribution are not perfectly Poissonian, though it is probably more likely due to the fact that the Cauchy distribution has infinite variance. The plot of the PDFs show that the spacings do eventually converge to the GOE, yet for a 300×300 matrix, there are still some differences in the PDFs which correspond to a surprisingly larger difference in the CDF.

No plots will be given for the Uniform distribution. The following table lists the discrepancies between the experimental and the GOE (*Disc* and *Disc2*) and also the discrepancies between the experimental and the CDF for the Poissonian distribution (*Disc3* and *Disc4*)

Band Radius	*Disc*	*Disc2*	*Disc3*	*Disc4*
1	0.220	0.117	0.006	0.002
3	0.202	0.110	0.015	0.008
5	0.149	0.082	0.072	0.039
7	0.115	0.064	0.106	0.058
9	0.087	0.048	0.136	0.075
11	0.061	0.035	0.163	0.087
13	0.049	0.027	0.177	0.094
17	0.033	0.019	0.192	0.102
300	0.013	0.005	0.211	0.113

The Uniform distribution seems to converge faster to the GOE as the band size is increased than the Laplace distribution. There does not seem to be any specific band size for which for which we can say is the size for which the spacings for all distributions have converged to the GOE.

2.3 Conclusion and What Comes Next

The spacings do not seem to converge to the GOE at a band of size $\sqrt{N}$ for all distributions. Of course, the size of the matrices used in this study is small. For the Laplace distribution it looked as if the spacings might have converged around $\sqrt{N}$, though tests for larger matrices would be needed. For the Cauchy and the Uniform distribution, it looked as if $\sqrt{N}$ was not a good estimate for the convergence. For the Cauchy $\sqrt{N}$ looks to be too small, while for the Uniform distribution $\sqrt{N}$ looks to be too large. Something that could be done in the future would be to investigate larger matrices in order to see if the convergences are really different as matrix approaches infinite size or the differences are just because a 300×300 is small.

Another interesting thing shown in this paper is that if N values are taken from a distribution and put in increasing order, then the nearest neighbor spacings of the middle $3/5$ of the values, when the spacings are normalized to mean 1, are Poissonian. It would be could to investigate how many values along the edges must be left out in order to see the Poissonian distribution for different distributions.

Bibliography

[CRC] D. Zwillinger, ed. *CRC Standard Mathematical Tables and Formulae, 30th ed*, CRC Press, Boca Raton, FL, 1996

[Gau] M. Gaudin, "Sur la Loi Limite de L'Espacement des Valuers Propres d'une Matrice Aleatoire", *Nuclear Physics*, vol 25 (1961), reprinted in *Statistical Theories of Spectra*

[Le] R. Lehman, *First-Order Spacings of Random Matrix Eigenvalues*, Junior Research Project 2001

[Liu] Yi-Kai Liu, *Statistical Behavior of the Eigenvalues of Random Matrices*, Junior Research Project 2001

[Me1] M. L. Mehta, *Random Matrices, 2nd ed* Academic Press, San Diego, CA 1991

[Me2] M. L. Mehta, "On the Statistical Properties of the Level- Spacings", *Nuclear Physics*, vol 18, 1960 reprinted in *Statistical Theories of Spectra*

[Mi] N. Miller, *Various Tendencies of Non-Poissonian Distributions Along Subsequences of Certain Transcendental Numbers*, Junior Research Project 2002-2003

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

[SMS] Y. Sinai, S.J. Miller, L. Silberman, *Junior Research Seminar: Probability, L-Functions, Random Matrix Theory, and Ramanujan Graphs*, Lecture Notes, 2003