

Advanced Statistical Physics:

4. Random matrices

Leticia F. Cugliandolo

`leticia@lpthe.jussieu.fr`

Sorbonne Université

Laboratoire de Physique Théorique et Hautes Energies

Institut Universitaire de France

November 26, 2019

Contents

| | | |
|----------|---|----------|
| 4 | Random matrices | 3 |
| 4.1 | Introduction | 3 |
| 4.2 | Some emblematic applications | 6 |
| 4.2.1 | Vibrations of a membrane | 6 |
| 4.2.2 | Energy splitting in a quantum system | 7 |
| 4.2.3 | Quantum chaos | 9 |
| 4.2.4 | Riemann's zeta function | 10 |
| 4.2.5 | QCD | 11 |
| 4.3 | Distribution of level spacings | 12 |
| 4.3.1 | Independent levels: Poisson distribution | 12 |
| 4.3.2 | Correlated levels: Wigner's argument | 14 |
| 4.4 | The Gaussian orthogonal ensemble | 15 |
| 4.4.1 | Joint probability function of eigenvalues | 16 |
| 4.4.2 | The Coulomb gas | 17 |

4 Random matrices

4.1 Introduction

A *random matrix* is a matrix some or all of whose elements are random variables, drawn from a probability distribution. Random matrix theory is a branch of mathematics but it is also applied to describe numerous physical systems, some of which we will discuss here.

Random matrices were first used in the early 1900's in the study of the *statistics of population characteristics* by Wishart [1] (for example, the analysis of correlations between pairs of features of the population, say height and income..., were set in matricial form). More details on these applications can be found in [2]. (This paper was published in a journal called *Biometrika*, devoted to the statistical studies of the physical traits of living organisms. If you are interested in history, have a look at the role played by some of the articles published in this journal in what is called *scientific racism*. Still, on the correlations between height and income see [3].)

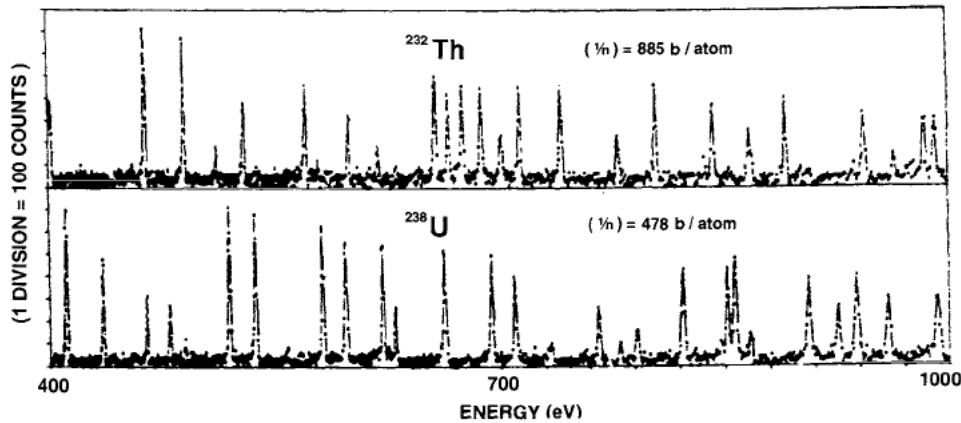


Figure 1.1. Slow neutron resonance cross-sections on thorium 232 and uranium 238 nuclei. Reprinted with permission from The American Physical Society, Rahn et al., Neutron resonance spectroscopy, X, *Phys. Rev. C* 6, 1854–1869 (1972).

Figure 4.1: This figure and caption were copied from Mehta's book [5].

The upsurge of physicists' interest in Random Matrix Theory came with its very successful application in the theory of *heavy nuclei*. Take Uranium. It is an atom with 92 protons, 92 electrons, and common isotopes have more than 140 neutrons. Its nucleus is

definitely very heavy, with over 200 protons and neutrons, each subject to and contributing to complex forces. If we completely understood the theory of the nucleus, we could predict all the energy levels of the spectrum; this, of course is not feasible.

Some insights into the nuclear structure of heavy nuclei were obtained with *scattering experiments* that consist in shooting neutrons (or protons) into the nucleus. The large numbers of energy levels of the nuclei appear in experimental data as peaks of the diffusion rate of the neutrons as a function of the energy, see Fig. 4.1. In the 50s data of this kind called the attention of theoreticians who tried to understand not really the low lying levels but the structure of the high energy spectra.

Since it appears to be impossible to derive the spectra of heavy nuclei one by one in full detail, the paradigm changed completely and the “revolutionary” idea put forward by Wigner and Dyson, mainly, was to construct a *statistical theory of energy levels*. In Dyson’s words: “such 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.” More still, the idea is

to develop a “new kind of statistical mechanics in which we renounce exact knowledge not of the state of a system but of the nature of the system itself. We picture a complex nucleus as a “black box” in which a large number of particles are interacting according to unknown laws. The problem is then to define in a mathematically precise way an ensemble of systems in which all possible laws of interaction are equally probable”.

This is the program initiated by Wigner [4] in the 50s. Since the overall energy scale is set by details of the nucleus at hand, the most generic or universal question to ask is what is the statistics of the separations between adjacent energy levels on a scale small compared to the energy. This is the *level spacing statistics*. The main and quite surprising content of this approach is that the statistical properties of the level statistics is related to the one between eigenvalues of conveniently chosen *random matrices*.

While independent random energy levels would yield a *Poisson distribution* of the distances between neighbouring energy levels, the experimental data were better described by what is called *Wigner’s surmise* [4]. The two probability densities are

$$p_{\text{Poisson}}(s) = e^{-s} \qquad p_{\text{Wigner}}(s) = c_{\beta} s^{\beta} e^{-a_{\beta} s^2} \qquad (4.1)$$

where $\beta = 1, 2, 4$, see Fig. 4.2, depending on the symmetry properties of the nucleus under time-reversal and spin rotation and the parameters a_{β} and c_{β} are fixed by $\int_0^{\infty} ds p(s) = \int_0^{\infty} ds s p(s) = 1$, the last identity ensuring that the normalization to unity of the average of s . (Note that, *a priori*, β has nothing to do with the inverse temperature in this

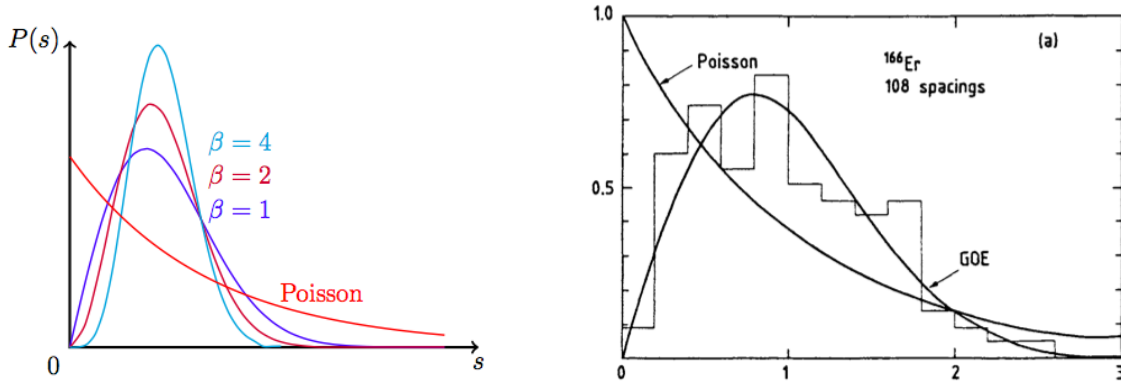


Figure 4.2: Left: the Poisson and Wigner distribution functions for the level spacings (the parameter β distinguishes different random matrix ensembles). Figure taken from [6]. Right: experimental data for the 108 level spacings in ^{166}Er (Erbium) favourably compared to the one of the Gaussian Orthogonal Ensemble and unfavourably compared to a Poisson distribution.

Chapter!) The main difference between the two forms is in the two extremes $s \simeq 0$ and $s \gg 1$, see Fig. 4.2. While Poisson’s pdf is finite for $s \rightarrow 0$, Wigner’s pdf has a dip close to $s = 0$, a feature that is associated with what is called *level repulsion*. In the large s limit, both distribution functions decay fast, with Wigner’s decaying faster than Poisson’s [5, 6, 7, 8, 9, 10]. We will discuss how one goes from the quantum mechanical problem to the diagonalisation of random matrices under certain conditions, and show a proof of these results below and in a particularly simple case in **TD 2**.

Random matrices appear in manifold fields of physics and mathematics and some of them, *vibrations of a membrane* [11], *quantum chaos* [12, 13], *Riemann’s zeta function* [14, 15, 16], *QCD* [17], will be discussed in the next section. Let us mention here that they also do in disordered condensed-matter systems, to describe *transport in disordered systems* [18]. In particular, random scattering matrices can provide a microscopic explanation for the mesoscopic conductance at low temperatures, a regime in which the quantum decoherence length becomes larger than the conductor’s size. Last but not least, random matrices also appear in *string theory*, *2d gravity* [19], *conformal field theory*, *integrable systems*, *RNA folding* [20], *glass theory* [21] and *stochastic processes* [22]. They are at the basis of many practical uses in *finance* [23] as well.

There are numerous excellent books and review articles that treat random matrix theory from different perspectives. Mehta’s book is the great classic in this field [5]. A more recent book is [8]. A concise presentation of the basics of random matrix theory appeared recently in [9]. Lectures by Eynard at IPhT [6] and Fyodorov [10] complement the physicists approach, I would say. More mathematical treatments can be found in the books [24, 25] and the lecture notes [26, 2]. Details on the numerical implementation of random matrix studies are in [27]. The connection with stochastic processes is discussed in [22] and a nice description of the historical development of the field is given in the first

Chapter written by Bohigas and Weidenmüller.

4.2 Some emblematic applications

We here briefly present several problems that lead to discrete energy spectra that can be analysed with random matrix theory. The examples belong to classical and quantum mechanics as well as mathematics.

4.2.1 Vibrations of a membrane

A membrane is a perfectly flexible and infinitely thin lamina of solid matter. Let us assume that it is made of uniform material and thickness, and that it is stretched in all directions by a tension so great as to remain unaltered during its vibrations. The principal subject in this field is the investigation of the transverse vibrations of membranes of different shapes, the boundaries of which are held fixed. One considers then a membrane with area Ω and boundary Γ . Imagine that the rest position of the membrane defines the zero displacement plane (x, y) and that, during its motion the time-dependent membrane's displacement relative to this plane is $\psi(x, y, t)$. This function obeys the *wave equation*

$$\frac{\partial^2 \psi(x, y, t)}{\partial t^2} = c^2 \nabla^2 \psi(x, y, t) \quad (4.2)$$

with c a velocity that depends on the physical properties of the membrane as well as on the tension imposed on it. The wave equation admits *harmonic solutions*

$$\psi(x, y, t) = \psi(x, y) e^{i\omega t} \quad (4.3)$$

or *normal modes*. The resulting equation for ψ (with an abuse of notation we keep the same symbol for the time-independent field) is an eigenvalue equation

$$\nabla^2 \psi(x, y) = -\frac{\omega^2}{c^2} \psi(x, y) = -k^2 \psi(x, y) = -E \psi(x, y) \quad (4.4)$$

with the *boundary condition* $\psi(x, y) = 0$ for $(x, y) \in \Gamma$. This equation possesses an infinite number of real positive eigenvalues with no accumulation point. One has

$$0 \leq E_1 \leq E_2 \leq E_3 \dots \quad \lim_{n \rightarrow \infty} E_n = \infty \quad (4.5)$$

The question is what are the properties of the energy spectrum. This problem can be generalised to higher dimensions and it can be made quantum with $\hbar k = \sqrt{2mE}$. It has attracted the attention of celebrated scientists such as Rayleigh, Sommerfeld, Lorentz, and Kac who asked the question ‘‘Can one hear the shape of a drum?’’ [11] More details on how to concretely analyse an energy spectrum are given with the help of the next example.

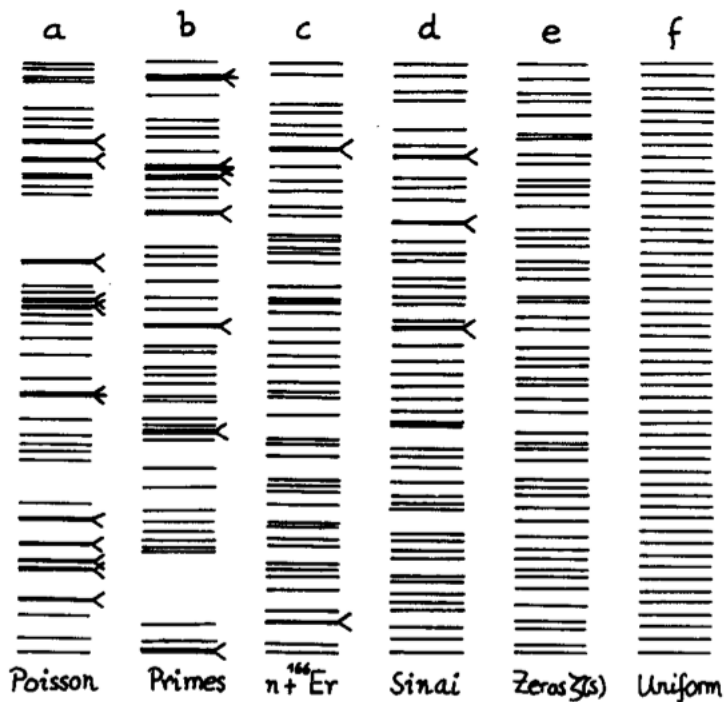


Figure 4.3: Some examples of level spectra taken from [13].

4.2.2 Energy splitting in a quantum system

Consider a quantum system with Hamiltonian \hat{H} . Its eigenstates are

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle \quad (4.6)$$

with the (joint set of) quantum number(s) n labelling the levels, that we consider to be discrete. In the matrix representation of quantum mechanics¹ this equation corresponds to the eigenvalue problem

$$H\psi_n = E_n\psi_n. \quad (4.7)$$

In general, the Hamiltonian is then represented by a *complex Hermitian matrix*. When there are some exact quantum numbers corresponding to exact integrals of motion, like angular momentum and parity, if the basis states are labelled by these exact quantum numbers, the Hamiltonian matrix splits into blocks, and the matrix elements connecting different blocks vanish. Once such a basis is chosen one can restrict attention to one of the

¹Let \hat{H} be the Hamiltonian operator, and $|\psi\rangle$ a generic state on which it will act. Calling $\{|u_i\rangle\}$ a complete orthonormal set of states, then $\hat{H}|\psi\rangle$ is represented as $H_{ij}\psi_j$ with $\psi_j = \langle u_j|\psi\rangle$ and the summation convention over the repeated index j adopted.

diagonal blocks, an $N \times N$ Hermitian matrix in which N is a large integer, for a system with many levels. The theoretical results are often derived in the limit $N \gg 1$.

The focus is then set on the energy spectrum. As the excitation energy increases, the energy levels occur on the average at smaller and smaller intervals. In other words, level density increases with the excitation energy. More precisely, let us focus on an interval of energy ΔE centred at E . This interval is much smaller compared to E whereas it is large enough to contain many levels since it is taken to be much larger than the mean distance between neighbouring levels D :

$$D \ll \Delta E \ll E. \quad (4.8)$$

In practice these scales can be very different from one nucleus to another. It then makes sense to normalise the level distances by their average (in the interval ΔE of interest). After normalisation, the average distance between the neighbouring levels is unity in all cases.

In the case of the nucleus we do not know the Hamiltonian and, on top, even if we did, the eigenvalue problem would be too difficult to solve. The random matrix approach relies on the assumption that one can take H to be a random matrix with elements restricted only by the symmetry properties of the problem. Imagine that the relevant kind of matrices to use were *real symmetric*. The first question one can ask is, which is the joint probability distribution, $\rho(H_{11}, H_{12}, \dots, H_{NN})$ of the elements of such a matrix? Due to the symmetry one needs to distinguish the diagonal elements, H_{ii} , from the off-diagonal ones on the which the symmetry has to be imposed. If all original H_{ij} are drawn from a Gaussian pdf, and the off-diagonal ones are constructed as $(H_{ij} + H_{ji})/2$, then

$$\rho(\{H_{ij}\}) = \prod_i \frac{1}{\sqrt{2\pi\sigma^2}} e^{-H_{ii}^2/(2\sigma^2)} \prod_{i \neq j} \frac{1}{\sqrt{\pi\sigma^2}} e^{-H_{ij}^2/\sigma^2}. \quad (4.9)$$

Exercise 4.1 Prove this result.

In order to make contact with the physical problem, one needs to study the eigenvalues of such a matrix, more precisely, the pdf of the eigenvalues of matrices thus constructed, and in a second stage, the pdf of the spacings between these eigenvalues.

A conjecture allows one to believe the results obtained in this way will be general:

Let H be an $N \times N$ real symmetric matrix, with off-diagonal elements H_{ij} with $i < j$ being independent identically distributed (i.i.d.) random variables with mean zero and variance σ^2 , i.e. $\langle H_{ij} \rangle = 0$ and $\langle H_{ij}^2 \rangle = \sigma^2$ with $0 < \sigma^2 < \infty$ (and similarly for the diagonal ones). In the limit of large N the statistical properties of the n eigenvalues of H become independent of the probability density of the H_{ij} i.e. when $N \rightarrow \infty$ the joint probability density of arbitrarily chosen n eigenvalues of H tends, for every finite n , with probability one, to the point correlation function of the *Gaussian orthogonal ensemble*.

This statement means that the local statistical properties of a few eigenvalues of a large random matrix are independent of the distribution of the individual elements, although it does depend on the symmetry properties of the matrix. (It has a flavour of the central limit theorem.) The same applies to other ensembles of random matrices.

The underlying space-time symmetries obeyed by the system put important restrictions on the admissible matrix ensembles. If the Hamiltonian is time-reversal invariant and invariant under rotations the Hamiltonian matrices can be chosen *real symmetric*. If the Hamiltonian is not time-reversal invariant then irrespective of its behaviour under rotations, the Hamiltonian matrices are *complex Hermitian*. Finally, if the system is time-reversal invariant but not invariant under rotations, and if it has half-odd-integer total angular momentum the matrices are *quaternion real*.

In these lectures we will focus on the case of real symmetric matrices. The real symmetry property is preserved under orthogonal transformations. Moreover, since we have already argued that for elements with finite variance the actual pdf used to draw them is immaterial one typically chooses to work with Gaussian probabilities. For these two reasons, this set is usually called the Gaussian orthogonal ensemble (GOE). The parameter that is commonly used to distinguish the three classes is called β and it is equal to $\beta = 1$ in the GOE class, see Fig. 4.2.

4.2.3 Quantum chaos

The great success of random matrix theory in the description of level spacings of heavy nuclei suggested that the same approach could be applied to other problems with complicated energy spectra. The quantum mechanics of a particle in a *2d chaotic billiard* is one such case. A classical particle performs free motion and it is simply reflected at the boundaries, it performs chaotic motion for many choices of the boundaries (like for Sinai's or the stadium shaped spaces), and the problem is therefore ergodic, see Fig. 4.4. A remarkable conjecture, that uses random matrices ideas to describe the quantum behaviour of these problems, is due to Bohigas, Giannoni and Schmit (BGS) [12]:

- if a billiard is classically integrable (for example a rectangle or an ellipse), the spacings between eigenvalues of the corresponding quantum Hamiltonian follow a Poisson law, and there are many level crossings.
- If it is chaotic, then the eigenvalues have the same statistical properties as those of random matrices, and they do not cross, repelling each other instead.

The numerical and experimental evidence for the validity of the BGS conjecture is overwhelming. However, the connection between *quantum chaos* and random matrix theory

is not very well understood yet, and there is no satisfying analytic proof of the BGS conjecture.

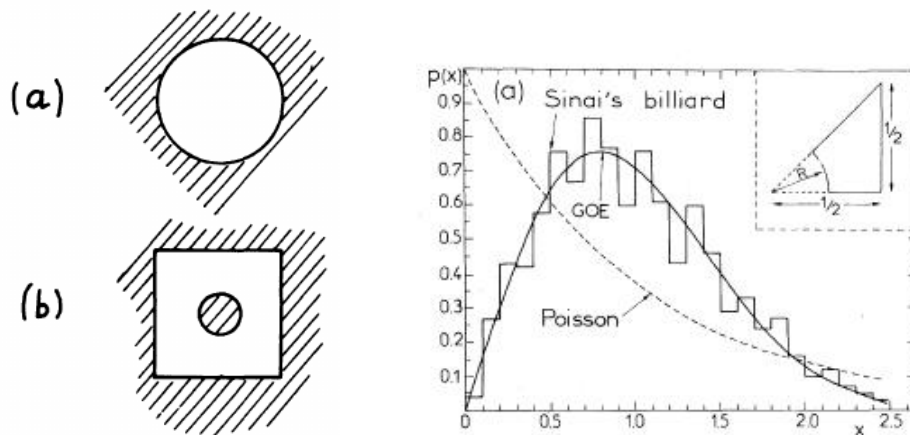


Figure 4.4: Left: an integral (a) and a chaotic (b) billiard. Right: numerical check of the BGS hypothesis in Sinai's billiard. The two figures are taken from [12].

4.2.4 Riemann's zeta function

In the early 1970's a remarkable connection was unexpectedly discovered between two very different fields, nuclear physics and *number theory*, when it was noticed that random matrix theory accurately modelled many problems in the latter as well [16]. The theory of numbers deals with the arithmetic properties of integers $1, 2, 3, \dots$, in particular of prime numbers. The spacing between prime numbers can be related to the one of the zeroes of a mathematical object, Riemann's zeta function, and the statistical properties of these can be set in contact with the ones of the eigenvalues of a certain class of random matrices.

The Riemann zeta function is defined as

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ primes}} \left(1 - \frac{1}{p^s}\right)^{-1} \quad (4.10)$$

for complex s with real part greater than 1.

The *prime number theorem* states that the number of primes smaller than x is $Li(x) + o(Li(x))$, where $Li(x) = \int_2^x dt/\ln t$ and for x large, $Li(x) \simeq x/\ln x$. Riemann observed that the error term in the prime number theorem [14] is expressed in terms of a sum over the zeroes of the Riemann zeta function.

The zeta function defined in eq. (4.10) has so-called trivial zeros at all negative even integers $s = -2, -4, -6, \dots$, and “nontrivial zeros” at complex values $s = \sigma + it$ for s in the *critical strip* $0 < \sigma < 1$. The *Riemann hypothesis* dates from 1859 and asserts that the nontrivial zeros of $\zeta(s)$ all have real part $\sigma = R[s] = 1/2$, a line called the *critical line*.

Riemann stated “...it is very probable that all roots are real. Of course one would wish for a rigorous proof here; I have for the time being, after some fleeting vain attempts, provisionally put aside the search for this, as it appears dispensable for the immediate objective of my investigation.” (He was discussing a version of the zeta function, modified so that its roots (zeros) are real rather than on the critical line.) The hypothesis still remains unproven and is one of the millennium prize problems stated by the Clay Mathematics Institute, see <http://www.claymath.org/millennium-problems>.

Assuming the validity of Riemann’s hypothesis, the non-trivial zeroes lie on a straight line along the imaginary direction in the complex plane. It therefore makes sense to study the distribution of distances between adjacent zeros. It was noticed by Montgomery and Dyson [15] that the agreement between the pair correlation of the zeroes of the ζ function and the eigenvalues of complex Hermitian random matrices is amazingly good. Since then the numerical checks of this hypothesis have been numerous and the accuracy of the agreement between the two kinds of correlation functions became just amazing.

4.2.5 QCD

Quantum chromodynamics (QCD) is a gauge theory with a gauge field that describes gluons. This gauge field lives in the Lie algebra of the group $G = U(N_c)$ with N_c the number of colors. Therefore, A is an anti-Hermitian matrix of size N_c .

The effective coupling constant in quantum chromodynamics becomes large at large distances where the perturbation theory is not applicable and there is no small parameter in this theory.

Although $N_c = 3$ in QCD, ’t Hooft proposed to study a generalization of QCD in the limit $N_c \rightarrow \infty$ [17] to allow for the control of the Feynman diagram perturbative expansion, and then recover $N_c = 3$ by perturbation theory in an $1/N_c$ expansion. The motivation was an expansion in the inverse number of field components N in statistical mechanics that allows to solve the $O(N)$ model (recall the Phase Transition Lectures) in the large N limit exactly. Indeed, the expansion of QCD in the inverse number of colours $1/N_c$ (known as the $1/N_c$ expansion) is such that only planar diagrams of the type in Fig. 4.5 survive in the perturbative expansion when the large- N_c limit is taken. Moreover, the diagrams of the perturbation theory grouped by the power of $1/N_c$ are also grouped according to their topology (e.g. genus number). As in statistical physics models, there is a large $N_c \rightarrow \infty$ limit with a proper choice of the scaling of coupling constants, in which the remaining theory is non-trivial. Following the line of reasoning of Wigner, the $N_c \times N_c$ matrix is random and the $N_c \rightarrow \infty$ generalization of QCD becomes a theory of large random matrices. For the simplest case of the Hermitian one-matrix model, which is

related to the problem of enumeration of graphs, an explicit solution at large N_c was first obtained by Brézin, Itzykson, Parisi & Zuber [28]. The problem consists in computing

$$Z = \int dM e^{-N_c \text{Tr} V(M)} \quad (4.11)$$

with M an $N_c \times N_c$ Hermitian matrix and a proper choice of the potential $V(M)$. The problem to solve is similar to the computation of Wigner's semicircle law for the GOE although perturbed by the interactions in $V(M)$.

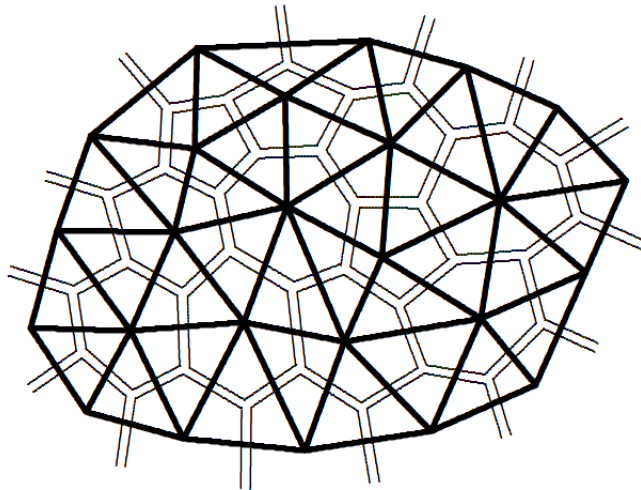


Figure 4.5: A planar graph in large N_c matrix model theory with cubic interactions. The double lines correspond to the two indices in the square matrix M . Its dual graph (depicted by bold lines) is constructed from equilateral triangles.

4.3 Distribution of level spacings

In this Section we explain the origin of the Poisson and Wigner distribution functions for the energy spacings. The problem is set as follows. Consider an energy interval ΔE , measure (or compute it if possible!) the energy levels that fall in this interval $E_1 \leq E_2 \leq \dots$ and let S_n be the consecutive splittings $S_n = E_{n+1} - E_n$. The average value $D = \langle S_n \rangle$ is the mean splitting over the energy interval ΔE and one considers the relative quantity $s_n = S_n / \langle S_n \rangle = S_n / D$. The probability density of s is called $p(s)$.

4.3.1 Independent levels: Poisson distribution

One can imagine the energies as being points placed on a real axis. If these positions were independent and not correlated, the probability that any E_n will fall between E and $E + dE$ is independent of E and is simply ρdE , where $\rho = D^{-1}$ is the average number of levels in a unit interval of energy.

Let us determine the probability of a spacing S ; that is, given a level at E , what is the probability of having no level in the interval $I = (E, E + S)$ and one level in the interval $dI = (E + S, E + S + dS)$. A sketch of this interval is shown in Fig. 4.6 (where $x_0 = E$, $x = E + S$ and $x + dx = E + S + dS$).

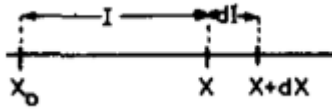


Figure 4.6: The interval partition, from [13].

We first look at the first interval $I = (E, E + S)$ and we investigate the condition of having no level in it. We divide the interval I into m equal parts each of length $dE = S/m$:

$$I = (E, E + S) = (E, E + S/m] \cup (E + S/m, E + 2S/m] \cup \dots \cup (E + (m - 2)S/m, E + (m - 1)S/m] \cup (E + (m - 1)S/m, E + S) \quad (4.12)$$

Since the levels are *independent*, the probability of having no level in $(E, E + S)$ is the product of the probabilities of having no level, $1 - \rho dE = 1 - \rho S/m$, in any of these m parts

$$P(\text{no level in } I = (E, E + S)) = \left(1 - \frac{\rho S}{m}\right)^m \quad (4.13)$$

and in the large m limit

$$\lim_{m \rightarrow \infty} P(\text{no level in } I = (E, E + S)) = e^{-\rho S} . \quad (4.14)$$

Instead, the probability of having a level in the remaining infinitesimal interval going from $E + S$ to $E + S + dS$ is ρdS . Therefore, the searched probability is

$$p(S)dS = \rho e^{-\rho S} dS \quad (4.15)$$

that, in terms of the normalised variable $s = S/D = \rho S$, becomes

$$p(s)ds = e^{-s} ds . \quad (4.16)$$

This is known as the Poisson distribution or the spacing rule for random independent levels. One can easily check $\int_0^\infty ds \rho(s) = \int_0^\infty ds s \rho(s) = 1$.

4.3.2 Correlated levels: Wigner's argument

We now revisit the same question. The probability we are searching is the one of occurrence of the two events, $B = \text{no-level in } (E, E + S) \cup A = \text{one level in } (E + S, E + S + dS)$. Therefore, we need to find this joint probability that, for brevity, we call $A \cap B$, and $P(A \cap B) = P(A/B)P(B)$, where $P(A/B)$ is the *conditional probability* of A given B . Going back to the notation of the problem at hand, the level spacing probability density $p(S)$ is such that

$$p(S)dS = \text{Prob}(\text{one level in } dI/\text{no level in } I) \text{Prob}(\text{no level in } I) \quad (4.17)$$

Now,

$$\text{Prob}(\text{no level in } I) = \int_S^{S_{\max}} dS' p(S'), \quad (4.18)$$

meaning that we need to have all level spacing larger than S not to have a level in the interval $(E, E + S)$. We called S_{\max} the maximal possible value of the level spacing. We call

$$\mu(S)dS \equiv \text{Prob}(\text{one level in } dI/\text{no level in } I). \quad (4.19)$$

Therefore, Eq. (4.17) reads

$$p(S) = \mu(S) \int_S^{S_{\max}} dS' p(S') \quad (4.20)$$

and can be solved² for $p(S)$ to yield

$$p(S) = C' \mu(S) e^{-\int_S^{S_{\max}} dS' \mu(S')}. \quad (4.21)$$

From here one can obtain the two limiting results, Poisson and Wigner.

For a Poisson case, the two events A and B are independent, $\mu(S) = \mu$ and

$$p_{\text{Poisson}}(S) = C \mu e^{-\mu S}. \quad (4.22)$$

The normalisation of the probability density $\int_0^{S_{\max}} dS p_{\text{Poisson}}(S) = 1$ and the fact that we want to deal with the splittings measured with respect to their mean, $\int_0^{S_{\max}} dS S p_{\text{Poisson}}(S) = \langle S \rangle = D$, fix $\mu = 1$ and $C = 1/\mu$. Finally,

$$p_{\text{Poisson}}(s) = e^{-s}. \quad (4.23)$$

If, instead, one assumes $\mu(S) = a_\beta S$, mimicking level repulsion, $p(S)$ takes the Wigner form

$$p_{\text{Wigner}}(S) = C a_\beta S e^{-\int_S^{S_{\max}} dS' a_\beta S'} = C' a_\beta S e^{\frac{a_\beta}{2} S^2}. \quad (4.24)$$

²First, write the equation as $y(S) \equiv p(S)/\mu(S) = \int_S^{S_{\max}} dS' y(S')\mu(S')$, next as $\dot{y}(S) = -y(S)\mu(S)$, and from here $\ln y(S) = -\int^S dS' \mu(S') + \text{cst.}$

If we now use the normalisation condition and we require that the average of s be one to work with the normalised splitting, we fix $a_\beta = -\pi/2$ and $C'a_\beta = \pi/2$ and

$$p_{\text{Wigner}}(s) = \frac{\pi}{2} s e^{-\frac{\pi}{4}s^2}. \quad (4.25)$$

In **TD2** we will study the pdf of the distances between eigenvalues of a 2×2 real symmetric matrix with elements drawn from a Gaussian probability function and the symmetry imposed as in Eq. (4.9). We will find that $p(s)$ exactly takes the form (4.25).

The take-home message from these calculations is that the eigenvalues of a random matrix are not independent.

In general, the joint probability distribution of the eigenvalues of a random matrix does not factorise: $\rho(\lambda_1, \dots, \lambda_N) \neq \rho(\lambda_1) \dots \rho(\lambda_N)$.

4.4 The Gaussian orthogonal ensemble

For a number of nuclear physics (and other applications) the relevant random matrix ensemble to use is the Gaussian orthogonal one in which the matrices are real and symmetric. The matrices have then $N(N+1)/2$ independent elements, M_{ij} with $i \leq j$, and the probability distribution is

$$p(M) = \prod_{i \leq j} p_{ij}(M_{ij}) \quad (4.26)$$

Since we argued that the choice of the p_{ij} is not important, as long as the variance is finite, one uses the same Gaussian pdf to draw each element in the (half of the) matrix. (In the nuclear physics context, the symmetry is a consequence of the invariance under space rotations, see the proof in [5] for example.)

The ensemble is invariant under orthogonal transformations

$$M' = O^T M O \quad (4.27)$$

with O any real orthogonal matrix. Precisely, this means that the probability $p(M)dM$ that a matrix M will be in the volume element dM is invariant under orthogonal transformations:

$$p(M)dM = p(M')dM'. \quad (4.28)$$

Exercise 4.2 A set of hints of how this result can be proven are given in **TD 2**.

The first numerical investigation of the distribution of successive eigenvalues associated with random matrices was carried out by Porter and Rozenzweig in the late 1950's [63]. They diagonalised a large number of matrices where the elements are generated randomly but constrained by a probability distribution. The analytical theory developed in parallel

with their work. At the time it was clear that the spacing distribution was not influenced significantly by the chosen form of the probability distribution of the individual elements. Remarkably, the distribution of the eigenvalue spacings of the $N \times N$ matrices had forms given almost exactly by the original Wigner distribution of the 2×2 matrices (see **TD2**).

The invariance of the measure with respect to the rotations imposes that it can only depend on N traces of the powers of M ,

$$\mathrm{Tr} M^j = \sum_{n=1}^N \lambda_n^j \quad \text{with} \quad j = 1, \dots, N \quad (4.29)$$

where $(\lambda_1, \dots, \lambda_N)$ are the N eigenvalues of M . Moreover, one can prove that for matrices constructed with statistical independent elements, the traces of the first two powers are enough, $j = 1, 2$, and these can only occur in the exponential; a more compact expression is

$$p(M) \propto e^{-a \mathrm{Tr} M^T M}, \quad (4.30)$$

see also **TD 2** for a proof of this statement, with the proportionality constant linked to the need to normalise the pdf.

4.4.1 Joint probability function of eigenvalues

The issue is to deduce the joint probability distribution function of the eigenvalues $p(\lambda_1, \dots, \lambda_N)$ implied by the Gaussian probability of the matrix elements (further constrained to build a symmetric matrix).

The real symmetric matrix has $N \times (N + 1)/2$ independent elements. Therefore, such a matrix can be parametrised with its N eigenvalues plus $N(N + 1)/2 - N = N(N - 1)/2$ other parameters that we call p_ℓ with $\ell = 1, \dots, N(N - 1)/2$. From the change of variables form M_{ij} to λ_n, p_ℓ

$$\prod_{i \leq j} p(\{M_{ij}\}) dM_{ij} = \prod_n \prod_\ell p(\{\lambda_n, p_\ell\}) d\lambda_n dp_\ell \quad (4.31)$$

that implies that the joint probability distribution of the eigenvalues is the integral over all parameters p_ℓ of the right-hand-side:

$$p(\{\lambda_n\}) = \int \prod_\ell dp_\ell p(\{\lambda_n, p_\ell\}) = \int \prod_\ell dp_\ell p(\{M_{ij}\}) \left| \frac{\partial(M_{11}, M_{12}, \dots, M_{NN})}{\partial(\lambda_1, \dots, \lambda_N, p_1, \dots, p_{N(N-1)/2})} \right|$$

The last factor is the Jacobian of the transformation from the M_{ij} to the λ_n and p_ℓ .

We now need to find the parameters p_ℓ . The matrix M can be diagonalised with an orthogonal transformation

$$M = O^T D O \quad \text{with} \quad O^T O = 1 \quad (4.32)$$

and D the diagonal matrix with elements equal to the eigenvalues of M (we will assume that the eigenvalues are non-degenerate, see [5] for the treatment of more general cases). O is a matrix with real elements whose columns are the normalised eigenvectors of M , that can be chosen to be orthogonal. By taking variations of the two sides of eq. (4.32) with respect to p_ℓ and manipulating in various ways the Jacobian (where the derivatives of M_{ij} with respect to p_ℓ appear) one ends with

$$p(\{\lambda_n\}) = \prod_{1 < j < k < N} |\lambda_j - \lambda_k| e^{a+b \sum_n \lambda_n + c \sum_n \lambda_n^2} \quad (4.33)$$

(a detailed derivation with this method can be found in Mehta's book [5]). With a simple linear change of variables from λ_n to x_n the linear term is eliminated and

$$p(\{x_n\}) = C \prod_{1 < j < k < N} |x_j - x_k| e^{-\frac{1}{2} \sum_n x_n^2}. \quad (4.34)$$

One reckons that the two factors in this form have different effects:

- the Gaussian one kills any configuration of eigenvalues that are far from zero;
- the numerator kills the configurations in which two eigenvalues get too close to each other, makes the eigenvalue ‘interact’ and it is at the heart of the level repulsion.

Exercise 4.3 Use the joint pdf (4.34) to compute the level spacing distribution in the case of a 2×2 random symmetric matrix (the example in **TD2**) and verify that the same result is found.

For the two other random matrix ensembles the product is replaced by the one of $|x_i - x_j|^\beta$ with $\beta = 1, 2, 4$.

4.4.2 The Coulomb gas

The joint probability density of the eigenvalues can be written as the Boltzmann factor for a gas of charged particles interacting via a two dimensional Coulomb force. The equilibrium density of this *Coulomb gas* is such as to make the potential energy a minimum, and this density is identified with the level density of the corresponding Gaussian ensembles. Consider a gas of N point charges with positions x_i free to move on the infinite straight line $-\infty < x_i < \infty$. Suppose that the potential energy of the gas is given by

$$V(\{x_i\}) = \frac{1}{2} \sum_{i=1}^N x_i^2 - \sum_{i < j} \ln |x_i - x_j| \quad (4.35)$$

The first term in V represents a harmonic potential which attracts each charge independently towards the origin $x = 0$; the second term represents an electrostatic repulsion

between each pair of charges. (The logarithmic function comes in if we assume the universe to be two-dimensional and therefore we have a *2d gas of charged particles confined to move on a line.*)

The potential energy is bounded from below

$$V \geq V_0 = -N(N-1)(1 + \ln 2) + \sum_{j=1}^N j \ln j \quad (4.36)$$

and this minimum is attained when the positions of the charges coincide with the zeros of the Hermite polynomial $H_N(x)$.

Let this charged gas be in thermodynamical equilibrium at a temperature T , so that the probability density of the positions of the N charges is given by the Boltzmann weight

$$P(\{x_i\}) = Z^{-1}(\beta) e^{-\beta V(\{x_i\})}. \quad (4.37)$$

(Thinking in terms of moving point charges one could add to the potential energy the kinetic energy. This, in any case, gives a trivial factor that depends on β and the mass of the particles. We can simply ignore it.) We note that this expression is the probability distribution of the eigenvalues provided that $\beta = 1$ in the GOE is interpreted as the inverse temperature here! The miracle now is that the integral over the positions of the point-like particles can be computed exactly for any N

$$Z(\beta) = \frac{(2\pi)^{N/2}}{\beta^{N/2} + \beta N(N-1)/4} \frac{1}{(\Gamma(1 + \beta/2))^N} \prod_{j=1}^N \Gamma(1 + \beta j/2). \quad (4.38)$$

The density of eigenvalues can then be simply computed by performing the integral over $N-1$ eigenvalues

$$\rho(x_N) = \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_{N-2} \int_{-\infty}^{\infty} dx_{N-1} P(x_1, \dots, x_N). \quad (4.39)$$

In this way, the eigenvalue density of the random matrices in the GOE can be derived for any finite N . In the limit $N \rightarrow \infty$ this distribution approaches the semi-circle law depicted in Fig. 4.7.

Saddle-point evaluation

There is another, simpler, way to compute the partition function in the large N limit. The idea is to transform the partition sum into a functional integral that will be evaluated by the saddle-point method. First of all, one rescales the eigenvalues

$$x_i \mapsto x_i \sqrt{\beta N} \quad (4.40)$$

so that the new x_i in the right-hand-side are now $\mathcal{O}(1)$. The potential energy now reads

$$V \mapsto V = \frac{\beta N}{2} \sum_{i=1}^N x_i^2 - \sum_{i < j} \ln |x_i - x_j|^\beta + \text{ct} \quad (4.41)$$

(where we recovered the power β in the logarithm, that would have been present for other random matrix ensembles, for convenience) and the partition sum

$$Z(\beta) = \prod_{n=1}^N \int dx_n e^{-\beta N^2 v(\{x_i\})} \quad (4.42)$$

$$v(\{x_i\}) \equiv \frac{1}{2N} \sum_{i=1}^N x_i^2 - \frac{1}{2N^2} \sum_{i \neq j} \ln |x_i - x_j|, \quad (4.43)$$

Next, one introduces a (coarse-grained) one-particle *counting function*, $\rho(x)$, through

$$\rho(x) = \frac{1}{N} \sum_{n=1}^N \delta(x - x_n) \quad (4.44)$$

that can be enforced into the discrete sum *via* the insertion of the identity

$$1 = \int \mathcal{D}\rho[x] \delta\left(\rho(x) - \frac{1}{N} \sum_{n=1}^N \delta(x - x_n)\right). \quad (4.45)$$

The partition function then reads

$$Z(\beta) \propto \int \mathcal{D}\rho[x] \prod_i \int dx_i \delta\left(\rho(x) - \frac{1}{N} \sum_{i=1}^N \int dx_i \delta(x - x_i)\right) e^{-\beta N^2 v(\{x_i\})}. \quad (4.46)$$

We now use the identities

$$\frac{1}{N} \sum_{i=1}^N f(x_i) = \int dx \rho(x) f(x), \quad (4.47)$$

$$\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N g(x_i, x_j) = \int dx \rho(x) \rho(y) g(x, y), \quad (4.48)$$

to write the potential as

$$v(\rho(x)) = \frac{1}{2} \int_{-\infty}^{\infty} dx x^2 \rho(x) - \frac{1}{2} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \ln |x - y| \rho(x) \rho(y) \quad (4.49)$$

where we have traded the $\{x_i\}$ dependence into a $\rho(x)$ one (a short distance cut-off is needed to cure the diverge at $x \rightarrow y$, we ignore it here).

Since there is no more $\{x_i\}$ dependence in v the mutiple-integral over the eigenvalues can now be computed. It yields an *entropy* (recall the calculations on the fully-connected Ising model) since it basically counts how many *micro-state configurations* yield the same

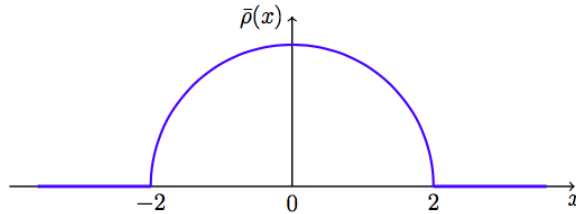


Figure 4.7: The semi-circle law eigenvalue distribution of matrices in the Gaussian orthogonal ensemble, with the horizontal and vertical axes normalised by \sqrt{N} to make the non-trivial interval be of finite extent. The limit $N \rightarrow \infty$ has been taken in this figure in such a way that the edges are sharp and ρ is identical to zero beyond -2 and 2 .

$\rho(x)$. Without entering into the details of this computation that can be found in [9], one obtains

$$I_N(\rho(x)) = e^{-N \int dx \rho(x) \ln \rho(x)} \quad (4.50)$$

the *Shannon entropy*, as one could have expected.

We note now that the potential energy contribution to the action is order N^2 while the entropic one is only order N . It is therefore negligible in this case.

The level density is then obtained from the minimisation of just the potential energy expression under the constraint

$$\int_{-\infty}^{\infty} dx \rho(x) = 1 \quad (4.51)$$

that can be imposed with a Lagrange multiplier that we call κ .

The functional saddle point leads to

$$0 = \frac{x^2}{2} - \int dx' \rho_{\text{sp}}(x') \ln |x - x'| - \kappa_{\text{sp}}, \quad (4.52)$$

$$1 = \int_{-\infty}^{\infty} dx' \rho_{\text{sp}}(x'), \quad (4.53)$$

and the task is now to solve the first of these two equations constrained by the second. After a number of steps detailed in [5, 9] one finds the semi-circle law

$$\rho(x) = \begin{cases} 1/\pi \sqrt{2 - x^2} & |x| < 2^{-1/2} \\ 0 & |x| \geq 2^{-1/2} \end{cases} \quad (4.54)$$

The edges of the semicircle are called soft: for large but finite N , there is always a nonzero probability of sampling eigenvalues exceeding the edge points. (Other ensembles have spectral densities with hard edges which the eigenvalues can never cross.)

There are many independent derivations of the semi-circle law for the eigenvalue density. In **TD 2** we will see a combinatorial one. Interestingly enough, the semi-circle law can also be derived with the *replica trick* used to study disordered systems [29].

The full calculation of the level spacing distribution is long and a bit tedious. It can be found in [5] for this and the other ensembles.

Dyson proposed to define various expressions that relate to the energy-level series in complete analogy with the classical notions of entropy, specific heat, and the like. One can measure these observables from the experimental data and check in this way the theory.

References

- [1] J. Wishart, *The generalized product moment distribution in samples from a normal multivariate population*, *Biometrika* **20A**, 32 (1928).
- [2] G. Borot, *An introduction to random matrix theory*, Lectures at National University of Mongolia, Ulaanbaatar (2015).
- [3] <https://www.theguardian.com/science/2016/mar/08/genetic-study-shows-mens-height-and-womens-weight-drive-earning-power>
- [4] E. P. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, *Ann. Math.* **62**, 548 (1955).
- [5] M. L. Mehta, *Random matrices*, 3rd edition (Elsevier, 2004).
- [6] B. Eynard, T. Kimura and S. Ribault, *Random matrices*, Lectures given at IPHT Saclay.
- [7] F. W. K. Firk and S. J. Miller, *Nuclei, Primes and the Random Matrix Connection*, *Symmetry* **1**, 64 (2005).
- [8] *The Oxford Handbook of Random Matrix Theory*, G. Akemann, J. Baik, and P. Di Francesco, eds. Oxford Handbooks in Mathematics (Oxford University Press, 2011).
- [9] G. Livan, M. Novaes and P. P. Vivo, *Introduction to Random Matrices Theory and Practice* (Springer-Verlag, 2017), arXiv:1712.07903.
- [10] Y. V. Fyodorov *Introduction to the Random Matrix Theory: Gaussian Unitary Ensemble and Beyond*, arXiv:math-ph/0412017.
- [11] M. Kac, *Can One Hear the Shape of a Drum?* *American Mathematical Monthly* **73**, 1 (1966).

-
- [12] O. Bohigas, M. J. Giannoni, and C. Schmit, *Characterization of Chaotic Quantum Spectra and Universality of Level Fluctuation Laws*, Phys. Rev. Lett. **52**, 1 (1984).
- [13] O. Bohigas and M-J Giannoni, *Chaotic motion and random matrix theories*, (Springer-Verlag)
- [14] B. Riemann, *Ueber die Anzahl der Primzahlen unter einer gegebenen GröÙsse* (*On the Number of Primes Less Than a Given Magnitude*) Monatsberichte der Berliner Akademie. In *Gesammelte Werke*, Teubner, Leipzig (1892), reprinted by Dover, New York (1953).
- [15] H. J. Montgomery, *The pair correlation of zeros of the zeta function*, Analytic number theory, Proc. Sympos. Pure Math., XXIV, Providence, R.I.: American Mathematical Society, pp. 181-193 (1973).
- [16] J. P. Keating and N. C. Snaith, *Random matrix theory and number theory*.
- [17] G. 't Hooft, *A Planar Diagram Theory for Strong Interactions*, Nucl. Phys. B **72**, 461 (1974).
- [18] C. W. J. Beenakker, *Random-matrix theory of quantum transport*, Rev. Mod. Phys. **69**, 731 (1997).
- [19] P. Di Francesco, P. H. Ginsparg, J. Zinn-Justin, *2-D Gravity and random matrices*, Phys. Rep. **254**, 1 (1995).
- [20] H. Orland and A. Zee, *RNA Folding and Large N Matrix Theory*, Nucl. Phys. B **620**, (2002)
- [21] L. F. Cugliandolo, J. Kurchan, G. Parisi and F. Ritort, *Matrix models as solvable glass models*, Phys. Rev. Lett. **74**, 1012 (1995).
- [22] *Stochastic Processes and Random Matrices*, Lecture Notes of the Les Houches Summer School, Volume 104, July 2015, ed. by G. Schehr, A. Altland, Y. V. Fyodorov, N. O'Connell, and L. F. Cugliandolo (Oxford University Press, 2017).
- [23] J.P. Bouchaud and M. Potters, *Theory of finance risks: from statistical physics to risk management* (Cambridge University Press, 2000).
- [24] T. Tao, *Topics in random matrix theory*, Lectures at UCLA.
- [25] G. W. Anderson, A. Guionnet and O. Zeitouni, *An Introduction to Random Matrices*, (Cambridge University Press).
- [26] C. A. Tracy and H. Widom, *Introduction to random matrices*, Lectures at 8th Scheveningen Conference (1992).

- [27] A. Edelman and N. Raj Rao, *Random matrix theory*, Acta Numerica 1 (2005).
- [28] E. Brézin, C. Itzykson, G. Parisi, and J.-B. Zuber, *Planar diagrams*, Commun. Math. Phys. **59**, 35 (1978).
- [29] S. F. Edwards and R. C. Jones, *The eigenvalue spectrum of a large symmetric random matrix*, J. Phys. A: Math. Gen. **9**, 1595 (1976).