Statistical Behavior of the Eigenvalues of Random Matrices

Yi-Kai Liu
Mathematics Junior Seminar, Spring 2001
Princeton University

1 Introduction

This paper will investigate the statistical behavior of the eigenvalues of real symmetric random matrices. In particular, we shall be interested in the spacings s between adjacent eigenvalues. Let $P(s)$ be the distribution of these spacings, in the limit of matrices of large dimension. Empirical evidence suggests that, for a large class of random matrices, $P(s)$ is given approximately by the "Wigner surmise":

$$P(s) \approx As \exp(-Bs^2) \tag{1}$$

However, the validity of this approximation has only been proved for matrices whose elements are i.i.d. Gaussian (the "Gaussian Ensembles"). This paper will summarize the theoretical understanding of the "Gaussian Orthogonal Ensemble" (GOE), and then describe some computer experiments which suggest that the eigenvalues of sparse and band-diagonal random matrices behave in a similar way.

2 The Theory of Random Matrices

2.1 Application to Nuclear Physics

To put the theory in context, we begin with some applications of random matrices to nuclear physics. Indeed, this was one of the original motivations for the study of random matrices. Loosely speaking, a quantum mechanical system is described by an eigenvalue problem

$$H\psi_n = E_n\psi_n \tag{2}$$

where H is a Hermitian operator on function-space, ψ_n is an eigenfunction, and E_n is the corresponding (scalar) eigenvalue. The operator H stands for

some physical measurement or observation, which can distinguish among different "states" of the system. Each state is represented by an eigenfunction, and the corresponding eigenvalue is the value that would be measured if the system were in that state. (Because H is Hermitian, its eigenvalues are real.) In the case of an atomic nucleus, H is the "Hamiltonian", and the eigenvalue E_n denotes the n-th energy level.

Most nuclei have thousands of states and energy levels, and are too complex to be described exactly. Instead, one must settle for a model that captures the statistical properties of the energy spectrum. Instead of dealing with the actual operator H, one can consider a family of random matrices, and compute the distribution of the eigenvalues of these matrices. This is a rather crude model, since it replaces the infinite-dimensional operator H with a set of matrices of finite dimension N. To make up for this shortcoming, one should look for the asymptotic behavior as $N \to \infty$. Also, each random matrix has only a finite spectrum of eigenvalues; but one might reasonably hope that the middle part of this spectrum, away from the edges, will still have similar properties to the actual infinite spectrum of the nucleus.

It is worth mentioning that some properties of the physical system are reflected in the choices of random matrices. For instance, if a physical system is invariant under time reversal, its Hamiltonian must be self-dual. The Gaussian Orthogonal Ensemble describes the special case of time-reversal invariance with integral total angular momentum. These issues are discussed fully in Chapter 2 of [Mehta, 1991]; see also the introduction in [*Statistical Theories of Spectra*].

We will now describe some results from random matrix theory on the distribution of eigenvalues, and the distribution of eigenvalue spacings.

2.2 The Semicircle Rule

Take a family of symmetric random matrices, of dimension N, chosen from some distribution D. Let $P_N(x)$ be the distribution of the eigenvalues, normalized so that the eigenvalues lie in the interval [-1,1], and the total area under the distribution is 1. Then, for suitable distributions D, the eigenvalues obey a "semicircle rule": As $N \to \infty$, $P_N(x)$ converges to a semicircular distribution $P(x)$. That is,

$$P_N(x) \to P(x) = \frac{2}{\pi}\sqrt{1 - x^2} \tag{3}$$

In particular, if the distribution D has finite moments of all orders, then we can prove the semicircle rule using the "method of moments". Here is an example:

Theorem 1 (Semicircle Rule) *Let D be a distribution having finite moments of all orders; specifically, if X is chosen from distribution D, then $\mathrm{E}(X) = 0$, $\mathrm{E}(X^2) = 1$, and, for all $k \geq 3$, $\mathrm{E}(X^k)$ is finite.*

Let us construct a family of real symmetric random matrices, of dimension N, as follows: To get a random matrix A, choose its elements A_{ij} (with $i \leq j$) independently from distribution D; the remaining elements are then determined by symmetry.

Let $\lambda_j/2\sqrt{N}$ be the normalized eigenvalues of A, and define their distribution

$$\mu_{A,N}(x) = \frac{1}{N}\sum_{j=1}^{N} \delta\Big(x - \frac{\lambda_j}{2\sqrt{N}}\Big) \tag{4}$$

We claim that, as $N \to \infty$,

$$\mu_{A,N}(x) \to P(x) = \frac{2}{\pi}\sqrt{1-x^2} \tag{5}$$

with probability $\to 1$.

Proof: Let $U(x^k)$ be the k-th moment of $\mu_{A,N}$:

$$\begin{aligned} U(x^k) &= \int_{-\infty}^{+\infty} x^k \mu_{A,N}(x)dx \\ &= \frac{1}{N}\sum_{j=1}^{N} \Big(\frac{\lambda_j}{2\sqrt{N}}\Big)^k \end{aligned} \tag{6}$$

Calculate the expected value of each moment:

$$\mathrm{E}(U(x^0)) = \mathrm{E}(1) = 1 \tag{7}$$

$$\begin{aligned} \mathrm{E}(U(x^1)) &= \mathrm{E}\Big(\frac{1}{N}\sum_{j=1}^{N} \frac{\lambda_j}{2\sqrt{N}}\Big) = \frac{1}{2N^{3/2}}\,\mathrm{E}(\mathrm{Tr}\,A) \\ &= \frac{1}{2N^{3/2}}\,\mathrm{E}\Big(\sum_{j=1}^{N} A_{jj}\Big) \\ &= \frac{1}{2N^{3/2}}\sum_{j=1}^{N} \mathrm{E}(A_{jj}) = 0 \end{aligned} \tag{8}$$

For $k = 2$, use the fact that $(A^2)_{jj} = \sum_{k=1}^{N} A_{jk}A_{kj} = \sum_{k=1}^{N}(A_{jk})^2$.

$$\begin{aligned}
\mathrm{E}(U(x^2)) &= \mathrm{E}\Big(\frac{1}{N}\sum_{j=1}^{N}\Big(\frac{\lambda_j}{2\sqrt{N}}\Big)^2\Big) = \frac{1}{4N^2}\,\mathrm{E}(\mathrm{Tr}(A^2)) \\
&= \frac{1}{4N^2}\,\mathrm{E}\Big(\sum_{j=1}^{N}(A^2)_{jj}\Big) \\
&= \frac{1}{4N^2}\,\mathrm{E}\Big(\sum_{j=1}^{N}\sum_{k=1}^{N}(A_{jk})^2\Big) \\
&= \frac{1}{4N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\mathrm{E}((A_{jk})^2) = \frac{N^2}{4N^2} = \frac{1}{4}
\end{aligned} \tag{9}$$

And so on, up to higher-order moments. The actual distribution is determined by its moments, provided that those moments do not increase too rapidly with k. Thus it will suffice to compare the moments of the eigenvalue distribution with the moments of the "semicircle".

Let $C(x^k)$ be the k-th moment of the semicircle, $P(x)$:

$$C(x^k) = \int_{-1}^{+1} x^k P(x)\,dx = \int_{-1}^{+1} \frac{2}{\pi}x^k\sqrt{1-x^2}\,dx \tag{10}$$

Make the substitution $x = \sin\theta$:

$$C(x^k) = \int_{-\pi/2}^{+\pi/2} \frac{2}{\pi}\sin^k\theta\,\cos^2\theta\,d\theta \tag{11}$$

For k odd, $C(x^k)$ vanishes by symmetry. For k even, we substitute in $\cos^2\theta = 1 - \sin^2\theta$, and find that:

$$C(x^k) = \int_{-\pi/2}^{+\pi/2} \frac{2}{\pi}\sin^k\theta\,d\theta - \int_{-\pi/2}^{+\pi/2} \frac{2}{\pi}\sin^{k+2}\theta\,d\theta \tag{12}$$

These integrals can be evaluated analytically (see Gradshteyn and Ryzhik, *Table of Integrals, Series, and Products, 5th ed.*, p.412). Let us define some notation: For n even, $n!! = 2\cdot 4\cdots n$; and for n odd, $n!! = 1\cdot 3\cdots n$. Then we can write:

$$\begin{aligned}
C(x^k) &= 2\frac{(k-1)!!}{k!!} - 2\frac{(k+1)!!}{(k+2)!!} \\
&= \frac{2\cdot(k-1)!!}{k!!}\Big(1 - \frac{k+1}{k+2}\Big) \\
&= \frac{2\cdot(k-1)!!}{(k+2)!!}
\end{aligned} \tag{13}$$

In particular, we have $C(x^0) = 1$, $C(x^1) = 0$, $C(x^2) = 1/4$ and $C(x^3) = 0$. These are equal to the corresponding moments (calculated above) for the eigenvalue distribution. By extending this approach to include higher moments, we can prove that the eigenvalue distribution goes asymptotically to the semicircle. ■

2.3 The Gaussian Orthogonal Ensemble

From this point on, we will concentrate specifically on random matrices belonging to the "Gaussian Orthogonal Ensemble" (GOE). First, consider the family of real symmetric random matrices, of dimension N. For any matrix in this family, the $N(N+1)/2$ matrix elements which lie on or above the diagonal can be chosen freely; the remaining elements are then determined by symmetry. So a random matrix H depends on $N(N+1)/2$ random variables, namely, the elements H_{ij} with $i \leq j$. We assume that these random variables are independent and identically distributed (i.i.d.).

Moreover, we seek to define a probability measure $P(H)$ on this family of matrices, subject to the following conditions:

1. Orthogonal invariance: For any real orthogonal matrix Q, we have:

$$P(Q^T H Q) = P(H) \tag{14}$$

2. The random matrix elements H_{ij} (with $i \leq j$) are statistically independent. Thus $P(H)$ can be written as a product:

$$P(H) = \prod_{i \leq j} f_{ij}(H_{ij}) \tag{15}$$

where f_{ij} is the probability distribution of H_{ij}.

The second condition is intended mainly to simplify things, but the first is absolutely essential, if we are to deal with random matrix eigenvalues. $P(H)$ must depend on abstract linear transforms, irrespective of basis, because these abstract transforms determine the eigenvalues.

So far we have not specified a probability distribution for the matrix elements H_{ij}. However, we shall prove the following result:

Theorem 2 *Suppose H is a real symmetric random matrix, of dimension N; suppose the matrix elements H_{ij} are i.i.d., as described above; and suppose there exists a probability measure $P(H)$ satisfying both condition 1 (orthogonal invariance) and condition 2 (independence of matrix elements).*

Then the matrix elements H_{ij} must be Gaussian distributed.

The resulting family of real symmetric matrices, whose elements are i.i.d. Gaussian, is called the Gaussian Orthogonal Ensemble (GOE).

Theorem 2 tells us that, by assuming conditions 1 and 2 above, we are restricting ourselves to work with Gaussian ensembles. Since our analysis of random matrix eigenvalues will rely heavily on these two conditions, it will be valid only for Gaussian ensembles; it cannot be generalized to other kinds of random matrices. This is one of the chief obstacles to any *general* theory of random matrices.

Proof of Theorem 2: The following argument is from [Porter and Rosenzweig, 1960]; see also Chapter 2 in [Mehta, 1991].

Consider a particular orthogonal transformation, namely the two-dimensional rotation through angle θ:

$$Q = \begin{pmatrix} \cos\theta & \sin\theta & 0 & \dots & 0 \\ -\sin\theta & \cos\theta & 0 & \dots & 0 \\ 0 & 0 & 1 & & \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & & & 1 \end{pmatrix} \tag{16}$$

The rotation Q acts as follows:

$$H = Q^T H' Q \tag{17}$$

It is straightforward to calculate the relations between the matrix elements H_{ij} and H'_{ij}.

$$\begin{aligned} H_{11} &= \frac{H'_{11} + H'_{22}}{2} + \frac{H'_{11} - H'_{22}}{2}\cos 2\theta - H'_{12}\sin 2\theta \\ H_{12} &= \frac{H'_{11} - H'_{22}}{2}\sin 2\theta + H'_{12}\cos 2\theta \\ H_{22} &= \frac{H'_{11} + H'_{22}}{2} - \frac{H'_{11} - H'_{22}}{2}\cos 2\theta + H'_{12}\sin 2\theta \\ H_{ij} &= H'_{ij} \text{ for all other } i, j \end{aligned} \tag{18}$$

In the product formula for $P(H)$, notice that the factors which depend on θ are f_{11}, f_{12} and f_{22}.

$$P(H) = f_{11}(H_{11})\ f_{12}(H_{12})\ f_{22}(H_{22}) \prod_{\text{all other } i \le j} f_{ij}(H_{ij}) \tag{19}$$

Orthogonal invariance of P implies that:

$$\frac{dP}{d\theta} = 0 \tag{20}$$

Therefore

$$\frac{dP}{d\theta} = \frac{f'_{11}}{f_{11}}\frac{dH_{11}}{d\theta}P + \frac{f'_{12}}{f_{12}}\frac{dH_{12}}{d\theta}P + \frac{f'_{22}}{f_{22}}\frac{dH_{22}}{d\theta}P = 0 \tag{21}$$

However, from equations (18) we can deduce:

$$\begin{aligned}
\frac{dH_{11}}{d\theta} &= -(H'_{11} - H'_{22})\sin 2\theta - 2H'_{12}\cos 2\theta \\
&= -2H_{12} \\
\frac{dH_{12}}{d\theta} &= (H'_{11} - H'_{22})\cos 2\theta - 2H'_{12}\sin 2\theta \\
&= H_{11} - H_{22} \\
\frac{dH_{22}}{d\theta} &= (H'_{11} - H'_{22})\sin 2\theta + 2H'_{12}\cos 2\theta \\
&= 2H_{12}
\end{aligned} \tag{22}$$

Substitute these into (21) to get:

$$\frac{f'_{11}}{f_{11}}(-2H_{12}) + \frac{f'_{12}}{f_{12}}(H_{11} - H_{22}) + \frac{f'_{22}}{f_{22}}(2H_{12}) = 0 \tag{23}$$

Divide across by $-H_{12}(H_{11} - H_{22})$ to get:

$$\frac{f'_{11}}{f_{11}}\frac{2}{H_{11} - H_{22}} - \frac{f'_{22}}{f_{22}}\frac{2}{H_{11} - H_{22}} = \frac{f'_{12}}{f_{12}}\frac{1}{H_{12}} = -C \tag{24}$$

where the constant C is introduced because the two sides of the equation depend on different variables. This allows us to solve for f_{12}:

$$f'_{12} = -CH_{12}f_{12} \text{ has a solution } f_{12} = B_{12}\exp(-CH_{12}^2/2) \tag{25}$$

We can rewrite equation (24) as:

$$\frac{f'_{11}}{f_{11}} - \frac{f'_{22}}{f_{22}} = -\frac{C}{2}(H_{11} - H_{22}) \tag{26}$$

$$\frac{f'_{11}}{f_{11}} + \frac{C}{2}H_{11} = \frac{f'_{22}}{f_{22}} + \frac{C}{2}H_{22} = K \tag{27}$$

The second step introduced another constant K because the two sides of the equation depend on different variables. Then we solve for f_{11} and f_{22}:

$$f_{11} = B_{11}\exp(-CH_{11}^2/4 + KH_{11}) \tag{28}$$

$$f_{22} = B_{22}\exp(-CH_{22}^2/4 + KH_{22}) \tag{29}$$

If H_{11} and H_{22} have mean 0, then $K = 0$ in the above solutions.

By choosing different rotations Q, the above argument can be made to cover all the elements of the matrix H. Hence, the matrix elements H_{ij} must follow a Gaussian distribution. So we have arrived at the Gaussian Orthogonal Ensemble. ■

We can derive many interesting results about the GOE. Using our solutions for the f_{ij}, we can re-write the probability density (see equation (15)) as:

$$\begin{aligned} P(H) &= C \exp\Big(-\frac{1}{4\sigma^2}\Big(\sum_j H_{jj}^2 + 2\sum_{i<j} H_{ij}^2\Big)\Big) \\ &= C \exp\Big(-\frac{1}{4\sigma^2}\sum_{i,j} H_{ij}^2\Big) \end{aligned} \tag{30}$$

This is the expression for $P(H)$ in the case where the matrix elements all have mean 0, the diagonal elements have variance $2\sigma^2$ and the off-diagonal elements have variance σ^2. When we discuss the theory of the GOE random matrices, we will usually be dealing with this case. (Note, however, that the off-diagonal elements greatly outnumber the diagonal elements. For purposes of numerical experimentation, it does not matter much if the variances of the diagonal and off-diagonal elements are not the same.)

We can write equation (30) very neatly, by observing that $(H^2)_{ii} = \sum_j H_{ij}H_{ji} = \sum_j (H_{ij})^2$.

$$\begin{aligned} P(H) &= C \exp\Big(-\frac{1}{4\sigma^2}\sum_i (H^2)_{ii}\Big) \\ &= C \exp\big(-\frac{1}{4\sigma^2}\operatorname{Tr}(H^2)\big) \end{aligned} \tag{31}$$

In this form, the orthogonal invariance of $P(H)$ is plainly visible.

2.4 Probability Distribution of the Eigenvalues

Using (30), we can derive the probability distribution of the eigenvalues, for random matrices belonging to the GOE. We will follow the argument given in [Porter and Rosenzweig, 1960]. Let E_1, E_2, ..., E_N be the eigenvalues, and for each E_k, let $V_k = (a_{1k}, a_{2k}, ..., a_{Nk})$ be the corresponding eigenvector. The eigenvectors make up the collumns of an $N \times N$ orthogonal matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{pmatrix} \tag{32}$$

The matrix A must satisfy N normalization constraints (one for each collumn), as well as $N(N-1)/2$ constraints to ensure that it is an orthogonal matrix. Thus, the matrix A can be determined by at most $N(N-1)/2$ independent parameters; call them α_1, α_2, ..., $\alpha_{N(N-1)/2}$.

We will take the probability measure $P(H)$ in equation (30), and rewrite it in terms of the N eigenvalues E_k and the $N(N-1)/2$ parameters α. This will involve substituting for the matrix elements H_{ij}, and calculating the Jacobian J.

The eigenvectors V_k and the eigenvalues E_k obey relations of the form $HV_k = E_k V_k$. We can combine these into a single matrix equation:

$$\begin{pmatrix} H_{11} & \dots & H_{1N} \\ \vdots & \ddots & \vdots \\ H_{N1} & \dots & H_{NN} \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \dots & a_{NN} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \dots & a_{NN} \end{pmatrix} \begin{pmatrix} E_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & E_N \end{pmatrix} \tag{33}$$

This can be written more succinctly as

$$HA = AE \tag{34}$$

where $E = \mathrm{diag}(E_1, E_2, ..., E_N)$ is a diagonal matrix containing the eigenvalues. Now multiply by A^T, keeping in mind that A is orthogonal, to get:

$$H = AEA^T \tag{35}$$

which gives the following expression for the matrix elements:

$$H_{ij} = \sum_k E_k a_{ik} a_{jk} \tag{36}$$

Also observe that, because A is orthogonal, $A^T A = I$, which implies the following:

$$\sum_k a_{ki} a_{kj} = \delta_{ij} \tag{37}$$

We use the above results to derive an expression which can be substituted into $P(H)$:

$$\begin{aligned} \sum_{i,j} H_{ij}^2 &= \sum_{i,j} \sum_{k,\ell} E_k E_\ell a_{ik} a_{jk} a_{i\ell} a_{j\ell} \\ &= \sum_k E_k^2 \end{aligned} \tag{38}$$

The Jacobian for this change of variables takes the form

$$J = \begin{vmatrix} \partial H_{11}/\partial E_1 & \dots & \partial H_{11}/\partial E_N & \partial H_{11}/\partial \alpha_1 & \dots & \partial H_{11}/\partial \alpha_{N(N-1)/2} \\ \partial H_{12}/\partial E_1 & \dots & \partial H_{12}/\partial E_N & \partial H_{12}/\partial \alpha_1 & \dots & \partial H_{12}/\partial \alpha_{N(N-1)/2} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \partial H_{NN}/\partial E_1 & \dots & \partial H_{NN}/\partial E_N & \partial H_{NN}/\partial \alpha_1 & \dots & \partial H_{NN}/\partial \alpha_{N(N-1)/2} \end{vmatrix} \tag{39}$$

Equation (36) shows that each H_{ij} is a linear function of all the eigenvalues E_k. So every entry of the form $\partial H_{ij}/\partial \alpha$ is linear in the eigenvalues. It follows that J is a polynomial of degree $N(N-1)/2$ in each of the eigenvalues.

Now suppose that two eigenvalues E_i and E_j are equal. Then their corresponding eigenvectors V_i and V_j are not uniquely determined. As a result, the inverse of the transformation in (36) is not unique. So the Jacobian J must vanish at $E_i = E_j$; therefore, J must contain the factor $(E_i - E_j)$.

This reasoning applies for any E_i and E_j, so J must contain all factors of the form $(E_i - E_j)$. Since there are exactly $N(N-1)/2$ such factors, and we know J is a polynomial of degree $N(N-1)/2$ in the eigenvalues, it follows that we have completely accounted for J's dependence on the eigenvalues. J can thus be written as:

$$J = \prod_{i<j} (E_i - E_j) \cdot h(\alpha_1, ..., \alpha_{N(N-1)/2}) \tag{40}$$

Equation (30) gives the probability distribution $P(H)$ on the matrix elements H_{ij}. Using (38) and (40), we transform it to a distribution on the eigenvalues E_k and the parameters α.

$$\begin{aligned} &P(E_1, ..., E_N, \alpha_1, ..., \alpha_{N(N-1)/2}) \\ &= P(H) \cdot |J| = C \exp\Big(-\frac{1}{4\sigma^2} \sum_{i,j} H_{ij}^2\Big) \cdot |J| \\ &= C \exp\Big(-\frac{1}{4\sigma^2} \sum_{k} E_k^2\Big) \cdot \Big|\prod_{i<j} (E_i - E_j) \cdot h(\alpha_1, ..., \alpha_{N(N-1)/2})\Big| \end{aligned} \tag{41}$$

Then integrate with respect to the α. This gives us the distribution of the eigenvalues E_k:

$$\boxed{P(E_1, ..., E_N) = K \cdot \prod_{i<j} |E_i - E_j| \cdot \exp\Big(-\frac{1}{4\sigma^2} \sum_{k} E_k^2\Big)} \tag{42}$$

Using the above result, one can rigorously derive the "semicircle law" stated earlier. Essentially, one takes equation (42), and integrates out each

of the variables E_k except for one:

$$P(E_1) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} P(E_1, ..., E_N)\, dE_2 \cdots dE_N \tag{43}$$

The "semicircle law" arises as we let $N \to \infty$. The above integration was first carried out by Mehta and Gaudin, using a number of remarkable techniques, including an integration over alternating variables, rewriting the integrand as a determinant, and eventually expressing it in terms of orthogonal Hermite polynomials. We will not describe the details, but refer to [Mehta and Gaudin, 1960].

2.5 Eigenvalue Spacings and the Wigner Surmise

We can sort the eigenvalues in ascending order, and compute the spacings s between adjacent eigenvalues. Let $P(s)$ be the distribution of these spacings, in the limit of matrices of large dimension. Wigner postulated that $P(s)$ can be approximated by:

$$P(s) \approx As \exp(-Bs^2) \tag{44}$$

For the GOE, the Wigner surmise becomes:

$$P(s) \approx \frac{\pi}{2} s \exp\big(-\frac{\pi}{4} s^2\big) \tag{45}$$

This distribution is quite different from the Poisson density $\rho(s) = \exp(-s)$ (which gives the intervals between random events). Wigner's surmise approaches zero for small values of s, implying that very small spacings are unlikely, and that the eigenvalues somehow "repel" each other.

Wigner guessed equation (44) by assuming that the energy levels of a nucleus behave like a modified Poisson process. Specifically, he postulated that the probability of finding a level in the interval dx is proportional to $f(x)\, dx$ given a level at $x = 0$, where $f(x)$ is a linear function. The physical interpretation is described in [Wigner, 1956].

Using the probability distribution on the eigenvalues (equation (42)), Gaudin was able to derive an exact expression for the spacing distribution $P(s)$. The derivation is lengthy, and utilizes many novel techniques; we will only summarize it here, and refer the reader to [Gaudin, 1961] and [Mehta, 1960] for details. Essentially, we start with the following integral, which describes random matrices of dimension N:

$$P(E_1, E_2) = \int \cdots \int P(E_1, ..., E_N)\, dE_3 \cdots dE_N \tag{46}$$

The integral is taken over all values of $E_3 ... E_N$ outside the interval between E_1 and E_2 (so that E_1 and E_2 are adjacent eigenvalues). Set $E_1 = -\theta$ and

$E_2 = \theta$; and for convenience, write $N = 2m$. Mehta proves the following result:

$$2m(2m-1)P(-\theta,\theta) = C\,\frac{2m!2^{m-1}}{m!}\,\frac{d^2}{d\theta^2}\phi_m(\theta) \tag{47}$$

where

$$\phi_m(\theta) = \int_\theta^\infty \cdots \int_\theta^\infty \exp(-2(y_1^2+\cdots+y_m^2)) \cdot \prod_{i<j}^{m}(y_i^2-y_j^2)^2\, dy_1 \cdots dy_m \tag{48}$$

Define $\Psi_m(\theta) = \phi_m(\theta)/\phi_m(0)$. Let $D = 2\sqrt{2m}/\pi$, and fix:

$$\frac{s}{D} = \frac{2\theta}{D} = \frac{2t}{\pi} \tag{49}$$

Let $\Psi(t)$ be the asymptotic form of $\Psi_m(\theta)$, as $m \to \infty$, under suitable conditions. It can then be deduced that:

$$\boxed{P(s) = \frac{\pi^2}{4}\,\frac{d^2\Psi}{dt^2}} \tag{50}$$

To evaluate (48), Gaudin rewrites the integrand using a Vandermonde determinant, then expresses it as a determinant involving orthonormal wavefunctions of the harmonic oscillator, and finally applies Gram's theorem. In this way, $\Psi_m(\theta)$ becomes a Fredholm determinant of a certain integral operator.

Letting $m \to \infty$, we ultimately express $\Psi(t)$ as the Fredholm determinant of the operator

$$Tf(x) = \int_{-t}^{+t} Q(x,y)f(y)\,dy \tag{51}$$

whose kernel is

$$Q(x,y) = \frac{1}{2\pi}\left(\frac{\sin(x-y)}{x-y} + \frac{\sin(x+y)}{x+y}\right) \tag{52}$$

$\Psi(t)$ can thus be written as an infinite product:

$$\boxed{\Psi(t) = \prod_{q=0}^{\infty}\Big(1 - \frac{t}{2\pi}\gamma_{2q}^2\Big)} \tag{53}$$

where the γ_{2q} are constants which can be computed numerically. In practice, this infinite product converges quite rapidly. As it turns out, Wigner's surmise is a very good approximation to the exact spacing distribution $P(s)$; in fact, the approximation is accurate to within 5% in the region where $P(s)$ is not too small.

3 Computer Experiments

Wigner's surmise that the spacing distribution $P(s)$ can be approximated by

$$P(s) \approx As \exp(-Bs^2) \tag{54}$$

has only been proven for Gaussian ensembles. However, empirical evidence suggests that it is almost universally applicable—for instance, it appears to hold for random matrices constructed with distributions other than the Gaussian. I performed some computer experiments with sparse and band-diagonal matrices, and the results support the Wigner surmise in these cases as well.

3.1 Notes on Implementation

The experiments were carried out using the software package MATLAB, which has built-in functions to do numerical linear algebra. In particular, I used MATLAB's functions to generate random numbers, and to compute matrix eigenvalues.

There were two performance-related issues which arose: First, MATLAB provides special data structures and functions which are optimized for sparse matrices. These sparse-matrix functions appear to save memory at the expense of speed. For extremely large sparse matrices, one may have no choice but to use the sparse-matrix data structures. However, I found that the regular (non-sparse) data structures worked for matrices as large as 500×500 (the largest in my experiments); and the regular matrix functions were two to three times faster than their sparse-matrix counterparts.

Secondly, MATLAB's built-in functions are significantly faster than MATLAB code written by the user. This is partly because the built-in functions work on vectors, calculating their components "in parallel", whereas user code usually works with scalars, using a loop to calculate them in sequence.[1]

In any case, many computations will run faster if they can be posed using vectors and the built-in MATLAB functions. For instance, the following two code fragments both generate a real symmetric matrix A whose elements are Gaussian distributed:

[1] This is not entirely true, though. On most microprocessors, a built-in vector function is not really computed "in parallel"—it usually contains a small loop. It is a peculiarity of MATLAB that a built-in loop is so much faster than a user-defined loop.

```
% Generate random matrix (method 1)
for i = 1 : N
  for j = 1 : i-1
    A(i,j) = A(j,i);
    % Always have j < i
    % (copy from upper triangle to lower triangle)
  end
  for j = i : N
    A(i,j) = randn;  % Gaussian distribution
  end
end

% Generate random matrix (method 2--using vectors)
A = triu(randn(N), 0);  % Generate upper triangle and diagonal
A = A + (triu(A,1))';   % Transpose the upper triangle to form
                        % the lower triangle
```

Using method 1, MATLAB will spend as much time generating the matrices as it will computing their eigenvalues! By contrast, method 2 has a negligible running time compared to the eigenvalue computation.

3.2 Sparse Matrices

I generated sparse random matrices according to the following rule: given a small constant p, each element is chosen independently from a distribution which returns 1 with probability $p/2$, -1 with probability $p/2$, and 0 otherwise. I tested values of p ranging from 0.05 to 0.25. The resulting eigenvalues obeyed the semicircle rule, and their normalized spacings agreed with the Wigner surmise. I did not observe any unexpected behavior.

For future work, there is another class of matrices which are "more sparse" than these, namely the adjacency matrices of sparse graphs. For instance, the adjacency matrix of a k-regular graph has $O(kN)$ nonzero elements, as compared to $O(pN^2)$ nonzero elements in the matrices that I tested. These adjacency matrices also suggest a possible relationship to the theory of random graphs.

3.3 Band-Diagonal Matrices

I also tested band-diagonal matrices: matrices whose nonzero elements lie only in a "band" above and below the main diagonal. Given a parameter r, I constructed matrices with exactly r nonzero diagonals above the main diagonal, and r nonzero diagonals below it; so $r = 0$ would give a diagonal matrix, and $r = N - 1$ would give a full matrix. In my experiments, the

elements within the "band" were either i.i.d. Gaussian, or i.i.d. uniform (in the interval [-1,1]).

The results are consistent with what is predicted by the theory. By restricting ourselves to band-diagonal matrices, we impose a fairly strong condition on the structure of the matrices. The behavior of the eigenvalues depends on the value of r. At one extreme, $r = 0$, the matrix is diagonal, and the eigenvalues are simply the nonzero matrix elements; so the eigenvalues have the same distribution as the matrix elements, and the spacings follow the Poisson density $\rho(s) = \exp(-s)$. At the other extreme, $r = N - 1$, the matrix is full, and so the eigenvalues obey the semicircle law, and their spacings follow Wigner's surmise.

It is of some interest to see when the behavior of the spacings changes from the Poisson density to the Wigner surmise, because the two are qualitatively so different. According to Professor Sarnak, some recent work suggests that the change-over occurs around $r = \sqrt{N}$. My own data are roughly consistent with this; but without measuring quantitatively the change in the distributions, it is difficult to draw any reliable conclusions.

3.4 Conclusions

These computer experiments appear to support the Wigner surmise for sparse and band-diagonal matrices. Unfortunately, they do not provide any clues about *why* the Wigner surmise is so universally valid, or how one might develop a more general theory to explain it. One intriguing point, however, is the behavior of band-diagonal matrices, which varies in a natural way between the opposite extremes of diagonal and full random matrices.

4 References

It is a pleasure to thank Professor Peter Sarnak and Steve Miller for helping me learn this material, and for their comments on my work. In addition, Rebecca Lehman did related work on random matrices with Gaussian, uniform and Cauchy distributions; her paper is also available at Princeton.

Billingsley, Patrick. *Probability and Measure.* John Wiley & Sons, 1979.

*Gaudin, M. "Sur la loi limite de l'espacement des valeurs propres d'une matrice aleatoire." *Nuclear Physics* 25, p.447 (1961).

*Mehta, M.L. and M. Gaudin. "On the Density of Eigenvalues of a Random Matrix." *Nuclear Physics* 18, p.420 (1960).

*Mehta, M.L. "On the Statistical Properties of the Level-Spacings in Nuclear Spectra." *Nuclear Physics* 18, p.395 (1960).

Mehta, Madan Lal. *Random Matrices, 2nd ed.* Academic Press, 1991.

*Porter, C.E. and N. Rosenzweig. "Statistical Properties of Atomic and Nuclear Spectra." *Ann. Acad. Sci. Fennicae* A-VI, No. 44 (1960).

Statistical Theories of Spectra: Fluctuations. (A collection of reprints, with an introuduction by Charles E. Porter.) Academic Press, 1965.

Tracy, Craig A., and Harold Widom. *Introduction to Random Matrices.* (A collection of notes.)

*Wigner, E.P. "Results and Theory of Resonance Absorption." Conference on Neutron Physics by Time-of-Flight, Gatlinburg, Tennessee, Nov. 1 and 2, 1956. *Oak Ridge Nat'l. Lab. Report* ORNL-2309, 1957, p.59.

*Wigner, E.P. "Statistical Properties of Real Symmetric Matrices with Many Dimensions." *Can. Math. Congr. Proc.*, p.174. Univ. of Toronto Press, 1957.

Papers marked with an asterisk (*) are included in *Statistical Theories of Spectra: Fluctuations*, which is also listed above.