Lectures on Random Matrices

Tiago Pereira

Department of Mathematics
Imperial College London, UK

Preface

Lecture notes for a short course on the school for complex systems in Sao Jose, Brazil. The goal of these lecture is to expose the student to the main concepts and tools of random matrices. This short course consists of a few lectures to students of various backgrounds. Therefore, I have chosen to include many elementary examples throughout the text. I tried to combine heuristic arguments and to communicate the main ideas. There are beautiful connections between many branches of mathematics, and the theory is aesthetically very pleasing. I tried to capture some of these connections and to highlight the main points, however, often I chose to sacrifice the precision of the statements. I also don't present the proofs, but I will offer the reference. I'm in debt with Alexei Veneziani and Daniel Maia for his critical reading of the text and for the exercises.

Contents

Motivations

In the mid 50s a large number of experiments with heavy nuclei was performed. These heavy atoms absorb and emit thousands of frequencies. So an experiment of this kind offers us a great number of differences in the energy levels, and difficult to find the set of levels behind the given differences.

In fact, it was virtually impossible to know the levels energy exactly and label them according to good numbers quantum. To tackle this problem one is required to understand the eigenvalue problem

$$H\psi_i = E_i\psi_i$$

where H is the Hamiltonian of the system, and E_i is the energy levels along with the eigenfunctions ψ_i. Not surprisingly, writing the Hamiltonian H is already a hard problem, as there are hundreds of nucleons involved. This large systems are typically non-integrable, so solving the eigenvalue problem is undoable.

Wigner and Dyson were the first to attack the problem through a statistical point of view. Instead of searching an approximate solution for the nuclear system, they focused on the distribution of energy levels. Dyson [1] summarizes the motivation behind the use of statistical methods:

The statistical theory will not predict the detailed sequence of levels in any one nucleus, but it will describe the general appearance and the degree of irregularity of the level structure, that is expected to occur in any nucleus which is too complicated to be understood in detail.

This view led Wigner to develop a theory based on random matrices for explaining the distribution of the energy levels [2]. Wigner assumed that the detailed knowledge of the system would not be relevant for statistical description of the system. Starting from these assumptions Wigner proposed the description of properties of a heavy nucleus through an ensemble random matrices, where the entries (elements) of the matrix would be independently chosen following a distribution. Additional system information could be obtained through the inherent symmetries, for example, invariance under time translation invariance and rotational. Such symmetries would place a distinct ensembles matrices.

This approach was indeed very successful. Much of the motivation for the study of random matrices comes from the fact that once removed the dependent part of the model used, the correlation of levels of different systems exhibit universal features in a variety of physical situations [3, 4].

Today, random matrices have a wide range of applications starting particle physics elementary [5] covering quantum hydrodynamics with applications in fluid Hele-Shaw [6] and applications detection of epilepsy [7]. Another important problem that can be addressed using the theory of random matrices is the emergence of collective behavior in complex networks [8, 9].

Zeros of the Riemann zeta function: Lets discuss this interesting example. Recall that the Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \ \ \text{Re } s > 1,$$

has a meromorphic continuation to all $\mathbb{C}$, and apart from the trivial zeros at $z = -2n$, on the negative axis, all the other zeros are conjectured to lie on the critical line Re $s = 1/2$, this is precisely the Riemann hypothesis: The non-trivial zeros zeros lie on the line Re$s = 1/2$, that is, $s = 1/2 + i\gamma$.

A surprising observation is that the random matrix theory describe the distribution of the non-trivial zeros of ζ. Assuming the Riemann hypothesis, Montgomery rescaled the imaginary parts of the zero

$$\gamma_j \to \tilde{\gamma}_j = \frac{\gamma_j \log \gamma_j}{2\pi},$$

to have a mean spacing of 1

$$\frac{\#\{j \geq 1 : \tilde{\gamma}_j < T\}}{T} \to 1.$$

He then obtained an expression for pairs of zeros

$$R(a,b) = \lim_{n \to \infty} \frac{1}{n} \#\{\text{ pairs } (j_1, j_2) : 1 \leq j_1, j_2 \leq n, \tilde{\gamma}_{j_1} - \tilde{\gamma}_{j_2} \in (a,b)\}$$

for any interval (a,b). Montgomery gave a talk in Princeton about his results, and Dyson could not attend the talk. However, they sope later and Montgomery went on explaining that he wants to obtain an expression for the pairs of zeros. Dyson then asked whether he found

$$R(a,b) = \int_a^b \left(1 - \left(\frac{\sin \pi u}{\pi u}\right)^2\right) du$$

which was precisely Montgomery's results... Wait a bit, how come? Dyson explain that this is what one should obtain if the zeros were to behaving as the eigenvalues of the GUE, see Ref. [10] for details.

Lecture 1: Wigner Semi-Circular Law

First lets fix some notation. Recall that a matrix $H = (H_{ij})_{i,j=1}^n$ is Hermitian if and only if

$$H = H^\dagger,$$

where † stands for the transpose conjugate. In terms of the matrix elements the hermitian properties reads

$$H_{ij} = \bar{H}_{ji},$$

where $\bar{\cdot}$ stands for the complex conjugate. If we need to explicit the real and complex components of the elements we denote

$$H_{ij} = H_{ij}^R + iH_{ij}^I,$$

where H_{ij}^R is the real part and H_{ij}^I the complex part.

A particular case is real symmetric matrices. A matrix H is real symmetric if and only if all its entries are real and

$$H = H^T,$$

where ^{T} stands for the transpose.

Exercise 1. *Let H be a Hermitian matrix. Show that all eigenvalues of H are real*

The spectral theorem for Hermitian matrices states that

Theorem 1. *Let H be a Hermitian matrix. Then, there exists an orthonormal basis consisting of eigenvectors of H, and each eigenvalue is real. Moreover, H admits the decomposition*

$$H = U\Lambda U^\dagger$$

where U is the matrix of eigenvectors and $\Lambda = diag\,(\lambda_1, \cdots, \lambda_n)$ is the matrix of eigenvalues. And

$$UU^\dagger = U^\dagger U = 1,$$

that is, the matrix U is unitary.

Hence, Hermitian matrices can be decomposed in terms of its spectral coordinates Now we are ready to define our object of study

Definition 1. *A Wigner matrix ensemble is a random matrix ensemble of Hermitian matrices $H = (H_{ij})_{i,j=1}^n$ such that*

- *the upper-triangular entries H_{ij}, $i > j$ are iid complex random variables with mean zero and unit variance.*

- *the diagonal entries H_{ii} are iid real variables, independent of the upper-triangular entries, with bounded mean and variance.*

Example 1 (Real-Symmetric Matrices)**.** *As we discussed real symmetric matrices are a particular case of Hermitian matrices. Lets see how the Wigner ensemble takes form for 2 by 2 matrices. Any two by two real symmetric matrix has the form*

$$H = \left(\begin{array}{cc} a & b \\ b & c \end{array} \right).$$

To have a Wigner ensemble we impose that a, b and c are independent and identically distributed. For example, they could be distributed according to a Gaussian with zero mean and variance 1. The collection of all these matrices will form the Wigner Ensemble.

There are many statistics of the Wigner ensemble one wishes to consider such as the eigenvalues. Of particular interest is the operator norm

$$\|H\| := \sup_{x \in \mathbf{C}^n : |x| = 1} |Hx|$$

where $|\cdot|$ is a vector norm. This is an interesting quantity in its own right, but also serves as a basic upper bound on many other quantities. For example, all eigenvalues $\lambda_i(H)$ of H have magnitude at most $\|H\|$. Because of this, it is particularly important to get good understanding of the norm.

Theorem 2 (Strong Bai-Yin theorem, upper bound)**.** *Let h be a real random variable with mean zero, variance 1, and finite fourth moment, and for all $1 \leq i \leq j$, let H_{ij} be an iid sequence with distribution h, and set $H_{ji} := H_{ij}$. Let $H := (H_{ij})_{i,j=1}^n$ be the random matrix formed by the top left $n \times n$ block. Then almost surely one has*

$$\limsup_{n \to \infty} \frac{\|H\|}{\sqrt{n}} \leq 2.$$

This means that operator norm of H is typically of size $O(\sqrt{n})$. So it is natural to work with the normalised matrix $H/\sqrt{n}$.

The Semi-Circular Law: A centerpiece in random matrix theory is the Wigner semi-circle law. It is concerned with the asymptotic distribution of the eigenvalues

$$\lambda_1 \left(\frac{H}{\sqrt{n}} \right) \leq \ldots \leq \lambda_n \left(\frac{H}{\sqrt{n}} \right)$$

of a random Wigner matrix H in the limit $n \to \infty$.

It appears that the histogram of eigenvalues, called the density of eigenvalues, converges to a deterministic shape. In fact, this is true. The density of eigenvalues of any Wigner matrix has a limiting distribution known as Wigner's semicircle law:

$$\mu'_{sc}(x) := \frac{1}{2\pi}(4 - |x|^2)_+^{1/2} dx,$$

where $(x)_+ = x$ if $x > 0$ and 0 otherwise

Example 2. *Let A be a $n \times n$ random matrix with real entries. Notice that*

$$H = \frac{A + A^T}{2}$$

is a real symmetric matrix. In matlab the command

```
n = 2000;
A = (randn(n)/sqrt(n))*sqrt(2);
H = (A+A')/2;
```

generates a real symmetric random matrix with Gaussian entries. We plot the distribution of eigenvalues we can use the command

```
d = eig(H);
[f,x] = hist(d,50);
bar(x,f/trapz(x,f))
```

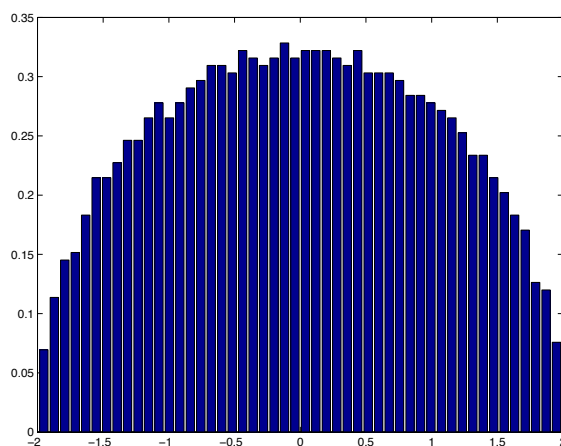*The result is well described by the semi-circle law. This would this need a proper normalization.*

Figure 1: Distribution of Eigenvalues

We will state two results on this convergence.

Theorem 3 (Wigner's Semicircle Law). *Let H_n be a sequence of Wigner matrices and I an interval. Then introduce the random variables*

$$E_n(I) = \frac{\#\{\lambda_j(H/\sqrt{n}) \in I\}}{n}. \tag{1}$$

Then $E_n(I) \to \mu_{sc}(I)$ in probability as $n \to \infty$.

Wigner realised that one can study the behavior of the random variables $E_n(I)$ without computing the eigenvalues directly. This is accomplished in terms of a random measure, the empirical law of eigenvalues.

8

Definition 2. *The empirical law of eigenvalues μ_n is the random discrete probability measure*

$$\mu_n := \frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_j(H/\sqrt{n})}.$$

Clearly this implies that for any continuous function $f \in C(\mathbb{R})$ we obtain

$$\int f d\mu_n = \frac{1}{n} \sum_{j=1}^{n} f(\lambda_j) \tag{2}$$

For the matrix ensemble the corresponding function μ_n is now a random measure, i.e. a random variable taking values in the space of probability measures on the real line. The semicircle law first proved by Wigner states that the eigenvalue distribution of the normalized matrices converges in probability as $n \to \infty$ to a non random distribution

$$\mu_n \to \mu_{sc}.$$

This last statement can be slightly confusing. The sequence of random measures μ_n (in the space of probability measures in the real line) converge in probability (resp. converge almost surely) to a deterministic limit, which is a deterministic probability measure! The precise statement is the following:

Theorem 4 (Wigners Semicircle Law). *Let H_n be a sequence of Wigner matrices. Then the empirical law of eigenvalues μ_n converges in probability to μ_{sc} as $n \to \infty$.*

Precisely, for any continuous bounded function f and each $\varepsilon > 0$,

$$\lim_{n \to \infty} \mathbb{P} \left(\left| \int f d\mu_n - \int f \mu_{sc} \right| > \varepsilon \right) = 0.$$

Comments about the proof: There are two basic schemes to prove the theorem, the so-called Moment approach and the resolvent approach. The classical Wigner method concerns with the moments. We will brief discuss the main ideas on this technique.

Exercise 2. *Without loss of generality (why?), let $f \in C(\mathbb{R})$ be a polynomial. Show that*

$$\int f d\mu_n = \frac{1}{n} \operatorname{Tr} f(H)$$

Hint: Use that

$$
\begin{aligned}
\int f d\mu_n &= \frac{1}{n} \sum_{j} f(\Lambda_{jj}) \\
&= \frac{1}{n} \operatorname{Tr} f(\Lambda) \\
&= \frac{1}{n} \operatorname{Tr} U^{\dagger} f(\Lambda) U \\
&= \frac{1}{n} \operatorname{Tr} f(H) \tag{3}
\end{aligned}
$$

In this formulation, we can use the spectral theorem to eliminate the explicit appearance of eigenvalues in the law. This is done using our last exercise. Consider the moments

$$M_j^n = \mathbb{E} \int_{\mathbb{R}} \lambda^j d\sigma_n(\lambda)$$

Notice that

$$M_j^n = \mathbb{E} \frac{1}{n} \operatorname{Tr} A^j.$$

After a long labor, one derives the relations

$$\lim_{n \to \infty} M_j^n = m_j = \begin{cases} t_k & \text{if } j = 2k \\ 0 & \text{if } j = 2k+1 \end{cases}$$

where $k \in \mathbb{N}$, and t_k's are given by the recurrence relation

$$t_0 = 1 \quad \text{and} \quad t_k = \sum_{j=0}^{k-1} t_{k-1-j} t_j$$

Actually, one can obtain

$$t_{2k} = C_k := \frac{1}{k+1} \binom{2k}{k}$$

The numbers C_k are the Catalan numbers. These are precisely the moments of the semicircle law

Exercise 3. *Let μ_{sc} be the semicircle law defined above. Let*

$$m_k = \int x^k \mu_{sc}(dx)$$

By symmetry, $m^{2k+1} = 0$ for all k. Use a trigonometric substitution to show that

$$m_0 = 1 \quad \text{and} \quad m_{2k} = \frac{2(2k-1)}{k+2} m_{2(k-1)}.$$

This recursion completely determines the even moments; show that, in fact,

$$m_{2k} = C_k$$

Hence, the moments are the same, and one can show that this is equivalent to the semi-circle law (by the problems of moments).

A neat example: Graph Theory

Consider random graphs of n nodes of labelled undirected graphs of n nodes. We will use a random graph model and terminology from references [11, 12]. This model is an extension of the Erdös-Rényi model for random graphs with a general degree distribution. The model consists in to prescribing the expected values of the node degrees. For convenience, any given sequence of expected degrees

$$\boldsymbol{w}_n = (w_1, w_2, \cdots, w_n).$$

We consider thus an ensemble of random graphs $G(\boldsymbol{w}_n)$ in which an edge between nodes i and j is independently assigned with success probability

$$p_{ij} = \frac{w_i w_j}{\sum_{k=1}^n w_k}.$$

In order to ensure that $p_{ij} \leq 1$, we assume that $\boldsymbol{w}_n$ is chosen so that $\left(\max_{1 \leq k \leq n} w_k\right)^2 \leq \sum_{k=1}^n w_k$.

A realisation of a graph in the ensemble $G(\boldsymbol{w}_n)$ is encoded in the adjacency matrix $A = (A_{ij})$ with $(0, 1)$-entries determining the connections among nodes of the graph. The degree κ_i of the ith node is the number of connections that it receives:

$$k_i = \sum_{j=1}^n A_{ij}.$$

Notice that κ_i is a random variable whose expected value is exactly the prescribed quantity w_i. In particular, $w_1 = \max_{1 \leq i \leq n} w_i$ is the largest expected value of a degree.

Now consider the the combinatorial Laplacian

$$L = D - A$$

where $D = \text{diag}(k_1, \cdots, k_n)$ is the matrix of degrees. This matrix is important for collective dynamics in networks [13, 14] and for counting the number of spanning trees. Now consider the normalised Laplacian

$$\mathcal{L} = 1 - D^{-1/2} A D^{-1/2}$$

this controls the isoperimetrical properties and the mixing rates of a random walk on a graph [12]. For graphs with uneven degrees, the above three matrices can have very different distributions.

The eigenvalues of the normalised Laplacian $\mathcal{L}$ satisfy the semicircle law under the condition that the minimum expected degree is relatively large (much larger than the square root of the expected average degree).

$$\max_{i \neq 0} |\lambda_i - 1| \leq [1 + o(1)] \frac{1}{\sqrt{\langle w \rangle}} + O(1/w_{\min})$$

Stability of Species

An important question in ecology is the stability of a given food web. In particular, it is important to understand how the stability of the food web depends on the number of species and the structure of interaction.

In early 70's, Gardner and Ashby suggested that large complex systems which are randomly connected may become unstable when the number of interacting species increase [15]. In a follow up, Robert May proposed an explanation using random matrices [16]. Lets discuss May's ideas here.

In an ecological system these are populations of n interacting species, which in general obey non-linear differential equations. However, the stability of the stable configurations is determined by the linearised dynamics around such configurations. Performing a linearisation one obtains the linear equation

$$\frac{dx}{dt} = Ax,$$

where x is a vector of disturbed populations x_j, and A a $n \times n$ interaction matrix with elements a_{jk} which characterise the effect of species k on species j near equilibrium.

Assume that in the absence of interaction disturbances are damped. That is, the system is stable. We can choose then $a_{ii} = -1$. This sets a time scales in which the disturbances decay. Then, the interactions are switched on. We also assume that the strength of interaction from species k to j is the same as j to k. That is, the matrix A is symmetric [1]. Moreover, the independent elements of A are is equally likely to be positive and negative and independent. The matrix elements is assigned from a distribution of random numbers, and this distribution has mean zero and mean square value α.

Here α is thought of as expressing the average interaction strength. May assumed no symmetry in the matrix interaction. Hence,

$$A = \alpha H - I$$

where H is a Wigner matrix. System is stable if and only if all the eigenvalues of A have negative real parts. The eigenvalues are $\lambda(A) = \alpha\lambda(H) - 1$. In particular we know that the largest eigenvalue of A must be

$$\lambda_{\max}(A) \leq 2\alpha\sqrt{n} - 1$$

Hence we obtain

$$\alpha < \frac{2}{\sqrt{n}}$$

Roughly speaking this suggests that within a web species which interact with many other should do this weakly.

---

[1] We make this assumption for simplicity as it is not quite realistic to think of a symmetric ecosystem.

Lecture 2: Gaussian & Unitary Ensembles

Lets for a moment consider the Wigner ensemble of orthogonal matrices with iid elements satisfying $\mathbb{E}(H_{ij}) = 0$ and $\mathbb{E}(H_{ij}^2) = 1$.

Exercise 4. *Show that independence implies that*

$$\mathbb{E}(H_{ij}H_{pq}) = \delta_{ij}\delta_{jq} + \delta_{ip}\delta_{jp}.$$

Lets assume for a moment that all elements has a Gaussian distribution From these assumptions on the matrix elements it is possible to obtain a probability density in the space of real-symmetric matrices. That is, the probability $P(H)dH$ that an symmetric matrix H lies in a small parallelopiped dH is Gaussian. The volume of the parallelopiped is given in terms in the independent coordinates

$$dH = \prod_{k \leq j} dH_{kj}.$$

Then, using the independence we can compute

$$
\begin{aligned}
P(H) &= \frac{1}{Z} \prod_{j<i} e^{-H_{ij}^2/2} \prod_{i} e^{-H_{ii}^2/2} \\
&= \frac{1}{Z} \exp\left\{ -\frac{1}{2}\left(\sum_{j<i} H_{ij}^2 + \sum_{i} H_{ii}^2 \right) \right\}
\end{aligned}
$$

where Z is the normalization constant. Recall that $H_{ij} = H_{ji}$, hence

$$P(H) = \frac{1}{Z} \exp\left\{ -\frac{1}{4} \sum_{i,j=1}^{n} H_{ij}^2 \right\}$$

or one can write this as a trace

Exercise 5. *Show that the above P(H) can be written as*

$$P(H) = \frac{1}{Z} \exp\left\{ -\frac{1}{4} \operatorname{Tr} H^2 \right\} \tag{4}$$

This means that Gaussianity in the matrix elements leads to a probability density represented a the trace of the matrix squared. Next, we will see that is this the case regardless the form of the distribution of the elements.

Orthogonal and Unitary Ensembles

Definition 3. *The Gaussian unitary ensemble GOE(n) is described by the Gaussian measure*

$$P(H)dH = \frac{1}{Z_o} e^{-\frac{n}{2} \operatorname{Tr} H^2} dH \tag{5}$$

on the space of $n \times n$ real symmetric matrices $H = (H_{ij})_{i,j=1}^{n}$.

Here, Z_o is a normalization constant, chosen so that the integral against the measure is equal to one. This ensemble can be redefined in terms of the matrix elements. The $GOE(n)$ defined on the space of real-symmetric matrices is characterized by the following properties

1. The ensemble is invariant under every transformation

$$H \mapsto W^t H W$$

   where W is any orthogonal matrix, that is $W^T W = W W^T = 1$.

2. The elements H_{kj} with $k \leq j$ are statistically independent. This means that the density $P(H)$ will be decomposed into a product form

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

Theorem 5. *The conditions* 1. *and* 2. *on the matrix elements uniquely determine the Gaussian measure $P(H)dH$, up to shifts $H \to H + const$ and a rescaling. This is regardless the form of the distribution (with finite variance).*

The proof is rather lengthy and can be found in Ref. [17] Therefore, invariance under orthogonal transformations and the statistical independence of matrix elements completely determine the Gaussian orthogonal ensemble. The motivation for the name of the ensemble becomes from the computation we just performed.

Exercise 6 (Why this Gaussian Ensemble?). *Let $H = U \Lambda U^T$, show that*

$$P(H) = P(\Lambda)$$

which implies that that P is a Gaussian density in the spectral coordinates

$$P(H) = \frac{1}{Z} \exp \left\{ -\frac{n}{2} \sum_i \lambda_i^2 \right\}$$

Hint: Use that $\operatorname{Tr} ABC = \operatorname{Tr} BCA$.

Remark 1. *Notice the factor n difference between the deduced Eq. (4) and the model Eq. (5). There are a few ways to see the rescaling*

– *Rescaling of the matrices $H \mapsto H/\sqrt{n}$. So we rescale the eigenvalues*

$$\frac{\lambda(H)}{\sqrt{n}} = \tilde{\lambda}\left(\frac{H}{\sqrt{n}}\right)$$

for any eigenvalue λ.

– *Rescaling the standard deviation*

$$\sigma \mapsto \frac{\sigma}{\sqrt{n}}$$

Notice this rescaling leads to the factor n in front of the trace, and since $\mathbb{E}H_{ij} = 0$, this variance rescaling corresponds precisely to the matrix rescaling.

This means that our definition of the ensemble already accounts for the normalisation. It is possible to show that that the volume element is invariant under real orthogonal transformation [17, 18]. Moreover, the invariance of $P(H)$ imposes that

$$P(H')dH' = P(H)dH$$

where $H' = W^t HW$, and W is any orthogonal matrix. This equation shows that this probability view by another observer (via W) the matrices in H realizing the same operators fill a parallelopiped dH'. The second observer computes the probability using the same laws as the first observer and obtains the same result.

Lets now introduce the unitary ensemble. Roughly speaking, the only differences are: i) we deal with Hermitian matrices which are complex as opposed to real symmetric, ii) ensemble invariance is with respect to unitary transformations, which can be think of as the generalisation of orthogonal transformations to Hermitian matrices.

Definition 4. *The Gaussian unitary ensemble GUE(n) is described by the Gaussian measure*
$$P(H) = \frac{1}{Z_u} e^{-\frac{1}{2}\operatorname{Tr} H^2}$$
on the space of $n \times n$ Hermitian matrices $H = (H_{ij})_{i,j=1}^n$.

Here, Z_u is a normalization constant, chosen so that the integral against the measure P is equal to one. The Gaussian unitary ensemble can also be redefined in terms of properties of the matrix elements by the following properties

U1 the ensemble is invariant under unitary transformations

$$H \mapsto U^t HU,$$

where U is a unitary matrix $U^\dagger U = UU^\dagger = 1$.

U2 linearly independent elements of H are statistically independent. This means that is written as products of functions which depend only on the independent matrix elements.

Notice that the volume element is also given in terms of the independent entries.

$$dH = \prod_{k \le j} dH_{kj}^R \prod_{k < j} dH_{kj}^I.$$

The conditions on the matrix elements U1 and U2 also completely charaterize the GUE. Similar to the GOE case, these conditions on the matrix elements yield the Guassian measure uniquely P, up to the the shifting and rescaling. This has a nice consequence: regardless the specific details of the distribution of the matrix elements the statistical properties of the ensemble is the same.

General Unitary Ensembles:

We will discuss some properties of the unitary ensembles. However, it makes more sense and yet it yields more interesting results to consider slightly more general classes of ensembles than just the GUE. Lets just rephrase our observations. Consider the polynomial

$$V(x) = \sum_{j=2}^{k} a_j x^j$$

and define the extension of this polynomial as a matrix function

$$V(H) = \sum_{j=2}^{k} a_j H^j.$$

Then, we can definite the unitary random matrix unitary ensemble

Definition 5. *The unitary ensemble is described by the measure*

$$P(H)d(H) = \frac{1}{Z} e^{-n\,\mathrm{Tr}\,V(H)} dH.$$

in the space of Hermitian matrices.

This ensemble is indeed invariant under unitary transformation. For any unitary matrix W we have that

$$dH' = dH$$

where $H^\dagger = WHW^\dagger$. Moreover, because the trace is cyclic $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA)$, for any matrices A, B and C, we obtain that

$$P(H) = P(H').$$

In particular, $P(H) = P(\Lambda)$, where Λ is the matrix of eigenvalues.

Our discussion up to now revealed *statistical independence* of the matrix elements leads uniquely to $V(H) = H^2$. For this observations we are guarantee that polynomials of degrees higher than 2 (or real-analytic functions) V will correspond the matrix ensembles with correlations in the matrix elements. We will take a close look in these models.

Joint Probability of Eigenvalues

In the last lecture we saw how the independence of the matrix entries and invariance leads to the Guassian measure. Our goal in this lecture will be to obtain a equation for the joint probability of eigenvalues from this probability measure.

The idea here is to write the matrix in its spectral form

$$H = U\Lambda U^\dagger \tag{6}$$

where U is the matrix of eigenvectors and Λ the matrix of eigenvalues. Notice that the probability $P(H)dH$ is given in terms of the matrix elements. One could try to express $P(H)dH$ in terms of the pair (U, Λ) and integrate over the variables U. At the end (of this laborious process) one is left with a probability depending only on the eigenvalues Λ. As we discussed in the previous lecture $P(H) = P(\Lambda)$, that is, this density does not depend on the eigenvector variables. On the other hand, we will discuss that the volume transforms as

$$dH = \prod_{i<j} |\lambda_j - \lambda_i|^2 d\Lambda dU$$

Combining these two results we obtain that in the spectral coordinates (Λ, U) the unitary ensemble reads

$$
\begin{aligned}
P(H)dH &= P(U\Lambda U^\dagger)d(U\Lambda U^\dagger) \\
&= P(\Lambda)Jd\Lambda dU \\
&= \frac{1}{Z}e^{-n\sum_i V(\lambda_i)} \prod_{i<j} |\lambda_i - \lambda_j|^2 d\Lambda dU.
\end{aligned}
$$

Integrating over the variables U we are left with a joint density of eigenvalues. The main result of the section is the following

Claim 1. *The joint distribution of eigenvalues is given by*

$$P(\lambda_1, \ldots, \lambda_n) = \frac{1}{Z}e^{-n\sum_i V(\lambda_i)} \prod |\lambda_i - \lambda_j|^2$$

The contribution $\prod |\lambda_i - \lambda_j|^2$ came from the volume element dH. This terms will play a major role in the analysis. And it is the main responsible for the beautiful behaviour of the ensemble such as universality in the eigenvalue statistics. To be more clear I should have change the letter for the joint probability of eigenvalues as it may be confusing since P used to denote the Gaussian density in the space of Hermitian matrices. However, this is the standard notation.

Claim 2. *For the GOE(n) the joint distribution of eigenvalues*

$$P(\lambda_1, \ldots, \lambda_n) = \frac{1}{Z}e^{-n\sum_i V(\lambda_i)} \prod |\lambda_i - \lambda_j|$$

We will discuss the main ideas of this construction. In mathematical language we wish to parametrize the probability in terms of the spectral coordinates. Consider the mapping from the spectral variables to the matrix elements

$$(\Lambda, U) \mapsto H = U \Lambda U^\dagger$$

This spectral parametrisation must be a bijective change of variables. There are two obstacles to this: $a)$ if there is a multiple eigenvalue then there is no canonical way to select orthonormal eigenvectors from the corresponding eigenspace. And $b)$ eigenvectors are is determined only up to an arbitrary sign or phase.

For the eigenvectors, we could simply say: we don't care about the phase (or sign), that is, we only consider eigenvectors up to an arbitrary change in the phase. It turns out, that this can be actually done in a rigorous way in terms of cosets in the unitary group [10, 17]. So I won't develop this ideas here. It is possible to overcome the challenge $a)$ if we take a measure theoretic perspective. It is possible to show that almost every hermitian matrix one picks *at random* will have distinct eigenvalues. This means that for all what matters the coordinate change can be done. I wish to be more precise here.

Recall that the map must be injective and surjective. By the spectral theorem, we can always write a Hermitian matrix in its spectral representation 6. This means that the map is surjective. It remains to show that the map is

Let $\mathcal{H}(n)$ be the space of hermitian matrices

$$\mathcal{H}(n) := \{H \in \mathbb{C}^{n \times n} : H^\dagger = H\}$$

Moreover, let $\widehat{\mathcal{H}}(n) \subset \mathcal{H}(n)$ be the subset of hermitian matrices with distinct eigenvalues. The first major observation is the following

Lemma 1. $\widehat{\mathcal{H}}(n) \subset \mathcal{H}(n)$ *is open, dense and has full measure.*

I won't present the proof of this claim here, see [10, 17, 18]. Hence, considering the eigenvectors coordinates (up to phases) and the eigenvalues is an honest parametrisation in $\widehat{\mathcal{H}}(n)$. Moreover, since one performs an integration to obtain Claim 1 the zero measure set of non simple eigenvalues plays no role.

Performing the change of variables, one obtains the Jacobian of the transformation:

$$dH = J(H)d\Lambda dU$$

As I discussed before the volume element is invariant under unitary similarity transformations. This implies that

$$dH = d(QHQ^*) = J(QHQ^*)d\Lambda d(QU) = J(QHQ^*)d\Lambda d(U)$$

Therefore $J(H) = J(QHQ^*)$, hence, $J(H)$ is a symmetric function of the eigenvalues of H only. Thus we can write $J(H) = J(\Lambda)$.

Proposition 1.

$$J(\lambda_1, \ldots, \lambda_n) = \prod_{i<j} |\lambda_j - \lambda_i|^2$$

Lecture 3: Universality

Physical systems in equilibrium obey the laws of thermodynamics. So, whatever the precise nature of the interaction between the particles at a detailed level, at the macroscopic level, physical systems exhibit universal behavior – they are all governed by the same thermodynamical laws. The random matrix theory share the same similarities.

Eigenvalues as a Gas of electrons

Note that we can rewrite the joint distribution of eigenvalues as

$$P(\lambda_1, \cdots, \lambda_n) = Z_n^{-1} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^2 \, e^{-n \sum_{i=1}^n V(\lambda_i)} \tag{7}$$

$$= Z_n^{-1} \exp \left\{ -\left(\sum_{1 \leq i \neq j \leq n} \log |\lambda_i - \lambda_j|^{-1} + n \sum_{i=1}^n V(\lambda_i) \right) \right\}$$

Those educated in statistical mechanics may be overjoyed. The normalisation constant Z is the same as the partition function of a gas of electrons in a plane when the electrons are confined to a one-dimensional straight wire under a quadratic potential. Lets explore this analogy for two reasons, first it is very nice, and second it will give us an nice intuition about the eigenvalues.

Just recall that the Couloumb potential in the plane is logarithmic (to see this you either solve the poisson equation in two dimensions or solve the Maxwell equations with a infinitely long wire passing through the origin). Hence that eigenvalues are charges of the same sign on a 2-d Coulomb potential. Therefore, the interaction between eigenvalues is repulsive. This is know as repulsion of eigenvalues.

If one consider the coulomb term alone the eigenvalues would then scape to infinity. Here is where the function V plays a role of a confining potential. And hence, the name. This gives a good reasoning for having the size n multiplying the potential, it provide a balance between coulomb and electrostatic. Otherwise, the equilibrium configuration could be found.

Variational Problem and Equilibrium Configurations: Lets just explore a bit more the relation the similarities between our joint distribution of electrons and the Coulomb gas. Notice that we could rewrite the joint distribution as

$$P(\lambda_1, \cdots, \lambda_n) = Z_n^{-1} e^{-n^2 I^V(\mu_n)} \tag{8}$$

where

$$\mu_n(\lambda) = \frac{1}{n} \sum_{i=1}^n \delta(\lambda - \lambda_i) \tag{9}$$

is the empirical law of eigenvalues and

$$I^V(\mu) \equiv \int (V(z) + U^\mu(z)) \, d\mu(z) \tag{10}$$

is the "electrostatic energy " associated with μ_n with

$$U^\mu(z) = \int \log |z - w|^{-1} d\mu(w) \tag{11}$$

being the logarithmic potential.

Remark 2. *We see this equation as the energy functional of n charges distributed according to the empirical law μ_n under the influence of an external potential V. Thus, the first contribution of (10) is the interaction energy of the n charges with the potential V and the second contribution comes from the of the Coulomb interaction energy between charges.*

Note that the scale n in the the potential V in (15) has an important effect, and balances the interaction energy with the potential of the chargers and the energy of interaction between the charges appear the same order n : this is the physical argument which justifies the scale of the eigenvalues of H for the support empirical law has a compact support as $n \to \infty$.

We can then turn this problem to a variational problem. The leading contribution to the integrals with respect to (8) in the thermodynamic limit $n \to \infty$ is given by the following variational problem:

$$E^V \equiv \inf_{\mu \in \mathcal{M}(\mathbb{R})} I^V(\mu), \quad E < \infty \tag{12}$$

where the infimum is taken over the set $\mathcal{M}(\mathbb{R})$ of all probability measure on $\mathbb{R}$. If such probability measure exists μ^V for the potential V

$$E^V = I(\mu^V),$$

we say that μ^V is the equilibrium measure associated with the potential V.

Notice that μ^V is the equilibrium distribution of the charges in the thermodynamic limit $n \to \infty$) and at the zero temperature limit ($T^{-1} \to \infty$) of a bi-dimensional Coulomb gas at a temperature $T = N^{-1}$.

Orthogonal Polynomials and Integral Kernel

There is a beautiful and extremely useful connection between the theory of orthogonal polynomials and random matrices. We can write the joint probability of eigenvalues as determinant

$$P_n(\lambda_1, \cdots, \lambda_n) = \frac{1}{n!} \det \left(K_n(\lambda_i, \lambda_j) \right)_{i,j=1}^n \tag{13}$$

of the integral kernel

$$K_n(x, y) = e^{-\frac{n}{2} V(x)} e^{-\frac{n}{2} V(y)} \sum_{j=1}^n \phi_j^n(x) \phi_j^n(y) \tag{14}$$

where $\left\{\phi_j^n\right\}_{j=1}^n$ is a set of orthogonal polynomials with respect to the weights

$$d\nu\left(x\right) = e^{-nV(x)}dx.$$

The demonstration is somewhat laborious and we won't present it here. Please, see Ref. [17].

Correlation Functions: We can introduce the n-point correlation function associated with P_n in a similar manner as in statistical mechanics

$$R_k^n(\lambda_1,\cdots,\lambda_k) = \frac{n!}{(n-k)!}\int P_n(\lambda_1,\cdots,\lambda_n)\prod_{i=k+1}^n d^2\lambda_i.$$

A remarkable result reveals that the $n-$point correlation function can also be represented as a determinant of the Integral kernel

$$R_k^n(\lambda_1,\cdots,\lambda_k) = \det\left(K_n\left(\lambda_i,\lambda_j\right)\right)_{i,j=1}^k.$$

Therefore, the asymptotic behavior of of the integral kernel K_N will dictate the statistics of the eigenvalues of the unitary ensemble as $n \to \infty$.

Universality – Heuristics

The universality in the eigenvalue statistics can be heuristically understood in terms of the Dyson interpretation of the eigenvalues as a Coulomb gas [17]. For a hermitian matrix the joint probability of the eigenvalues reads as

$$P_n(\lambda_1, \cdots, \lambda_n) \propto \exp\left\{ \sum_{1 \leq i \neq j \leq n} \ln|\lambda_i - \lambda_j|^{-1} - n \sum_{i=1}^{n} V(\lambda_i) \right\}. \tag{15}$$

As we explored last section, this joint distribution of eigenvalues has the same form as a 2 dimensional Coulomb gas restricted to one dimension. The eigenvalue interaction is given by a Coulomb term $\ln|\lambda_i - \lambda_j|^{-1}$ under a potential $V(\lambda)$ that confines the eigenvalues in a bounded set of the real line.

Since the eigenvalues are restricted to one dimension, if an eigenvalue is located between two eigenvalues the Coulomb potential does not allow it to switch its position with the surrounding eigenvalues. As the number of eigenvalues increase (the size of the matrix) the eigenvalues become closer and the effect of the Coulomb potential overtakes the potential $V(\lambda)$ (due to the logarithm behavior near zero).

If we then rescale the mean distance between eigenvalues to the unity, heuristically this is equivalence to *switch off* the external potential. Thus no matter the external potential, after the rescaling only the Coulomb potential affects the eigenvalues. Notice that the "Coulomb" interaction comes front he volume element and therefore, is intrinsic of the matrix ensemble. Hence, the universality. Recently, all this reasoning has been rigorously proven (see e.g. [10] and references therein). We will discuss a sketch of these ideas here.

Rescaled Integral Kernel and Universal Limit

Lets discuss in slightly more detail the meaning of universality for Hermitian ensembles. The quantity playing a major role in the analysis of random ensembles is the integral kernel

$$K_n(x, y) = e^{-\frac{n}{2}(V(x) + V(y))} \sum_{j=1}^{n} \phi_j(x) \phi_j(y) \tag{16}$$

where $V(x)$ is the potential [see Eq. (15)] and $\{\phi_j\}_{j=1}^{n}$ denotes the set of polynomials up to order $n-1$, orthogonal with respect to the weight $\exp\{-nV(x)\}$. The important statistical quantities associated with the matrix ensemble such as the eigenvalue density $\rho(\lambda_1) = \int P_n(\lambda_1, \cdots, \lambda_n) d\lambda_2 \cdots \lambda_n$ can be obtained by the K_n [10]. For large n the relation reads

$$K_n(\lambda, \lambda) = n\rho(\lambda)(1 + o(1)),$$

where $o(1)$ converges to 0 as $n \to \infty$.

To proceed the universality analysis, we must rescale the integral kernel. The scale is chosen such that the fraction of eigenvalues in an interval of length s close to a point λ equals s, in other words, the average spacing between eigenvalues is

unity. It can be shown that the correct scale is $K_n(\lambda, \lambda)$. Indeed, since the fraction of eigenvalues in the interval $A \subset \mathbb{R}$ is given by

$$f_n(A) = \frac{1}{n} \sum_{i=1}^{n} \chi_A(\lambda_i), \tag{17}$$

it is possible to show that [19] for large n we obtain

$$\left\langle n f_n \left(\lambda, \lambda + \frac{s}{K_n(\lambda, \lambda)} \right) \right\rangle = s[1 + o(1)].$$

Once we have the proper scale $K_n(\lambda, \lambda)$, we proceed the analysis by rescaling the integral kernel

$$\widetilde{K}_n(x, y) = \frac{1}{K_n(\lambda, \lambda)} K_n \left(\lambda + \frac{x}{K_n(\lambda, \lambda)}, \lambda + \frac{y}{K_n(\lambda, \lambda)} \right). \tag{18}$$

The astonishing result [19] is then

$$\lim_{n \to \infty} \widetilde{K}_n(x, y) = \frac{\sin \pi(x - y)}{\pi(x - y)} \tag{19}$$

exists pointwise for Freud–type or real analytic potentials V. Since it does not depend on V the Hermitian ensembles in this sense are universal. The natural questions is whether normal ensembles also display universal eigenvalue statistics.

Lecture 4: Non-Hermitian Random Matrices

The case of non-Hermitian ensembles are significantly less developed than the Hermitian. If the matrix is normal[2] some deeper understanding is possible. For example, Hermitian ensembles and the Wigner semi-circle law can be obtain as a limit of the normal case [20, 21]. Moreover, these ensemble have physical interpretations to quantum Hall effects and pattern formation [24–26]. Universality questions for normal ensemble have been addressed recently [6, 27, 28].

Lets is discuss a model introduced in Ref. [29]. Consider an ensemble of random matrices such that

$$\langle H_{ij} H_{ji} \rangle = \tau$$

where $\tau \in [-1, 1]$. This means that for $\tau \to -1$ this mimics a totally anti-symmetric ensemble, whereas for $\tau \to 1$, we would have a symmetric ensemble. Moreover, assume that the distribution of the independent elements is Guassian. Then, it is possible to show that

$$P(H)dH = \frac{1}{Z} \exp \left\{ -\frac{n}{2(1 - \tau^2)} \operatorname{Tr}(HH^T - \tau H^2) \right\} dH$$

---

[2]A matrix is normal if and only if it commutes with its adjoint

In this case, the eigenvalues are complex, but tt can be shown that the eigenvalues are uniformly distributed inside ellipsoids determined by τ. Indeed, density of eigenvalues reads as

$$\mu'(x + iy) = \begin{cases} \frac{1}{\pi ab}, & \text{if}(x/a)^2 + (y/b)^2 \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

where $a = 1 + \tau$ and $b = 1 - \tau$. In the case of totaly assymetric ensembles $\tau = 0$, we recover (almost surely) the Girko ensemble, with eigenvalues being uniformly distributed inside the circle.

References

[1] F.J. Dyson, J. Math. Phys. **3**, 140 (1962).

[2] E.P. Wigner, Conference on Neutron Physics, Tenessee, Lab, Rept. ORNL-2309 p.67 (1959). Reimpresso em [3].

[3] C.E. Porter, *Statistical Theories of Spectra: Fluctuations*, (Academic Press, New York, 1965).

[4] P. Deift, *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach*, (New York University 1999).

[5] C. Fernández-Ramírez and A. Relaño, Phys. Rev. Lett. **98**, 062001 (2007).

[6] H. Hedenmalm e N. Makarov, *Quantum Hele-Shaw flow*, arXiv:math.PR/0411437 (2004).

[7] Y. Lai, M. G. Frei, e I. Osorio, Phys. Rev. Lett. **98**, 108102 (2007).

[8] M. Timme, F. Wolf, and T. Geisel, Phys. Rev. Lett 92, 07410 (2004).

[9] T Pereira, D Eroglu, G Bagci, U Tirnakli, and HJ Jensen, Phys. Rev. Lett. 110, 234103 (2013).

[10] P. Deift. "Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach". American Mathematical Society. New York University, New York (2000).

[11] B. Bollobás, *Random Graphs*, Cambridge University Press; 2 edition (2001).

[12] F. R. K. Chung and L. Lu, *Complex Graphs and Networks*, American Mathematical Society (2006).

[13] L. M. Pecora and T. Carrol, Phys. Rev. Lett. **80**, 2109 (1998), H. Fujisika and T. Yamada, Progr. Theoret. Phys. **69**, 32 (1983).

[14] T. Pereira, J. Eldering, M. Rasmussen, A. Veneziani, Nonlinearity 27, 501 (2014).

[15] M.R Gardner, and W.R. Ashby, Nature 228, 784 (1970).

[16] R. May, Nature 238, 413 (1972).

[17] D. L. Mehta. "Random Matrices - Revised and Enlarged". 2nd Edition. Academic Press. 1991.

[18] A Venezianni, T Pereira, *A note on the volume form in normal matrix space*, (2013).

[19] P. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venekides, and X. Zhou, J. Comput. Appl. Math. **133**, 47-63 (2001).

[20] AM Veneziani, T Pereira, and DHU Marchetti, Journal of Physics A: Math. Theor. 44, 075202 (2011).

[21] P. Elbau e G. Felder, arXiv:math.QA/0406604v2 (2005).

[22] P.B. Wiegmann, A. Zabrodin, Comm. Math. Phys. **213**, 523 (2000).

[23] Ellison R, Gardner V, Lepak J, et ali., Inter. Jour. Bif. Chaos **15** (2005) 2283; W.J. Freeman and J.M. Barrie, Jour. Neurophys. **84** (2000) 1266.

[24] L.L. Chau and Y. Yu, Phys. Lett. A **167**, 452 (1992).

[25] A. Marshakov, P. Wiegmann, A. Zabrodin, Commun. Math. Phys. **227**, 131 (2002).

[26] A. Zabrodin, arxiv:0210331 (2002).

[27] AM Veneziani, T Pereira, and DHU Marchetti, Journal of Mathematical Physics 53, 023303 (2012).

[28] L.L. Chau and O. Zaboronsky, Comm. Math. Phys. 196, 203 (1998).

[29] H.J. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein, Phys. Rev. Letts. 60, 1895 (1998).