Statistical Mechanics and Random Matrices

Alice Guionnet

IAS/Park City Mathematics Series
Volume XX, XXXX

Statistical Mechanics and Random Matrices

Alice Guionnet

UMPA, ENS Lyon, 46 allée d'Italie, 69364 Lyon Cedex 07, FRANCE.
E-mail address: aguionne@umpa.ens-lyon.fr

Contents

Statistical Mechanics and Random Matrices
Alice Guionnet 1

Statistical Mechanics and Random Matrices 3
1. Introduction 6
2. Motivations 7
3. The different scales; typical results 12

Lecture 1. Wigner matrices and moments estimates 15
1. Wigner's theorem 16
2. Words in several independent Wigner matrices 23
3. Estimates on the largest eigenvalue of Wigner matrices 25

Lecture 2. Gaussian Wigner matrices and Fredholm determinants 27
1. Joint law of the eigenvalues 28
2. Joint law of the eigenvalues and determinantal law 29
3. Determinantal structure and Fredholm determinants 31
4. Fredholm determinant and asymptotics 31

Lecture 3. Wigner matrices and concentration inequalities 35
1. Concentration inequalities and logarithmic Sobolev inequalities 36
2. Smoothness and convexity of the eigenvalues of a matrix and of traces of matrices 39
3. Concentration inequalities for random matrices 42
4. Brascamp-Lieb inequalities; Applications to random matrices 43

Lecture 4. Matrix models 49
1. Combinatorics of maps and non-commutative polynomials 51
2. Formal expansion of matrix integrals 55
3. First order expansion for the free energy 59
4. Discussion 66

Lecture 5. Random matrices and dynamics 69
1. Free Brownian motions and related stochastic differential calculus 70
2. Consequences 76
3. Discussion 78

Bibliography 81

1. Introduction

In these lecture notes, we wish to show how classical ideas and tools from statistical mechanics can be used in the framework of random matrices. The lecture that is more specific to random matrices is Lecture 2 where we show how the eigenvalues of complex Gaussian matrices have a determinantal law and where we relate their statistics with Fredholm determinants. The other lectures are based on the following ideas from probability and statistical mechanics;

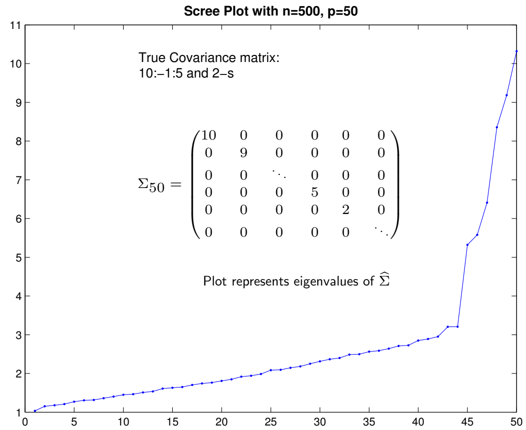
- In Lecture 1, we use the **method of moments** to show the convergence of the spectral measure of random matrices. This type of techniques is quite popular in random matrices theory since random matrices are often modeled through their entries, the joint law of the eigenvalues being in general not explicit (whereas a moment under the spectral measure of a matrix is the trace of a power of the matrix, and therefore an explicit function of the entries). We shall show some generalizations of these techniques.
- In Lecture 3, we use **concentration theory** to obtain concentration inequalities for matrix quantities such as the spectral measure or the largest eigenvalue.
- In Lecture 4, we study natural **Gibbs measures** for random matrices. These Gibbs measures can be seen as small perturbations of the law of independent Gaussian matrices. Because we are typically in a "high temperature regime", uniqueness of saddle points and concentration inequalities will be the key tools of the proofs.
- In Lecture 5, we shall use **dynamics** to improve the results of the last section and obtain non-perturbative results, as well as study the previous saddle points. This analysis takes place in the natural 'limiting setting' of large random matrices provided by free probability; we, however, reduced the 'free' part of the theory to its minimum in these notes.

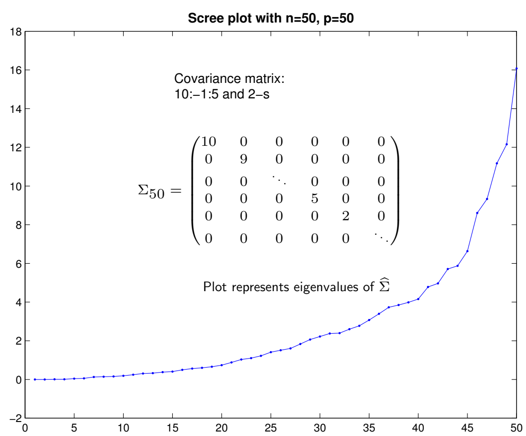
Even though the ideas and techniques developed in these notes are often borrowed from standard probability or statistical mechanics, the motivations behind the results we shall obtain are quite specific. In the next section, we shall review some of the typical motivations and questions in the field.

2. Motivations

Large random matrices have been studied since the thirties when Wishart [**105**] considered them to analyze some statistical problems. Since then, random matrices appeared in various fields of mathematics. Let us briefly summarize some of them and the mathematical questions they raised.

(1) **Large random matrices and statistics**: In 1928, Wishart [**105**] considered matrices of the form $\mathbf{Y}^{N,M} = \mathbf{X}^{N,M}(\mathbf{X}^{N,M})^*$ with an $N \times M$ matrix $\mathbf{X}^{N,M}$ with random entries. Typically, the matrix $\mathbf{X}^{N,M}$ is made of independent equidistributed vectors $\{X^1, \cdots, X^N\}$ in $\mathbb{C}^M$ with covariance matrix Σ, $(\Sigma)_{ij} = \mathbb{E}[X_i^1 X_j^1]$ for $1 \leq i, j \leq M$. Such random vectors naturally appear in multivariate analysis context where $\mathbf{X}^{N,M}$ is a data matrix the column vectors of which represent an observation of a vector in $\mathbb{C}^M$. In such a setup, one would like to find the effective dimension of the system, that is the smallest dimension with which one can encode all the variations of the data. Such a principal components analysis is based on the study of the eigenvalues and eigenvectors of the covariance matrix $\mathbf{Y}^{N,M}$. A central question in this domain is also to estimate the matrix Σ from the observation of $\mathbf{X}^{N,M}$. When M is much smaller than N, it is rather clear that Σ can well be estimated by the empirical sums $\hat{\Sigma}_{ij} := N^{-1} \sum_{i=1}^N X_i^\ell X_j^\ell$ according to the law of large numbers. Moreover, the eigenvalues of the random matrix $X^{N,M}$ have a few large eigenvalues and many small ones, suggesting that the variations of the data takes mainly place in the eigenspace corresponding to these few large eigenvalues (see below the figure, kindly provided by N. El Karoui, representing the eigenvalues of $\mathbf{Y}^{N,M}$ when $\mathbf{X}^{N,M}$ is Gaussian with covariance matrix Σ). It used to be reasonable to assume that N/M was large. However, the case where N/M is of order one is nowadays commonly considered; it corresponds to the cases where either the number of observations is rather small or when the dimension of the observation is very large. In this case, the picture is much less transparent and the above estimator $\hat{\Sigma}$ is biased. Moreover, the spectrum is much more continuous (see the figure below, provided by N. El Karoui) and analysis in principal components need deeper thoughts.





Tests can for instance be done by observing the k largest eigenvalues, to decide whether the distribution of the eigenvalues can resemble the distribution of the eigenvalues of a matrix $\mathbf{Y}^{N,M}$ with the $\mathbf{X}^{N,M}(ij)$ independent equidistributed Gaussian variables, see [**67**] (corresponding to Σ being a multiple of the identity). Hence, the study of the local properties of the spectrum as well as the related eigenvectors is particularly interesting (see [**39, 99**] and references therein).

Similar questions on random matrices arise in finance to model portfolio optimization (see e.g. the work of Bouchaud et al [**86**]). In the same spirit, random matrices appear in problems related with telecommunications and more precisely with the analysis of cellular phones data where a very large number of customers have to be treated simultaneously. The problem is to retrieve the signal from a noised observation, often assuming that the observation is a linear function of the signal (see [**99**] and references therein).

The smallest eigenvalue of $\mathbf{Y}^{N,M}$ is as well of great interest, for instance because when it is sufficiently apart from the origin, the spectral measure (or Brown measure) of $\mathbf{X}^{N,M}$ is then intimately related with the moments of $\mathbf{X}^{N,M}$, a fact that was used in [**8, 95**] to prove its convergence to the circular law.

In this setting, the main questions concern local properties of the spectrum (such as the study of the large N, M behavior of the spectral radius of $\mathbf{Y}^{N,M}$, see [**67**], of its smallest eigenvalue [**8, 95**], the asymptotic behavior of the k largest eigenvalues etc), or the form of the eigenvectors of $\mathbf{Y}^{N,M}$, but also macroscopic questions concerning the estimation of Σ.

(2) **Large random matrices and quantum mechanics**: Wigner, in 1951 [**104**], suggested to approximate the Hamiltonians of highly excited nuclei by large random matrices. The basic idea is that there are so many phenomena going on in such systems that they can

not be analyzed exactly and only a statistical approach becomes reasonable. The random matrices should be chosen as randomly as possible within the known physical restrictions of the model. For instance, Wigner considered Hermitian (since the Hamiltonian has to be Hermitian) matrices with i.i.d entries (modulo the symmetry constraint). In the case where the system is invariant by time inversion, one can consider real symmetric matrices etc... As Dyson pointed out, the general idea is to choose the most random model within the imposed symmetries and to check if the theoretical predictions agree with the experiment, a disagreement pointing out that an important symmetry of the problem has been neglected.

It turned out that experiments agreed exceptionally well with these models ; for instance, it was shown that the energy states of the atom of hydrogen submitted to a strong magnetic field can be compared with the eigenvalues of a symmetric matrix with i.i.d Gaussian entries (the so-called GOE). The book [**47**] summarizes a few similar experiments as well as the history of random matrices in quantum mechanics.

In quantum mechanics, the eigenvalues of the Hamiltonian represent the energy states of the system. It is therefore important to study, following Wigner, the spectral distribution of the random matrix under study, but even more important, is its spacing distribution which represents the energy gaps. Such questions were addressed in the reference book of M.L. Mehta [**79**]. In the last fifteen years, a rigorous treatment of these questions was given around the work of C. Tracy et H. Widom [**97, 96**] . It is also important to make sure that the results obtained do not depend on the details of the large random matrix models such as the law of the entries ; this important field of investigation is often referred to as universality. An important effort of investigation was made in the last ten years in this direction for instance in [**91, 64, 98, 81**]... For instance, it is now known [**88**] that a real symmetric matrix with independent equidistributed entries with 36^{th} moment finite has a largest eigenvalue with Tracy-Widom fluctuations whereas if these entries have not a finite fourth moments, the largest eigenvalue does not converge anymore to the edge of the limiting spectral measure and has a Frechet limit distribution [**93, 6**]. It is conjectured that this behaviour really only depends on the existence of the moment of order four.

(3) **Large random matrices and Riemann Zeta function:** The Riemann Zeta function is given by

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$$

with $\mathrm{Re}(s) > 1$ and can be analytically continued to the complex plane. The study of the zeroes of this function in the strip $0 \le \mathrm{Re}(s) < 1$ furnishes one of the most famous open problems. It is well known that ζ has trivial zeroes at $-2, -4, -6....$ and that its zeroes are distributed symmetrically with respect to the line $\mathrm{Re}(s) = 2^{-1}$. The Riemann conjecture is that all the non trivial zeroes are located on this line. It was suggested by Hilbert and Polya that these zeroes might be related to the eigenvalues of a Hermitian operator, a fact that would immediately imply that they are aligned. To comfort this idea, H. Montgomery (1972), assuming the Riemann conjecture, studied the number of zeroes of the zeta function in $\mathrm{Re}(s) = 2^{-1}$ up to a distance T of the real axis. His result suggests a striking similarity with corresponding statistics of the distribution of the eigenvalues of random Hermitian or unitary matrices when T is large (in fact the spacing distribution seems to be given by a determinantal law with Sine kernel, as is the case for the eigenvalues of Wigner matrices in the bulk). Since then, an extensive literature was devoted to understand this relation (see e.g. [**70**]). It was in particular discovered by Keating et al [**71**] that moments of characteristic polynomials of random matrices play a key role in predicting the asymptotic behavior of moments of L functions computed to a height T on the critical line, as T goes to infinity. This discovery introduces the idea that not only the zeroes of the zeta functions should be related with the eigenvalues of a random matrix, but the functions themselves behave like random (characteristic) polynomials. In a similar spirit, it was shown that (pseudo)moments of the characteristic polynomial of unitary matrices is related with (pseudo)moments of the Riemann zeta functions and the enumeration of (pseudo)magic squares [**32, 33**].

In somewhat the same direction, there is numerical evidence that the eigenvalues distribution of large Wigner matrices is also related with the large eigenvalues of the Laplacian in some bounded domain such as the cardioid. This is related to quantum chaos since these eigenvalues describe the long time behavior of the classical ray dynamics in this domain (i.e the billiard dynamics).

(4) **Large random matrices and free probability**: Free probability is a probability theory in a non-commutative framework. Probability measures are replaced by tracial states on von Neumann algebras. Free probability also contains the central notion of freeness which can be seen as a non-commutative analogue of the notion of independence. At the algebraic level, it can be related with the usual notion of freeness for groups. This is why free probability could be well suited to solve important questions in von Neumann algebras, such as the question of isomorphism between free group factors. Even though this goal is not yet achieved, let us quote a few results on von Neumann algebras that were proved thanks to free probability machinery [**44, 45, 102**].

In the 1990's, Voiculescu [**101**] proved that large random matrices are asymptotically free as their size goes to infinity. Hence, large random matrices became a source for constructing many non-commutative laws, with nice properties with respect to freeness. Thus, free probability can be considered as the natural asymptotic framework to consider large random matrices as their size goes to infinity. Conversely, if one believes that any tracial state could be approximated by the empirical distribution of large matrices (a notion that we shall define more precisely later), which would answer to a well known question of A. Connes in the affirmative, then any tracial state could be obtained as such a limit.

In this context, one often studies the asymptotic behavior of traces of polynomial functions of several random matrices with size going to infinity, trying to deduce from this limit either intuition or results concerning tracial states. For instance, free probability and large random matrices can be used to construct counter examples to some operator algebra questions [**58**].

(5) **Combinatorics, enumeration of maps and matrix models**: It is well known that the evaluation of the expectation of traces of random matrices possesses a combinatorial nature. For instance, if one considers an $N \times N$ symmetric or Hermitian matrix $\mathbf{X}_N$ with i.i.d centered entries with covariance N^{-1}, Wigner's theorem asserts that $E[N^{-1}\mathrm{Tr}(\mathbf{X}_N^p)]$ converges toward 0 if p is odd and toward the Catalan number $C_{\frac{p}{2}}$ if p is even. C_p is the number of non crossing partitions of $\{1, \cdots, 2p\}$ and arises very often in combinatorics. This idea was pushed forward by J. Harer and D. Zagier [**60**] who computed exactly moments of the trace of $\mathbf{X}_N^p$ to enumerate maps with given number of vertices and genus, and then in the seminal article of Kontsevich [**73**].

This approach of combinatorial problems by using large random matrices is inherited from theoretical physics and takes its inspiration in the seminal work of 't Hooft [**94**]. In this article, 't Hooft showed that Gauge theory with local Gauge group $U(N)$ simplifies in the limit N going to infinity. More precisely, such a theory is often described in quantum field theory by some integral over complicated spaces such as connections (integrals that rarely make proper mathematical sense). It is then customary in quantum field theory to give a meaning to these integrals by performing a formal expansion of all terms that are not quadratic in order to obtain integrals under a Gaussian law. Because of Wick calculus, or equivalently Feynman diagrams, Gaussian expectation can in general be written as a generating function of some graphs (see Lecture 4). 't Hooft discussion shows that if the original system is invariant under the action of $U(N)$ with N as large as wished, then such expansion should only depend on planar graphs. Each correction to this infinite N limit should depend of graphs with a given genus.

This point of view was specified in [**26**] to the case where the integral is perfectly well defined as an integral over complex Gaussian Wigner matrices (whose law is invariant under the action of the unitary group); they considered matrix integrals such as

$$Z_N(P) = E[e^{N\mathrm{Tr}(P(\mathbf{X}_N^1,\cdots,\mathbf{X}_N^k))}]$$

with a polynomial function P and independent copies $\mathbf{X}_N^i$ of an $N \times N$ matrix $\mathbf{X}_N$ with complex Gaussian entries and law invariant under the action of the unitary group (in fact they considered the special case where $k = 1$ and P is a quartic polynomial). Then, one can see that if $P = \sum t_i q_i$ with some (complex) parameters t_i and some monomials q_i, $\log Z_N(P)$ expands formally (as a function of the parameters t_i) and in fact the coefficients of this formal expansion enumerate certain maps. The formal proof follows from Feynmann diagrams expansion. This relation is nicely summarized in an article by A. Zvonkin [**107**] and we shall describe it more precisely in Lecture 4. One-matrix integrals can be used to enumerate various maps of arbitrary genus, and several matrix integrals can serve to consider the case where the edges of these maps are colored, i.e can take different states. For example, two-matrix integrals can therefore serve to define an Ising model on random graphs.

Matrix models were also used in physics to construct string theory models. According to the last remark, since string theory concerns maps with arbitrary genus, matrix models have to be considered at criticality and with temperature parameters well tuned with the dimension in order to have any relevance in this domain (the so-called double scaling limit).

This subject had also a great revival in connection with Gromov-Witten invariants. In fact, some of these invariants were shown to be connected with matrix integrals, for instance in the seminal work of Kontsevich [**73**] who proved that some generating function of some intersection numbers of stable classes of the moduli space of algebraic curves is given by a matrix integral as above. According to Witten, such relations should hold in a much larger generality (we refer also to works of Marino, Vafa, Dijkgraaf etc). In connection with this approach, A. Okounkov and coauthors proved that a discrete analogue of matrix models, provided for instance by tiling models, are related with some other Gromov-Witten invariants.

In this domain, one tries to estimate integrals such as $Z_N(P)$, and, more precisely, to obtain the full expansion of $\log Z_N(P)$ in terms of the dimension. This follows on the formal level by Feynman diagrams techniques. A large dimension expansion (meaning where issues on the convergence of series are addressed) was obtained rigorously for one-matrix models by use of Riemann-Hilbert problem techniques by J. Mc Laughlin and N. Ercolani [**40**] and for several-matrix models by E. Maurel Segala and myself. We shall review this topic in Lecture 4.

Another interest in the relation between random matrices and the enumeration of interesting combinatorial objects comes from numbers theory. This is one of the motivations of the article of Diaconis and Gamburd [**36**] relating the enumeration of magic squares and moments of the secular coefficients of a matrix following the Haar measure on the unitary group (see [**32**]).

Another question is to use the representation of the combinatorial problem in terms of matrix models to actually solve or analyze this problem. First order asymptotics for a few several-matrix models could be obtained by orthogonal polynomial methods by M. L. Mehta [**79, 30**] and by large deviations techniques in [**48**]. We shall discuss this issue in Lecture 4. The physics literature on the subject is much more consistent as can be seen on the arxiv (see work by B. Eynard, P. Di Francesco, V. Kazakov, I. Kostov, M. Staudacher, P. Zinn Justi, J.B Zuber etc).

(6) **Large random matrices, random partitions and determinantal laws**: The laws of the eigenvalues of complex Gaussian matrices have a determinantal form [**79**], i.e., the law of the eigenvalues $(\lambda_1, \cdots, \lambda_N)$ of a Wigner matrix with complex Gaussian entries (also called the **GUE**) is absolutely continuous with respect to Lebesgue measure and the interaction between the eigenvalues described by the square of a Vandermonde determinant Δ.

Because Δ is a determinant, specific techniques can be used to study for instance the law of the top eigenvalue or the spacing distribution in the bulk or next to the top (cf. [**96**], and Lecture 2). Such laws appear actually in diverse contexts such as random partitions [**80, 21**], tilling problems [**66**], longest increasing subsequence [**9, 65**], directed polymers and the totally asymmetric simple exclusion process [**63**]. Extension of this connection allowed recently to study the totally asymmetric exclusion process with various initial conditions [**20, 19**]. These relations are often based on bijections with pairs of Young tableaux. Recently, it was shown [**84**] that the zeroes of random analytic functions also have a determinantal form. More examples will be given in Y. Peres course.

In fact, determinantal laws appear naturally when non-intersecting paths are involved. Indeed, following [**69**](see also Gessel-Viennot in a discrete setting [**46**]), if k_T is the transition probability of a homogeneous continuous Markov process, and P_T^N the distribution of N independent copies $X_t^N = (x_1(t), \cdots, x_N(t))$ of this process, then for any $X = (x_0, \cdots, x_N), x_1 < x_2 < \cdots < x_N$, $Y = (y_0, \cdots, y_N), y_1 < y_2 < \cdots < y_N$,

$$P(X_N(0) = X, X_N(T) = Y | \forall t \geq 0, x_1(t) \leq x_2(t) \leq \cdots x_N(t))$$

$$(1) \qquad = C(x) \det \left((k_T(x_i, y_j))_{\substack{1 \leq i \leq N \\ 1 \leq j \leq N}} \right)$$

with

$$C(x)^{-1} = \int \det \left((k_T(x_i, y_j))_{\substack{1 \leq i \leq N \\ 1 \leq j \leq N}} \right) dy.$$

This might provide an additional motivation to study determinantal laws.

In this set of problems, one often meets the problem of analyzing the asymptotics of the largest particles and/or spacing distribution.

3. The different scales; typical results

As we have seen in the previous section, according to the different settings, one can be interested to study the spectrum of random matrices at different scales. In what follows $(\lambda_1 \leq \lambda_2 \cdots \leq \lambda_N)$ denotes the N eigenvalues of a Hermitian random matrix X^N.

3.1. Macroscopic regime

A typical question is to study empirical quantities in the eigenvalues such as the spectral measure

$$L_N := \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i}.$$

If X^N is a Wigner matrix, that is a symmetric or Hermitian matrix such that $X_{ij}^N = y_{ij}/\sqrt{N}$ for $i \leq j$, with $(y_{ij}, i \leq j)$ an infinite triangular array of independent equidistributed random variables with law μ, we shall see in section 1 the following result due to Wigner [**104**]. If $\mu(x^2) = 1$, L_N converges almost surely towards σ the semi-circular law

$$\sigma(dx) = \frac{1}{\pi} 1_{[-2,2]} \sqrt{4 - x^2} dx.$$

Convergence of a probability measure in these lecture notes will always be understood in the sense of weak convergence, that is μ_n converges towards μ iff $\mu_n(f)$ converges to $\mu(f)$ for all bounded continuous functions.

Such a theorem can be refined in many ways; by proving a central limit limit theorem (see e.g. [**62, 4**]), concentration inequalities (see Lecture 3) or a large deviation principle (see [**10**]). Such results have been proved under more and more restrictive hypothesis; central limit theorem requires a finite fourth moment for μ (see e.g. [**4, 89**]), concentration inequalities need sub-Gaussian, sub -exponential or compactly supported laws μ and the large deviation result has only been proved for Gaussian laws, based on an explicit formula for the joint law of the eigenvalues which only exists in this case (see Lecture 2).

This theorem can be generalized also to other random matrices or to more general observables. For instance, if one considers Wishart matrices $\mathbf{Y}^{N,M} = \mathbf{X}^{N,M}(\mathbf{X}^{N,M})^*$ and denotes L_N the empirical measure of the eigenvalues of $\mathbf{Y}^{N,M}$, if the entries of $\mathbf{X}^{N,M}$ are independent and equidistributed, it is well known (see e.g. [**76, 7**]) that L_N converges, if M/N goes to $y \in [0,1]$, to the Marchenko-Pastur law

$$\pi_y(dx) = \frac{1}{2xy\pi}\sqrt{4y-(x-y-1)^2}1_{[(1-\sqrt{y})^2,(1+\sqrt{y})^2]}dx.$$

Other random matrices appear in different context; band matrices where the entries are independent but their covariance depends on the indices, sparse matrices where entries may be zero with some probability, circular matrices $\mathbf{X}^{N,N}$ where all the entries are independent (leading to a complex spectrum)(see [**8**]).

My point of view, borrowed from free probability, is that such questions can very often be rephrased in terms of the Wigner matrices $\mathbf{X}^N$ provided we consider it jointly with some other matrices. Namely, instead of considering only the spectral measure of a matrix, it is natural to consider it together with a bunch say $(\Delta_i^N, 1 \le i \le m)$ of deterministic matrices and to wonder when $\frac{1}{N}\mathrm{Tr}(P(D_i^N, 1 \le i \le m, \mathbf{X}^N))$ converges as N goes to infinity for some polynomial P in these matrices. One can recover like this most classical ensembles quoted above. For instance, take $D_1^N(ij) = 1_{i=j}1_{i\le[\alpha N]}$ to be the projection on the first $[\alpha N]$ indices. Then, with 1 denoting the identity matrix

$$\mathbf{Z}^N = D_1^N\mathbf{X}^N(1-D_1^N) + (1-D_1^N)\mathbf{X}^N D_1^N = \begin{pmatrix} 0 & X^{N-[\alpha N],[\alpha N]} \\ (X^{N-[\alpha N],[\alpha N]})^* & 0 \end{pmatrix}$$

with $X^{N-[\alpha N],[\alpha N]}$ the corner $(\mathbf{X}^N)_{1\le i\le[\alpha N],[\alpha N]+1\le j\le N}$ of the matrix $\mathbf{X}^N$. Then,

$$(\mathbf{Z}^N)^2 = \begin{pmatrix} X^{N-[\alpha N],[\alpha N]}(X^{N-[\alpha N],[\alpha N]})^* & 0 \\ 0 & (X^{N-[\alpha N],[\alpha N]})^* X^{N-[\alpha N],[\alpha N]} \end{pmatrix}$$

has the eigenvalues of the Wishart matrix $X^{N-[\alpha N],[\alpha N]}(X^{N-[\alpha N],[\alpha N]})^*$ with multiplicity 2 plus $N-2[\alpha N]$ null eigenvalues (if $\alpha \ge 1/2$ so that $N-[\alpha N] \le [\alpha N]$). Many other classical ensembles can be derived in this way. Adopting this point of view is often fruitful because it allows to decipher some general structure such as freeness (see Lectures 1 and 5) which at the end simplifies the analysis of the convergence of moments for instance. The backdraw is that you consider more general objects than the empirical measure of the eigenvalues (which eventually converges towards a probability measure), namely the trace of polynomials in random matrices, which possibly converges towards a linear functional on polynomials, called a tracial state. Analysis of such objects may then require free probability tools.

Exercise 0.1. *The idea of this exercise is to see that also band matrices can often be decomposed as functions of a Wigner matrix and diagonal matrices.*

Take $\sigma(x,y) := \int f_s(x)f_s(y)dp(s)$ for some probability measure p on a probability space (Ω, P) and bounded continuous functions $(f_s, s \in \Omega)$ on $[0,1]$. Show that the band matrix with entries $\sigma(\frac{i}{N},\frac{j}{N})X_{ij}$ can be written as $\int D_s^N\mathbf{X}^N D_s^N dp(s)$ with D_s^N the diagonal matrix with entries $(f_s(\frac{i}{N}), 1 \le i \le N)$ on the diagonal.

3.2. Microscopic regime

At the opposite, one would like to study what happens at a very microscopic regime, for instance study the consecutive spacing distribution $\Delta_i := \lambda_{i+1} - \lambda_i$. This can be done for instance when the matrix $\mathbf{X}_N$ is a Gaussian complex Wigner matrix. Then, as we already noticed, the joint law of the eigenvalue is a determinantal law

$$dP(\lambda_1, \cdots, \lambda_N) = Z_N^{-1}\Delta(\lambda)^2 e^{-\frac{N}{2}\sum_{i=1}^N \lambda_i^2}\prod d\lambda_i$$

with $\Delta(\lambda)$ the Vandermonde determinant. Then, it was proved [**96**] that the spacing distribution around the origin converges to the Sine kernel law, i.e., that for any compact set A

$$\lim_{N\to\infty} P[N\lambda_i^N \notin A, i=1,\ldots,N] = 1 + \sum_{k=1}^{\infty}\frac{(-1)^k}{k!}\int_A\cdots\int_A \det_{i,j=1}^k \tfrac{1}{\pi}\tfrac{\sin(x_i-x_j)}{x_i-x_j}\prod_{j=1}^k dx_j. \tag{2}$$

A similar result holds for the spacing distribution inside the bulk (namely around any point within $(-2,2)$). The picture changes at the boundary where one gets

$$\lim_{N\to\infty} P\left[N^{2/3}\left(\lambda_i^N-2\right)\notin[t,t'], i=1,\cdots,N\right] = 1+\sum_{k=1}^{\infty}\frac{(-1)^k}{k!}\int_t^{t'}\cdots\int_t^{t'}\det_{i,j=1}^{k} A\left(x_i,x_j\right)\prod_{j=1}^{k}dx_j$$

with A the Airy kernel defined by

$$A(x,y) := \frac{Ai(x)Ai'(y)-Ai'(x)Ai(y)}{x-y}$$

if Ai denotes the Airy function.

We shall show the general structure of the proof of this sort of analysis in part 2 via Fredholm determinants (see also the more recent process approach discussed by Ramirez, Rider and Virag [**87**]).

3.3. Mesoscopic regime

It is also natural to wonder what happens in the intermediate scales, for instance how many eigenvalues fall into a set $A_N=[N^{-\alpha}a, N^{-\alpha}b]$ for $\alpha\in(0,1)$. Based also on the determinantal structure of the law it is possible to show that if $a<0<b$,

$$\frac{\sum_{i=1}^N 1_{\lambda_i\in A_N}-\mathbb{E}[\sum_{i=1}^N 1_{\lambda_i\in A_N}]}{\mathbb{E}\left[\left(\sum_{i=1}^N 1_{\lambda_i\in A_N}-\mathbb{E}[\sum_{i=1}^N 1_{\lambda_i\in A_N}]\right)^2\right]^{\frac{1}{2}}}$$

converges in law towards a standard Gaussian variable as N goes to infinity.

We shall not discuss this scaling at all in these notes (such a result can be derived by determinantal law analysis).

Acknowledgments: I am particularly grateful to UC Berkeley, and in particular to D. Voiculescu and V. Jones, who welcomed me during spring 2007 when I wrote these lecture notes. My visit was supported in part by funds from NSF Grants DMS-0405778, DMS-0605166 and DMS-0079945. I also wishes to thank S. Sheffield and T. Spencer for inviting me to give these lectures.

LECTURE 1

Wigner matrices and moments estimates

In this lecture, we shall analyse moments of large random matrices and describe the fine combinatorics needed to evaluate them. This analysis will be used to study the spectral measure and the largest eigenvalue of Wigner's matrices.

1. Wigner's theorem

We consider in this section an $N \times N$ matrix $\mathbf{X}^N$ with real or complex entries such that $\left(\mathbf{X}^N_{ij}, 1 \le i \le j \le N\right)$ are independent and $\mathbf{X}^N$ is self-adjoint; $\mathbf{X}^N_{ij} = \bar{\mathbf{X}}^N_{ji}$. We assume further that

$$\mathbb{E}[\mathbf{X}^N_{ij}] = 0, \ \lim_{N\to\infty} \max_{1\le i,j\le N} |N\mathbb{E}[|\mathbf{X}^N_{ij}|^2] - 1| = 0.$$

We shall show that the eigenvalues $(\lambda_1, \cdots, \lambda_N)$ of $\mathbf{X}^N$ satisfy the almost sure convergence

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N f(\lambda_i) = \int f(x) d\sigma(x) \tag{3}$$

where f is a bounded continuous function or a polynomial function, when the entries have some finite moments properties. σ is the semi-circular law

$$\sigma(dx) = \frac{1}{2\pi}\sqrt{4 - x^2} 1_{|x|\le 2} dx.$$

We shall prove this convergence for polynomial functions and rely on the fact that for all $k \in \mathbb{N}$, $\int x^k d\sigma(x)$ is null when k is odd and given by the Catalan number $C_{k/2}$ when k is even. Deducing (3) from convergence of moments is done in section 1.5.

1.4. Wigner's theorem

In this section, we use the same notation for complex and for real entries since both cases will be treated at once and yield the same result. The aim of this section is to prove

Theorem 1.1. *[Wigner's theorem [**104**]] Assume that for all $k \in \mathbb{N}$,*

$$B_k := \sup_{N\in\mathbb{N}} \sup_{ij\in\{1,\cdots,N\}^2} \mathbb{E}[|\sqrt{N}\mathbf{X}^N_{ij}|^k] < \infty. \tag{4}$$

Then,

$$\lim_{N\to\infty} \frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right) = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ C_{\frac{k}{2}} & \text{otherwise,} \end{cases}$$

where the convergence holds in expectation and almost surely. $(C_k)_{k\ge 0}$ *are the Catalan numbers;*

$$C_k = \frac{\begin{pmatrix} 2k \\ k \end{pmatrix}}{k+1}.$$

The Catalan number C_k will appear here as the number of non-crossing pair partitions of $2k$ elements. Namely, recall that a partition of the (ordered) set $S := \{1, \cdots, n\}$ is a decomposition

$$\pi = \{V_1, \cdots, V_r\}$$

such that $V_i \cap V_j = \emptyset$ if $i \ne j$ and $\cup V_i = S$. The $V_i, 1 \le i \le r$ are called the blocks of the partition and we say that $p \sim_\pi q$ if p, q belong to the same block of the partition π. A partition is said to be a pair partition if each of its block has exactly two elements. A partition π of $\{1, \cdots, n\}$ is said to be *crossing* if there exist $1 \le p_1 < q_1 < p_2 < q_2 \le n$ with

$$p_1 \sim_\pi p_2 \not\sim_\pi q_1 \sim_\pi q_2.$$

It is *non-crossing* otherwise. We give as an exercise to the reader to prove that C_k as given in the theorem is exactly the number of non-crossing pair partitions of $\{1, 2, \cdots, 2k\}$.

Proof. We start the proof by showing the convergence in expectation, for which the strategy is simply to expand the trace over the matrix in terms of its entries. We then use some (easy) combinatorics on trees to find out the main contributing term in this expansion. The almost sure convergence is obtained by estimating the covariance of the considered random variables.

(1) *Expanding the expectation.*

Setting $\mathbf{Y}^N = \sqrt{N}\mathbf{X}^N$, we have

$$\mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right] = \sum_{i_1,\cdots,i_k=1}^{N} N^{-\frac{k}{2}-1}\mathbb{E}[Y_{i_1i_2}Y_{i_2i_3}\cdots Y_{i_ki_1}] \tag{5}$$

where $Y_{ij}, 1 \leq i,j \leq N$, denote the entries of $\mathbf{Y}^N$ (which may eventually depend on N). We denote $\mathbf{i} = (i_1,\cdots,i_k)$ and set

$$P(\mathbf{i}) := \mathbb{E}[Y_{i_1i_2}Y_{i_2i_3}\cdots Y_{i_ki_1}].$$

By (4) and Hölder's inequality, $P(\mathbf{i})$ is bounded by B_k, independently of $\mathbf{i}$ and N. Since the random variables $(Y_{ij}, i \leq j)$ are independent and centered, $P(\mathbf{i})$ equals zero unless for any pair (i_p, i_{p+1}), $p \in \{1,\cdots,k\}$, there exists $l \neq p$ such that $(i_p, i_{p+1}) = (i_l, i_{l+1})$ or (i_{l+1}, i_l). Here, we used the convention $i_{k+1} = i_1$. To find more precisely which set of indices contributes to the first order in the right hand side of (5), we next provide some combinatorial insight into the sum over the indices.

(2) *Connected graphs and trees.*

$V(\mathbf{i}) = \{i_1,\cdots,i_k\}$ will be called the vertices. An edge is a pair (i,j) with $i,j \in \{1,\cdots,N\}^2$. At this point, edges are directed in the sense that we distinguish (i,j) from (j,i) when $j \neq i$ and we shall precise later when we consider undirected edges. We denote by $E(\mathbf{i})$ the collection of the k edges $(e_p)_{p=1}^k = (i_p, i_{p+1})_{p=1}^k$.

We consider the graph $G(\mathbf{i}) = (V(\mathbf{i}), E(\mathbf{i}))$. $G(\mathbf{i})$ is connected since there exists an edge between any two consecutive vertices. Note that $G(\mathbf{i})$ may contain loops (i.e., cycles, for instance edges of type (i,i)) and multiple undirected edges.

The skeleton $\tilde{G}(\mathbf{i})$ of $G(\mathbf{i})$ is the graph $\tilde{G}(\mathbf{i}) = \left(\tilde{V}(\mathbf{i}), \tilde{E}(\mathbf{i})\right)$ where vertices in $V(\mathbf{i})$ appears only once, edges in $E(\mathbf{i})$ are undirected and appear at most once.

In other words, $\tilde{G}(\mathbf{i})$ is the graph $G(\mathbf{i})$ where multiplicities and orientation have been erased. It is connected, as is $G(\mathbf{i})$.

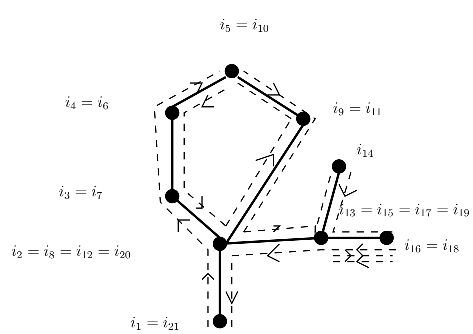


FIGURE 1. Figure of $G(\mathbf{i})$ (in dash) versus $\tilde{G}(\mathbf{i})$ (in bold), $|\tilde{E}(i)| = 9, |\tilde{V}(\mathbf{i})| = 9$

We now state and prove a well known inequality concerning undirected connected graphs $G = (V,E)$. If we let, for a discrete finite set A, $|A|$ be the number of its distinct elements, we have the following inequality

$$|V| \leq |E| + 1. \tag{6}$$

Let us prove this inequality as well as the fact that equality implies that G is a tree. This relation is straightforward when $|V| = 1$ and can be proved by induction as follows. Assume $|V| = n$ and consider one vertex v of V. This vertex is contained in $l \geq 1$ edges of E that we denote $(e_1,\cdots,e_l)$. The graph G then decomposes into $(v, e_1,\cdots,e_l)$ and $r \leq l$ undirected connected graphs $(G_1,\cdots,G_r)$. We denote $G_j = (V_j, E_j)$ for $j \in \{1,\cdots,r\}$. We have

$$|V| - 1 = \sum_{j=1}^{r}|V_j|, \quad |E| - l = \sum_{j=1}^{r}|E_j|.$$

Applying the induction hypothesis to the graphs $(G_j)_{1\le j\le r}$ gives

$$\begin{aligned} |V|-1 &\le \sum_{i=1}^{r}(|E_j|+1) \\ &= |E|+r-l \le |E| \end{aligned} \tag{7}$$

which proves (6). In the case where $|V|=|E|+1$, we claim that G is a tree, namely does not have loop. In fact, for the equality to hold, we need to have equalities when performing the previous decomposition of the graph, a decomposition that can be reproduced until all vertices have been considered. If the graph contains a loop, the first time that we erase a vertex of this loop when performing this decomposition, we will create one connected component less than the number of edges we erased and so a strict inequality occurs in the right hand side of (7) (i.e., $r<l$).

(3) *Convergence in expectation.*

Since we noticed that $P(\mathbf{i})$ equals zero unless each edge in $E(\mathbf{i})$ is repeated at list twice, we have that

$$|\tilde{E}(\mathbf{i})| \le 2^{-1}|E(\mathbf{i})| = \frac{k}{2},$$

and so by (6) applied to the skeleton $\tilde{G}(\mathbf{i})$ we find

$$|\tilde{V}(\mathbf{i})| \le [\frac{k}{2}]+1$$

where $[x]$ is the integer part of x. Thus, since the indices are chosen in $\{1,\cdots,N\}$, there are at most $N^{[\frac{k}{2}]+1}$ indices that contribute to the sum (5) and so we have

$$\left|\mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right]\right| \le B_k N^{[\frac{k}{2}]-\frac{k}{2}}.$$

where we used (4). In particular, if k is odd,

$$\lim_{N\to\infty}\mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right]=0.$$

If k is even, the only indices that will contribute to the first order asymptotics in the sum are those such that

$$|\tilde{V}(\mathbf{i})| = \frac{k}{2}+1,$$

since the other indices will be such that $|\tilde{V}(\mathbf{i})|\le\frac{k}{2}$ and so will contribute at most by a term $N^{\frac{k}{2}}B_kN^{-\frac{k}{2}-1}=O(N^{-1})$. By the previous considerations, when $|\tilde{V}(\mathbf{i})|=\frac{k}{2}+1$, we have that

(a) $\tilde{G}(\mathbf{i})$ is a tree,

(b) $|\tilde{E}(\mathbf{i})|=2^{-1}|E(\mathbf{i})|=\frac{k}{2}$ and so each edge in $E(\mathbf{i})$ appears exactly twice.

We can explore $G(\mathbf{i})$ by following the path P of edges $i_1\to i_2\to i_3\cdots\to i_k\to i_1$. Since $\tilde{G}(\mathbf{i})$ is a tree, $G(\mathbf{i})$ appears as a fat tree where each edge of $\tilde{G}(\mathbf{i})$ is repeated exactly twice. We then see that each pair of directed edges corresponding to the same undirected edge in $\tilde{E}(\mathbf{i})$ is of the form $\{(i_p,i_{p+1}),(i_{p+1},i_p)\}$ (since otherwise the path of edges has to form a loop to return to i_0). Therefore, for these indices, $P(\mathbf{i})=\prod_{e\in\tilde{E}(\mathbf{i})}E[|\sqrt{N}X_e^N|^2]$ converges uniformly to one by hypothesis.

Finally, observe that $G(\mathbf{i})$ gives a pair partition of the edges of the path P (since each undirected edges have to appear exactly twice) and that this partition is non crossing (as can be seen by unfolding the path keeping track of the pairing between edges by drawing an arc between paired edges). Therefore we have proved

$$\lim_{N\to\infty}\mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right]=\sharp\{\text{ non-crossing pair partitions of } k \text{ edges }\}.$$

(4) *Almost sure convergence.*

To prove the almost sure convergence, we estimate the variance and then use Borel Cantelli's lemma. The variance is given by

$$\begin{aligned}\mathrm{Var}((\mathbf{X}^N)^k) &:= \mathbb{E}\left[\frac{1}{N^2}\left(\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right)^2\right] - \mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right]^2 \\ &= \frac{1}{N^{2+k}} \sum_{\substack{i_1,\dots,i_k=1 \\ i'_1,\dots,i'_k=1}}^{N} [P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')]\end{aligned}$$

with

$$P(\mathbf{i},\mathbf{i}') := \mathbb{E}[Y_{i_1 i_2} Y_{i_2 i_3} \cdots Y_{i_k i_1} Y_{i'_1 i'_2} \cdots Y_{i'_k i'_1}].$$

We denote $G(\mathbf{i},\mathbf{i}')$ the graph with vertices $V(\mathbf{i},\mathbf{i}') = \{i_1,\cdots,i_k,i'_1,\cdots,i'_k\}$ and edges $E(\mathbf{i},\mathbf{i}') = \{(i_p,i_{p+1})_{1\le p\le k},(i'_p,i'_{p+1})_{1\le p\le k}\}$. For $\mathbf{i},\mathbf{i}'$ to contribute to the sum, $G(\mathbf{i},\mathbf{i}')$ must be connected. Indeed, if $E(\mathbf{i})\cap E(\mathbf{i}')=\emptyset$, $P(\mathbf{i},\mathbf{i}') = P(\mathbf{i})P(\mathbf{i}')$. Moreover, as before, each edge must appear at least twice to give a non zero contribution so that $|\tilde{E}(\mathbf{i},\mathbf{i}')|\le k$. Therefore, we are in the same situation as before, and if $\tilde{G}(\mathbf{i},\mathbf{i}') = (\tilde{V}(\mathbf{i},\mathbf{i}'),\tilde{E}(\mathbf{i},\mathbf{i}'))$ denotes the skeleton of $G(\mathbf{i},\mathbf{i}')$, we have the relation

$$|\tilde{V}(\mathbf{i},\mathbf{i}')| \le |\tilde{E}(\mathbf{i},\mathbf{i}')| + 1 \le k+1. \tag{8}$$

This already shows that the variance is at most of order N^{-1} (since $P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')$ is bounded uniformly, independently of $(\mathbf{i},\mathbf{i}')$ and N), but we need a slightly better bound to prove the almost sure convergence. To improve our bound let us show that the case where $|\tilde{V}(\mathbf{i},\mathbf{i}')| = |\tilde{E}(\mathbf{i},\mathbf{i}')|+1 = k+1$ can not occur. In this case, we have seen that $\tilde{G}(\mathbf{i},\mathbf{i}')$ must be a tree since then equality holds in (8). Also, $|\tilde{E}(\mathbf{i},\mathbf{i}')| = k$ implies that each edge appears with multiplicity exactly equals to 2. For any contributing set of indices $\mathbf{i},\mathbf{i}'$, $\tilde{G}(\mathbf{i},\mathbf{i}')\cap G(\mathbf{i})$ and $\tilde{G}(\mathbf{i},\mathbf{i}')\cap G(\mathbf{i}')$ must share at least one edge (i.e., one edge must appear with multiplicity one in each of this subgraph) since otherwise $P(\mathbf{i},\mathbf{i}') = P(\mathbf{i})P(\mathbf{i}')$. This is a contradiction. Indeed, if we explore $\tilde{G}(\mathbf{i},\mathbf{i}')$ by following the path $i_1 \to i_2 \to \cdots \to i_1$, we see that either each (non-oriented) visited edge appears twice, which is impossible if $\tilde{G}(\mathbf{i},\mathbf{i}')\cap G(\mathbf{i})$ and $\tilde{G}(\mathbf{i},\mathbf{i}')\cap G(\mathbf{i}')$ share one edge, or it this path makes a loop, which is also impossible since $\tilde{G}(\mathbf{i},\mathbf{i}')$ is a tree. Therefore, we conclude that for all contributing indices,

$$|\tilde{V}(\mathbf{i},\mathbf{i}')| \le k$$

which implies

$$\mathrm{Var}((\mathbf{X}^N)^k) \le p_k N^{-2}$$

with p_k a uniform bound on $P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')$. Applying Chebychev's inequality gives for any $\delta>0$

$$\mathbb{P}\left(\left|\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right) - \mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right]\right| > \delta\right) \le \frac{p_k}{\delta^2 N^2},$$

and so Borel-Cantelli's lemma implies

$$\lim_{N\to\infty}\left|\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right) - \mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left((\mathbf{X}^N)^k\right)\right]\right| = 0 \quad a.s.$$

The proof of the theorem is complete.

□

Exercise 1.2. *Take for $L\in\mathbb{N}$, $\mathbf{X}^{N,L}$ the $N\times N$ self-adjoint matrix such that*

$$\mathbf{X}^{N,L}_{ij} = (2L)^{-\frac{1}{2}} 1_{|i-j|\le L} X_{ij}$$

with $(X_{ij}, 1 \le i \le j \le N)$ *independent centered random variables having all moments finite and* $E[X_{ij}^2] = 1$. *The purpose of this exercise is to show that for all* $k \in \mathbb{N}$,

$$\lim_{L\to\infty}\lim_{N\to\infty} \mathbb{E}[\frac{1}{N}\mathrm{Tr}((\mathbf{X}^{N,L})^k)] = C_{k/2}$$

with C_x *null if* x *is not integer. Moreover, if* $L(N) \in \mathbb{N}$ *is a sequence going to infinity with* N *so that* $L(N)/N$ *goes to zero, prove that*

$$\lim_{N\to\infty} \mathbb{E}[\frac{1}{N}\mathrm{Tr}((\mathbf{X}^{N,L(N)})^k)] = C_{k/2}.$$

If $L(N) = [\alpha N]$, *one can also prove the convergence of the moments of* $\mathbf{X}^{N,L(N)}$. *Show that this limit can not be given by the Catalan numbers* $C_{k/2}$ *by considering the case* $k = 2$.

Hint: Show that for $k \ge 2$

$$\mathbb{E}[\frac{1}{N}\mathrm{Tr}((\mathbf{X}^{N,L})^k)] = (2L)^{-k/2} \sum_{\substack{|i_2-[\frac{N}{2}]|\le L,\\ |i_{p+1}-i_p|\le L, p\ge 2}} \mathbb{E}[X_{[\frac{N}{2}]i_2}\cdots X_{i_k[\frac{N}{2}]}] + O(N^{-1}).$$

Then prove that the contributing indices to the above sum correspond to the case where $G(0, i_2, \cdot, i_k)$ *is a tree with* $k/2$ *vertices and show that being given a tree there are approximately* $(2L)^{\frac{k}{2}}$ *possible choices of indices* $i_2, \cdots, i_k$.

1.5. To learn more

1.5.1. *Catalan numbers.* We can also define the Catalan numbers as the number of (oriented) rooted trees. Actually, Catalan numbers count many other combinatorial objects, such as non-crossing partitions or Dick paths. We shall see that they also give the moments of the semi-circular law.

Let us recall that a graph is given by a set of vertices (or nodes) $V = \{i_1, \cdots, i_k\}$ and a set of edges $(e_i)_{i\in I}$. An edge is a couple $e = (i_{j_1}, i_{j_2})$ for some $j_1, j_2 \in \{1, \cdots, k\}^2$. An edge is directed if (i_1, i_2) and (i_2, i_1) are distinct when $i_1 \ne i_2$, which amounts to write edges as directed arrows. It is undirected otherwise. A cycle (or loop) is a collection of distinct undirected edges $e_i = (v_i, v_{i+1})$, $1 \le i \le p$ such that $v_1 = v_{p+1}$ for some $p \ge 1$.

A *tree* is a connected graph with no loops (or cycles).

We will say that a tree is oriented if it is drawn (or embedded) into the plane; it then inherits the orientation of the plane. A tree is rooted if we specify one oriented edge, called the root. Note that if each edge of an oriented tree is seen as a double (or fat) edge, the connected path drawn from these double edges surrounding the tree inherits the orientation of the plane (see Figure 2). A root on this oriented tree then specifies a starting point in this path.

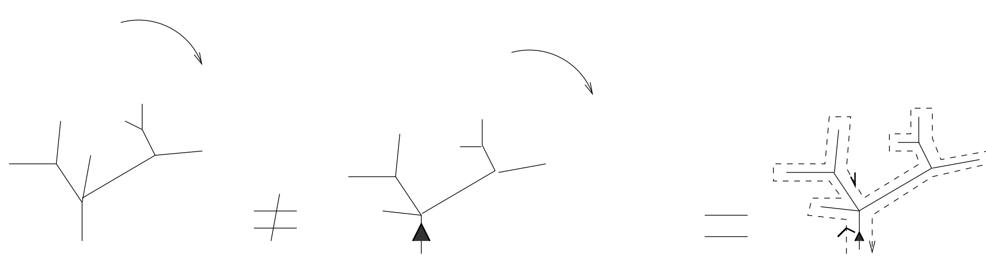


Figure 2. Embedding rooted trees into the plane

Definition 1.3. *We denote* C_k *the number of rooted oriented trees with* k *edges.*

Exercise 1.4. *Show that* $C_2 = 2$ *and* $C_3 = 5$ *by drawing the corresponding graphs.*

Exercise 1.5. *A Dick path of length* $2n$ *is a path starting and ending at the origin, with increments* $+1$ *or* -1, *and that stays above the non negative real axis. Prove that there exists a bijection between the set of rooted oriented trees with* n *edges and the set of Dick paths of length* $2n$.

Hint: Define the walk as the walk around the tree of Figure 2, count $+1$ *when you arrive to an edge that was not visited,* -1 *otherwise*

The following property of the Catalan numbers will be useful later

Property 1.6. *The standard semicircle law is given by*

$$\sigma(dx) = \frac{1}{2\pi}\sqrt{4-x^2}\mathbf{1}_{|x|\le 2}dx.$$

Then for all $k \geq 0$*,*

$$m_{2k} = C_k.$$

A proof can be given following the exercises below.

Exercise 1.7. *Show that with the convention* $C_0 = 1$*, for all* $k \geq 1$

$$C_k = \sum_{l=0}^{k-1} C_{k-l-1}C_l. \tag{9}$$

Exercise 1.8. *For all* $k \geq 0$*,* $C_k \leq 2^{2k}$ *and*

$$C_k = \frac{\begin{pmatrix} 2k \\ k \end{pmatrix}}{k+1}.$$

Count Dick paths or use the induction relation of (9) *to compute the generating function* $S(z) = \sum_{n\geq 0} z^n C_n$

Exercise 1.9. *Prove Property 1.6 by deriving an explicit formula for the* m_k*'s.*

1.5.2. *Weak convergence of the spectral measure.* We now consider weak convergence of the spectral measure rather than convergence in moments and then weaken the hypothesis on the entries.

Theorem 1.10. *Let* $(\lambda_i)_{1\le i\le N}$ *be the* N *(real) eigenvalues of* $\mathbf{X}^N$ *and define*

$$L_{\mathbf{X}^N} := \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}$$

to be the spectral measure of $\mathbf{X}^N$*.* $L_{\mathbf{X}^N}$ *belongs to the set* $\mathcal{P}(\mathbb{R})$ *of probability measures on* $\mathbb{R}$*. Assume that* (4) *holds for all* $k \in \mathbb{N}$*. Then, for any bounded continuous function* f*,*

$$\lim_{N\to\infty}\int f(x)dL_{\mathbf{X}^N}(x) = \int f(x)d\sigma(x) \quad a.s.$$

Proof. By Weierstrass' theorem, we can find for any $B > 2$ and $\delta > 0$, a polynomial P_δ such that $g_\delta := f - P_\delta$ satisfies

$$\sup_{|x|\le B}|g_\delta(x)| \le \delta.$$

Using the previous convergence in moments, one shows that for any $q \in \mathbb{N}$,

$$|\int_{|x|\ge B} g_\delta(x)dL_{\mathbf{X}^N}(x)| \le C\int_{|x|\ge B}(1+|x|^p)dL_{\mathbf{X}^N}(x) \le CB^{-p-2q}\int[1+x^{2(p+q)}]dL_{\mathbf{X}^N}(x)$$

is as small as wished when N goes to infinity and $B > 2$ since the right hand side is then bounded by $B^{-p-2q}2^{2(p+q+1)}$ (since σ is supported in $[-2,2]$) that goes to zero as p goes to infinity. Consequently,

$$\begin{aligned}\left|\int f(x)d(L_{\mathbf{X}^N}(x)-\sigma(x))\right| &\le \left|\int P_\delta(x)d(L_{\mathbf{X}^N}(x)-\sigma(x))\right| \\ &\quad +\delta + \left|\int_{|x|\ge B}(f-P_\delta)(x)dL_{\mathbf{X}^N}(x)\right|\end{aligned} \tag{10}$$

goes to zero as N goes to infinity. ∎

1.5.3. *Relaxation over the number of finite moments.* In this section, we relax the assumptions on the moments of the entries while keeping the hypothesis that $(X_{ij}^N)_{1\le i\le j\le N}$ are independent. The generalization of Wigner's theorem to possibly mildly dependent entries can be found for instance in [**24**]. A nice, simple, but finally optimal way to relax the assumption that the entries of $\sqrt{N}\mathbf{X}^N$ possess all their moments, relies on the following observation.

Lemma 1.11. *Let A, B be $N\times N$ Hermitian matrices, with eigenvalues $\lambda_1(A)\ge\lambda_2(A)\ge\ldots\ge\lambda_N(A)$ and $\lambda_1(B)\ge\lambda_2(B)\ge\ldots\ge\lambda_N(B)$. Then,*

$$\sum_{i=1}^N |\lambda_i(A)-\lambda_i(B)|^2 \le \mathrm{Tr}(A-B)^2\,.$$

The proof is left to the reader; an idea is to observe that this inequality means that the maximum over matrices A,B with a given spectrum of the right hand side is achieved when the two matrices have the same basis of eigenvectors and more precisely the k-th eigenvector correspond to the k-th largest eigenvalues of the matrices. This fact can be shown by induction over the dimension N of the matrices (see [**13, 3**]).

Corollary 1.12. *Assume that $\{\sqrt{N}\mathbf{X}_{ij}^N, i\le j\}$ are independent, equidistributed with law μ such that $\mu(x)=0$, $\mu(x^2)=1$. Then, for any bounded continuous function f*

$$\lim_{N\to\infty}\int f(x)dL_{\mathbf{X}^N}(x)=\int f(x)d\sigma(x)\quad a.s.$$

The proof is left to the reader; it amounts to approximate the original matrix $\sqrt{N}\mathbf{X}^N$ by a matrix $\sqrt{N}\mathbf{Y}^N$ with bounded entries in such a way that $\frac{1}{N}\mathrm{Tr}(\mathbf{X}^N-\mathbf{Y}^N)^2$ goes to zero as N goes to infinity and then use Lemma 1.11.

Remark. When the entries are not equidistributed, the convergence in probability can be proved when $\{\sqrt{N}\mathbf{X}_{ij}^N, i\le j\}$ are uniformly integrable. The almost sure convergence can be proved when moments of order four are uniformly bounded for instance.

Remark Let us remark that if $\sqrt{N}X^N(ij)$ has no moments of order 2, then the theorem is not valid anymore (see the heuristics of Cizeau-Bouchaud [**31**] and rigorous studies in [**106, 11**]). Even though under some assumptions the spectral measure of the matrix X^N, once properly normalized, converges, its limit is not the semicircle law but a heavy tailed law with unbounded support.

1.5.4. *Relaxation of the hypothesis on the centering of the entries.* A last generalization concerns the hypothesis on the mean of the variables $\sqrt{N}X_{ij}^N$ which, as we shall see, is irrelevant in the statement of Corollary 1.12. More precisely, we shall prove that (proof originated from [**57**])

Lemma 1.13. *Let X^N, Y^N be $N\times N$ Hermitian matrices for $N\in\mathbb{N}$ such that $\mathbf{Y}^N$ has rank $r(N)$. Assume that $N^{-1}r(N)$ converges to zero as N goes to infinity. Then, for any bounded continuous function f with compact support,*

$$\limsup_{N\to\infty}|\int f(x)dL_{\mathbf{X}^N+\mathbf{Y}^N}(x)-\int f(x)dL_{\mathbf{X}^N}(x)|=0.$$

Proof. We first prove the statement for bounded increasing functions. To this end, we shall first prove that for any Hermitian matrix $\mathbf{Z}^N$, any $e\in\mathbb{C}^N$, $\lambda\in\mathbb{R}$, and for any bounded measurable increasing function f,

$$\left|\int f(x)dL_{\mathbf{Z}^N}(x)-\int f(x)dL_{\mathbf{Z}^N+\lambda ee^*}(x)\right|\le\frac{2}{N}\|f\|_\infty. \tag{11}$$

We denote $\lambda_1^N\le\lambda_2^N\cdots\le\lambda_N^N$ (resp. $\eta_1^N\le\eta_2^N\cdots\le\eta_N^N$) the eigenvalues of $\mathbf{Z}^N$ (resp. $\mathbf{Z}^N+\lambda ee^*$). By the following theorem due to Lidskii

Theorem 1.14. *[Lidskii] Let $A\in\mathcal{H}_N^{(2)}$ and $z\in\mathbb{C}^N$. We order the eigenvalues of $A\pm zz^*$ in increasing order. Then*

$$\lambda_k(A\pm zz^*)\le\lambda_{k+1}(A)\le\lambda_{k+2}(A\pm zz^*).$$

As a consequence, the eigenvalues λ_i and η_i are interlaced;

$$\lambda_1^N\le\eta_2^N\le\lambda_3^N\cdots\le\lambda^N_{2[\frac{N-1}{2}]+1}\le\eta^N_{2[\frac{N}{2}]}.$$

$$\eta_1^N \leq \lambda_2^N \leq \eta_3^N \cdots \leq \eta^N_{2[\frac{N-1}{2}]+1} \leq \lambda^N_{2[\frac{N}{2}]}.$$

Therefore, if f is an increasing function,

$$\sum_{i=1}^N f(\lambda_i^N) \leq \sum_{i=2}^N f(\eta_i^N) + \frac{1}{N}\|f\|_\infty \leq \sum_{i=1}^N f(\eta_i^N) + \frac{2}{N}\|f\|_\infty$$

but also

$$\sum_{i=1}^N f(\lambda_i^N) = f(\lambda_1^N) + \sum_{i=2}^N f(\lambda_i^N) \geq f(\lambda_1^N) + \sum_{i=2}^N f(\eta_{i-1}^N) = f(\lambda_1^N) - f(\eta_i^N) + \sum_{i=1}^N f(\eta_i^N)$$

These two bounds prove (11). We leave the reader extend this result from $\mathbf{Y}^N = \lambda e e^*$ with rank 1 to $\mathbf{Y}^N$ with rank $r(N)$.

□

By Corollary 1.12 and Lemma 1.13, we find that

Corollary 1.15. *Assume that the matrix $\left(\mathbb{E}[\mathbf{X}_{ij}^N]\right)_{1\leq i,j\leq N}$ has rank $r(N)$ so that $N^{-1}r(N)$ goes to zero as N goes to infinity, and that the variables $\sqrt{N}(X_{ij}^N - \mathbb{E}[X_{ij}^N])$ satisfy the hypotheses of Corollary 1.12 and have covariance 1. Then, for any bounded continuous function f,*

$$\lim_{N\to\infty} \int f(x) dL_{\mathbf{X}^N}(x) = \int f(x) d\sigma(x) \quad a.s.$$

This result holds in particular if $\mathbb{E}[\mathbf{X}_{ij}^N] = x^N$ is independent of $i,j \in \{1,\cdots,N\}^2$, in which case $r(N)=1$. It extends to the case where $\mathbb{E}[\mathbf{X}_{ij}^N] = x^N 1_{i\neq j} + y^N 1_{i=j}$ with y^N going to zero as N goes to infinity.

The last comment is simply due to the fact that $\int f(x) d(L_{\mathbf{X}^N} - L_{\mathbf{X}^N - y^N I}$ goes to zero by Lemma 1.11 when y^N goes to zero.

2. Words in several independent Wigner matrices

In this section, we consider m independent Wigner $N\times N$ matrices $\{\mathbf{X}^{N,\ell}, 1\leq \ell\leq m\}$ with real or complex entries. In other words, the $\mathbf{X}^{N,\ell}$ are self-adjoint random matrices with independent entries $\left(\mathbf{X}_{ij}^{N,\ell}, 1\leq i\leq j\leq N\right)$ above the diagonal that are centered and with variance one. Moreover, the $\left(\mathbf{X}_{ij}^{N,\ell}, 1\leq i\leq j\leq N\right)_{1\leq \ell\leq m}$ are independent. We shall generalize Theorem 1.17 to the case where one considers words in several matrices, that is show that $N^{-1}\mathrm{Tr}\left(\mathbf{X}^{N,\ell_1}\mathbf{X}^{N,\ell_2}\cdots\mathbf{X}^{N,\ell_k}\right)$ converges for all choices of $\ell_i \in \{1,\cdots,m\}$ and give a combinatorial interpretation of the limit. We generalize Theorem 1.1 to the context of several matrices as a first step towards part 4. Let us first describe the combinatorial objects that we shall need.

2.6. Partitions of colored elements

Because we now have m different matrices, the partitions that will naturally show up are partitions of elements with m different colors; in the following, each $\ell \in \{1,\cdots,m\}$ will be assigned a color, said 'color ℓ'. Also, because matrices do not commute, the order of the elements is important. This leads us to the following definition.

Definition 1.16. *Let $q(X_1,\cdots,X_m) = X_{\ell_1}X_{\ell_2}\cdots X_{\ell_k}$ be a monomial in m non-commutative indeterminates.*

We define the set $S(q)$ associated with q as the set of k colored points on the real line so that the first point has color ℓ_1, the second one has color ℓ_2 till the last one that has color ℓ_k.

$NP(q)$ is the set of non-crossing pair partitions of $S(q)$ such that two points of $S(q)$ can not be in the same block if they have different colors.

Note that S defines a bijection between non-commutative monomials and set of colored points on the real line (i.e., ordered set of points).

2.7. Voiculescu's theorem

The aim of this section is to prove that if $\{\mathbf{X}^{N,\ell}, 1 \le \ell \le m\}$ are m independent Wigner matrices such that

$$\mathbb{E}[\mathbf{X}^{N,\ell}_{ij}] = 0, \forall 1 \le i,j \le N, 1 \le \ell \le m, \quad \lim_{N\to\infty} \max_{1\le i,j\le N} |N\mathbb{E}[|\mathbf{X}^{N,\ell}_{ij}|^2] - 1| = 0$$

Theorem 1.17. *[Voiculescu [**101**]] Assume that for all $k \in \mathbb{N}$,*

$$B_k := \sup_{1\le \ell \le m} \sup_{N\in\mathbb{N}} \sup_{ij\in\{1,\cdots,N\}^2} \mathbb{E}[|\sqrt{N}\mathbf{X}^{N,\ell}_{ij}|^k] < \infty. \tag{12}$$

Then, for any $\ell_j \in \{1,\cdots,m\}, 1 \le j \le k$,

$$\lim_{N\to\infty} \frac{1}{N}\mathrm{Tr}\left(\mathbf{X}^{N,\ell_1}\mathbf{X}^{N,\ell_2}\cdots\mathbf{X}^{N,\ell_k}\right) = \sigma^m(X_{\ell_1}\cdots X_{\ell_k})$$

where the convergence holds in expectation and almost surely. $\sigma^m(X_{\ell_1}\cdots X_{\ell_k})$ is the number $|NP(X_{\ell_1}\cdots X_{\ell_k})|$ of non-crossing pair partitions of $S(X_{\ell_1}\cdots X_{\ell_k})$.

Remark 1.18. *σ^m, once extended by linearity to all polynomials, is called the law of m free semi-circular variables.*

Proof. The proof is very close to that of Theorem 1.1 and is left to the reader. The only point is to notice that the main contribution is again given by indices described by non-crossing partitions but that now these partitions come with a weight given by a product of covariances that vanishes when edges of different colors have been paired.

Exercise 1.19. *The next exercise concerns a special case of what is called 'Asymptotic freeness' and was proved in greater generality by D. Voiculescu.*

Let $(\mathbf{X}^N_{ij}, 1 \le i \le j \le N)$ be independent real variables and consider $\mathbf{X}^N$ the self-adjoint matrix with this entries. Assume

$$\mathbb{E}[\mathbf{X}^N_{ij}] = 0 \qquad \mathbb{E}[(\sqrt{N}\mathbf{X}^N_{ij})^2] = 1 \quad \forall i \le j.$$

Assume that for all $k \in \mathbb{N}$,

$$B_k = \sup_{N\in\mathbb{N}} \sup_{ij\in\{1,\cdots,N\}^2} \mathbb{E}[|\sqrt{N}\mathbf{X}^N_{ij}|^k] < \infty \tag{13}$$

Let D^N be a deterministic diagonal matrix such that

$$\sup_{N\in\mathbb{N}} \max_{i\le j} |D^N_{ii}| < \infty \qquad \lim_{N\to\infty} \frac{1}{N}\mathrm{Tr}((D^N)^k) = m_k \text{ for all } k \in \mathbb{N}$$

Show that

(1)
$$\lim_{N\to\infty} \mathbb{E}[\frac{1}{N}\mathrm{Tr}(D^N(\mathbf{X}^N)^k)] = C_{k/2}m_1$$

(2) *Prove that*
$$\lim_{N\to\infty} \mathbb{E}[\frac{1}{N}\mathrm{Tr}((D^N)^{l_1}(\mathbf{X}^N)^{k_1}(D^N)^{l_2}(\mathbf{X}^N)^{k_2})]$$
$$= C_{k_1/2}C_{k_2/2}(m_{l_1+l_2} - m_{l_1}m_{l_2}) + C_{(k_1+k_2)/2}m_{l_1}m_{l_2}$$

(3) *(more difficult)Prove in general that*
$$\mathbb{E}[\frac{1}{N}\mathrm{Tr}\left((D^N)^{l_1} - \frac{1}{N}\mathrm{Tr}(D^N)^{l_1}\right)\left((\mathbf{X}^N)^{k_1} - \mathbb{E}[\frac{1}{N}\mathrm{Tr}(\mathbf{X}^N)^{k_1}]\right)$$
$$\cdots\left((D^N)^{l_p} - \frac{1}{N}\mathrm{Tr}(D^N)^{l_p}\right)\left((\mathbf{X}^N)^{k_p} - \mathbb{E}[\frac{1}{N}\mathrm{Tr}(\mathbf{X}^N)^{k_p}]\right)]$$

goes to zero as N goes to infinity for any integer numbers $l_1,\cdots,l_p,k_1,\cdots,k_p$.

Hint: Expand the trace in terms of a weighted sum over the indices and show that the main contribution comes from indices whose associated graph is a tree. Conditioning on the tree, average out the quantities in the D^N and conclude (be careful that the D^N's can come with the same indices but show then that the main contribution comes from independent entries of the $(\mathbf{X}^N)^k_{ii})$'s because of the tree structure).

3. Estimates on the largest eigenvalue of Wigner matrices

In this section, we derive estimates on the largest eigenvalue of a Wigner matrix with real entries

$$X_{ij}^N = \frac{Y_{ij}}{\sqrt{N}}$$

with $(Y_{ij}, 1 \leq i \leq j \leq N)$ independent equidistributed centered random variables with marginal distribution P. The idea is to improve the moments estimates of the previous section.

We shall assume that P is a symmetric law (see the recent article [**82**] for a relaxation of this hypothesis);

$$P(-Y \in .) = P(Y \in .).$$

We take the normalization $E[Y^2] = 1$. Further we make the common assumption that there exists a finite constant c such that for all $k \in \mathbb{N}$,

$$E[Y^{2k}] \leq (ck)^k.$$

We follow the article of S. Sinaï and A. Soshnikov [**90**] to prove the following result :

Theorem 1.20. *[S. Sinaï - A. Soshnikov [**90**]] For all $\epsilon > 0$, all $N \in \mathbb{N}$, there exists a finite function $o(s,N)$ such that $\lim_{N\to\infty} \sup_{N^\epsilon \leq s \leq N^{\frac{1}{2}-\epsilon}} o(s,N) = 0$ and*

$$\mathbb{E}[\mathrm{Tr}((X^N)^{2s})] = \frac{N 2^{2s}}{\sqrt{\pi s^3}}(1 + o(s,N)). \tag{14}$$

In particular, for all $\epsilon > 0$, if we let $\lambda_{max}(X^N)$ denote the spectral radius of X^N,

$$\lim_{N\to\infty} P(|\lambda_{max}(X^N) - 2| \geq \epsilon) = 0.$$

A previous result of the same nature (but under weaker hypothesis (the symmetry hypothesis of the distribution of the entries being removed) under which the moments estimate (14) holds for a smaller range of s), was proved by Komlós and Füredi [**43**]. A later result of Soshnikov [**92**] improves the range of s under which (14) holds to s of order less than $n^{\frac{2}{3}}$, a result that captures the fluctuations of $\lambda_{\max}(X^N)$. We emphasize here that the proof below heavily depends on the assumption that the distribution of the entries is symmetric.

Proof. Let us first derive the convergence in probability from the moment estimates. First, note that

$$P(\lambda_{\max}(X^N) \leq 2 - \epsilon) \leq P(\int f(x) dL_{X^N} = 0)$$

for all functions f supported on $]2-\epsilon, \infty[$. Taking f bounded continuous. null on $]-\infty, 2-\epsilon]$ and strictly positive in $[2-\epsilon/2, 2]$, we see that $P(\int f(x) dL_{X^N} = 0)$ goes to zero by Theorem 1.10. For the upper bound on $\lambda_{\max}(X^N)$, we shall use Chebychev's inequality and the moment estimates (14) as follows.

$$\begin{aligned} P(\lambda_{\max}(X^N) \geq 2 + \epsilon) &\leq \frac{1}{(2+\epsilon)^{2s}} \mathbb{E}[\lambda_{\max}(X^N)^{2s}] \leq \frac{1}{(2+\epsilon)^{2s}} \mathbb{E}[\mathrm{Tr}((X^N)^{2s})] \\ &\leq \frac{N 2^{2s}}{(2+\epsilon)^{2s}\sqrt{\pi s^3}}(1 + o(s,N)) \end{aligned}$$

where the right hand side goes to zero with N when $s = N^\epsilon$ for some $\epsilon > 0$.

The proof of (14) is based on the expansion of the moments as in the proof of Theorem 1.1 and a good control on the graphs given by the indices that contribute to the resulting sum. The main point is to show that when s is much smaller than $\sqrt{N}$, these graphs are still trees. The interested reader can find the proof in the original article or in [**49**]. □

2
Gaussian Wigner matrices and Fredholm determinants

In this lecture, we shall consider the case where the entries of the matrix $\mathbf{X}^{N,\beta}$ are real or complex Gaussian variables. Moreover, since the results will depend upon the fact that the entries are real or complex, we now make the difference in the notations. We consider $N \times N$ self-adjoint random matrices with entries

$$X_{kl}^{N,\beta} = \frac{g_{kl} + i(\beta-1)\tilde{g}_{kl}}{\sqrt{\beta}} \quad 1 \le k < l \le N, \quad X_{kk}^{N,\beta} = \sqrt{\frac{2}{\beta}} g_{kk}, \quad 1 \le k \le N$$

where the $(g_{kl}, \tilde{g}_{kl}, k \le l)$ are independent equidistributed centered Gaussian variables with variance 1. $(\mathbf{X}^{N,2}, N \in \mathbb{N})$ is commonly referred to as the Gaussian Unitary Ensemble (**GUE**) and $(\mathbf{X}^{N,1}, N \in \mathbb{N})$ as the Gaussian Orthogonal Ensemble (**GOE**) since they can be characterized by the fact that their laws are invariant under the action of the unitary and orthogonal group respectively (see [**79**]). We denote $P_N^{(\beta)}$ the law of $\mathbf{X}^{N,\beta}$.

The goal of this lecture will be to show that
- the law of the eigenvalues of the (**GUE**) is a determinantal law,
- the eigenvalues statistics are described by a Fredholm determinant,
- this description permits to derive the asymptotics of local statistics (see (2) and (3)).

Note here that the eigenvalues are not normalized and so the previous lecture implies that $\frac{1}{N}\sum_{i=1}^N \delta_{\frac{\lambda_i(\mathbf{X}^{N,\beta})}{\sqrt{N}}}$ converges as N goes to infinity (here $(\lambda_i(\mathbf{X}^{N,\beta}), 1 \le i \le N)$ denotes the eigenvalues of the matrix $\mathbf{X}^{N,\beta}$). We denote $P_N^{(\beta)}$ the law of $\mathbf{X}^{N,\beta}$. $\mathcal{H}_N^{(\beta)}$ denotes the set of symmetric (resp. Hermitian) matrices when $\beta = 1$ (resp. $\beta = 2$).

The content of this lecture is borrowed from a book in progress with G. Anderson and O. Zeitouni [**3**]. We shall only sketch here the arguments and refer the interested reader to this book for details.

1. Joint law of the eigenvalues

Lemma 2.1. *Let $\mathbf{X} \in \mathcal{H}_N^{(\beta)}$ be random with law $P_N^{(\beta)}$. The joint distribution of the eigenvalues $\lambda_1(X) \le \cdots \le \lambda_N(X)$, has density with respect to Lebesgue measure proportional to*

$$1_{x_1 \le \cdots \le x_N} \prod_{1 \le i < j \le N} |x_i - x_j|^\beta \prod_{i=1}^N e^{-\beta x_i^2/4}. \tag{15}$$

We denote $\mathcal{P}_N^{(\beta)}$ the unordered law of the eigenvalues;

$$\int f(\{x_i\}_{1\le i \le N}) d\mathcal{P}_N^{(\beta)}(x_1, \cdots, x_N) = \frac{1}{N!}\sum_{\sigma \in S_N} \int f(\{\lambda_{\sigma(i)}(X)\}_{1 \le i \le N}) dP_N^{(\beta)}(X)$$

with S_N the set of permutations of $\{1, \cdots, N\}$.

We shall not prove in details this lemma here but emphasize the ideas of a proof in the case $\beta = 1$. It is simply to write the decomposition $X = UDU^*$, with the eigenvalues matrix D that is diagonal and with real entries, and with eigenvectors matrix U (that is a unitary matrix). Suppose this map was a bijection (which it is not, at least at the matrices that do not possess all distinct eigenvalues) and that one can parametrize the orthonormal basis of eigenvectors by $\beta N(N-1)/2$ parameters in a smooth way (which one cannot in general). Then, it is easy to deduce from the formula $X = UDU^*$ that the Jacobian of this change of variables will depend polynomially on the entries of D and will be of degree $\beta N(N-1)/2$ in these variables. Since the bijection must break down when $D_{ii} = D_{jj}$ for some $i \ne j$, the Jacobian must vanish on that set. When $\beta = 1$, this imposes that the polynomial must be proportional to $\prod_{1 \le i<j\le N}(x_i - x_j)$. Further degree and symmetry considerations allow to generalize this to $\beta = 2$. We refer the reader to [**3**] for a full proof, which shows that the set of matrices for which the above manipulations are not permitted has Lebesgue measure zero.

2. Joint law of the eigenvalues and determinantal law

We now restrict our attention to the case $\beta = 2$ and show that the law $\mathcal{P}_N^{(2)}$ is a determinantal law. More precisely, we let $\mathcal{P}_{p,N}^{(2)}$ be the distribution of p unordered eigenvalues of the GUE;$\mathcal{P}_{p,N}^{(2)}$ is the probability measure on $\mathbb{R}^p$ so that for any $f \in C_b(\mathbb{R}^p)$,

$$\int f(\theta_1, \cdots, \theta_p) d\mathcal{P}_{p,N}^{(2)}(\theta_1, \cdots, \theta_p) = \int f(\theta_1, \cdots, \theta_p) d\mathcal{P}_N^{(2)}(\theta_1, \cdots, \theta_N)$$

2.8. Hermite polynomials

We now introduce the Hermite polynomials and associated normalized harmonic oscillator wave-function.

Definition 2.2.

a) The n^{th} Hermite polynomial $\mathrm{h}_n(x)$ is defined as

$$\mathrm{h}_n(x) := (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}.$$

b) The n^{th} normalized harmonic oscillator wave-function is the function

$$\psi_n(x) = \frac{e^{-x^2/4}\mathrm{h}_n(x)}{\sqrt{\sqrt{2\pi}\, n!}}.$$

For our needs, the most important property of the harmonic oscillator wave-function are their orthogonality relations

$$\int \psi_k(x)\psi_\ell(x) dx = \delta_{k\ell} \tag{16}$$

that we leave as an exercise.

2.9. Determinantal structure

We are finally ready to describe the determinantal structure of $\mathcal{P}_{p,N}^{(2)}$.

Lemma 2.3. *For any $p \leq N$, the law $\mathcal{P}_{p,N}^{(2)}$ is absolutely continuous with respect to Lebesgue measure, with density*

$$\rho_{p,N}^{(2)}(\theta_1, \cdots, \theta_p) = \frac{(N-p)!}{N!} \det_{k,l=1}^{p} K^{(N)}(\theta_k, \theta_l),$$

where

$$K^{(N)}(x, y) = \sum_{k=0}^{N-1} \psi_k(x)\psi_k(y). \tag{17}$$

Proof. Lemma 2.1 shows that $\rho_{p,N}^{(2)}$ exists and equals, if $x_i = \theta_i$ for $i \leq p$ and ζ_i for $i > p$, to

$$\rho_{p,N}^{(2)}(\theta_1, \cdots, \theta_p) = C_{N,p} \int \prod_{1 \leq i < j \leq N} |x_i - x_j|^2 \prod_{i=1}^{N} e^{-x_i^2/2} \prod_{i=p+1}^{N} d\zeta_i \tag{18}$$

for some constant $C_{N,p}$. The fundamental remark is that this density depends on the Vandermonde determinant

$$\prod_{1 \leq i < j \leq N} (x_j - x_i) = \det_{i,j=1}^{N} x_i^{j-1} = \det_{i,j=1}^{N} \mathrm{h}_{j-1}(x_i) \tag{19}$$

where we used in the last equality that the Hermite polynomials are monic.

We first consider the case $p = N$. Then,

$$\begin{aligned} \rho_{N,N}^{(2)}(\theta_1, \cdots, \theta_N) &= C_{N,N} \left(\det_{i,j=1}^{N} \mathrm{h}_{j-1}(\theta_i) \right)^2 \prod_{i=1}^{N} e^{-\theta_i^2/2} \\ &= \tilde{C}_{N,N} \left(\det_{i,j=1}^{N} \psi_{j-1}(\theta_i) \right)^2 = \tilde{C}_{N,N} \det_{i,j=1}^{N} K^{(N)}(\theta_i, \theta_j), \end{aligned} \tag{20}$$

where in the last equality we used the formula $\det(AB) = \det(A)\det(B)$. Here,

$$\tilde{C}_{N,N} = \prod_{k=0}^{N-1} (\sqrt{2\pi} k!) C_{N,N}$$

is given by the inverse of

$$\begin{aligned} \int \left(\det_{i,j=1}^{N} \psi_{j-1}(\theta_i) \right)^2 \prod d\theta_i &= \sum_{\sigma,\sigma'} \epsilon(\sigma)\epsilon(\sigma') \prod_{i=1}^{N} \int \psi_{\sigma(i)-1}(\theta_i)\psi_{\sigma'(i)-1}(\theta_i) d\theta_i \\ &= \sum_{\sigma,\sigma'} 1_{\sigma=\sigma'} = N! \end{aligned}$$

For $p < N$, using (18) and (19), we find that for some constant $\tilde{C}_{N,p}$, with $x_i = \theta_i$ if $i \le p$ and ζ_i otherwise,

$$\rho^{(2)}_{p,N}(\theta_1, \cdots, \theta_p) = \tilde{C}_{N,p} \int (\det_{i,j=1}^{N} \psi_{j-1}(x_i))^2 \prod_{i=p+1}^{N} d\zeta_i$$

$$= \tilde{C}_{N,p} \sum_{\sigma,\tau \in S_N} \varepsilon(\sigma)\varepsilon(\tau) \int \prod_{j=1}^{N} \psi_{\sigma(j)-1}(x_j)\psi_{\tau(j)-1}(x_j) \prod_{i=p+1}^{N} d\zeta_i \, .$$

Therefore, letting $\mathcal{S}(p, \nu)$ denote those bijections τ, σ of $\{1, \cdots, p\}$ into $\{\nu_1, \cdots, \nu_p\}$, we get

$$\begin{aligned} &\rho^{(2)}_{p,N}(\theta_1, \cdots, \theta_p) \\ &= \tilde{C}_{N,p} \sum_{1 \le \nu_1 < \ldots < \nu_p \le N} \sum_{\sigma,\tau \in \mathcal{S}(p,\nu)} \varepsilon(\sigma)\varepsilon(\tau) \prod_{i=1}^{p} \psi_{\sigma(i)-1}(\theta_i)\psi_{\tau(i)-1}(\theta_i) \\ &= \tilde{C}_{N,p} \sum_{1 \le \nu_1 < \ldots < \nu_p \le N} \left(\det_{i,j=1}^{p} \psi_{\nu_j - 1}(\theta_i) \right)^2 , \end{aligned} \tag{21}$$

where in the first equality we used the orthogonality of the family $\{\psi_j\}$ to conclude that the contribution comes only from permutations of $\mathcal{S}_N$ so that $\tau(i) = \sigma(i)$ for $i > p$, and we put $\{\nu_1, \cdots, \nu_p\} = \{\tau(1), \cdots, \tau(p)\} = \{\sigma(1), \cdots, \sigma(p)\}$.

We next need the following generalization of the formula $\det(AB) = \det(A)\det(B)$ given by the Cauchy-Binet Theorem;

Theorem 2.4 (Cauchy-Binet Theorem)**.** *Suppose A is an m by k matrix, B a k by n matrix, $C = AB$, and, with $r \le \min\{m, k, n\}$, set $I = \{i_1, \ldots, i_r\} \subset \{1, \ldots, m\}$, $J = \{j_1, \ldots, j_r\} \subset \{1, \ldots, n\}$. Then, letting $\mathcal{K}_{r,k}$ denote all subsets of $\{1, \ldots, k\}$ of cardinality r,*

$$\det C_{I,J} = \sum_{K \in \mathcal{K}_{r,k}} \det A_{I,K} \det B_{K,J} \, . \tag{22}$$

Using this theorem with $A = B^*$ and $A_{i,j} = \psi_{\nu_j - 1}(\theta_i)$, we get from (21) that

$$\rho^{(2)}_{p,N}(\theta_1, \cdots, \theta_p) = \tilde{C}_{N,p} \det_{i,j=1}^{p} (K^{(N)}(\theta_i, \theta_j)).$$

To compute $\tilde{C}_{N,p}$, note that by integrating both sides of (21), we obtain

$$1 = \tilde{C}_{N,p} \sum_{1 \le \nu_1 < \ldots < \nu_p \le N} \int \left(\det_{i,j=1}^{p} \psi_{\nu_j - 1}(\theta_i) \right)^2 d\theta_1 \cdots d\theta_p = \tilde{C}_{N,p} \sum_{1 \le \nu_1 < \ldots < \nu_p \le N} p! \tag{23}$$

so that $\tilde{C}_{N,p} = (N-p)!/N!$. □

3. Determinantal structure and Fredholm determinants

Now we arrive at the main point, on which the study of the local properties of the GUE is based.

Lemma 2.5. *For any measurable subset A of $\mathbb{R}$,*

$$P_N^{(2)}(\cap_{i=1}^N\{\lambda_i \in A\}) = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{A^c} \cdots \int_{A^c} \det_{i,j=1}^{k} K^{(N)}(x_i, x_j) \prod_{i=1}^{k} dx_i. \tag{24}$$

The last expression appearing in (24) is a *Fredholm determinant.*

Proof. By using in the first equality Lemmas 2.3, and the orthogonality relations of the harmonic functions in the second equality, we have

$$\begin{aligned}
P[\lambda_i \in A, i = 1, \ldots, N] &= \frac{1}{N!} \int_A \cdots \int_A (\det_{i,j=0}^{N-1} \psi_i(x_j))^2 \prod dx_i \\
&= \frac{1}{N!} \sum_{\sigma \in S_N} \epsilon(\sigma) \int_A \cdots \int_A \det_{i,j=0}^{N-1} (\psi_i(x_j)) \prod \psi_i(x_{\sigma(i)}) dx_i \\
&= \frac{1}{N!} \sum_{\sigma \in S_N} \int_A \cdots \int_A \det_{i,j=0}^{N-1} (\psi_i(x_{\sigma(j)})) \prod \psi_i(x_{\sigma(i)}) dx_i \\
&= \int_A \cdots \int_A \det_{i,j=0}^{N-1} (\psi_i(x_j))) \prod_{i=1}^{n} \psi_i(x_i) \prod dx_i \\
&= \det_{i,j=0}^{N-1} \int_A \psi_i(x)\psi_j(x) dx = \det_{i,j=0}^{N-1} \left(\delta_{ij} - \int_{A^c} \psi_i(x)\psi_j(x) dx\right) \\
&= 1 + \sum_{k=1}^{N} (-1)^k \sum_{0 \le \nu_1 < \cdots < \nu_k \le N-1} \det_{i,j=1}^{k} \left(\int_{A^c} \psi_{\nu_i}(x)\psi_{\nu_j}(x) dx\right),
\end{aligned}$$

Therefore,

$$\begin{aligned}
&P[\lambda_i \in A, i = 1, \ldots, N] \qquad (25)\\
&= 1 + \sum_{k=1}^{N} \frac{(-1)^k}{k!} \int_{A^c} \cdots \int_{A^c} \sum_{0 \le \nu_1 < \ldots < \nu_k \le N-1} \left(\det_{i,j=1}^{k} \psi_{\nu_i}(x_j)\right)^2 \prod_{i=1}^{k} dx_i \\
&= 1 + \sum_{k=1}^{N} \frac{(-1)^k}{k!} \int_{A^c} \cdots \int_{A^c} \det_{i,j=1}^{k} K^{(N)}(x_i, x_j) \prod_{i=1}^{k} dx_i \\
&= 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{A^c} \cdots \int_{A^c} \det_{i,j=1}^{k} K^{(N)}(x_i, x_j) \prod_{i=1}^{k} dx_i \,, \qquad (26)
\end{aligned}$$

where we used the Cauchy-Binet Theorem 2.4 and the last step is trivial since the determinant $\det_{i,j=1}^k K^{(N)}(x_i, x_j)$ has to vanish identically for $k > N$ because the rank of $\{K^{(N)}(x_i, x_j)\}_{i,j=1}^k$ is at most N (for instance because $\{K^{(N)}(x_i, x_j)\}_{i,j=1}^k$ can be seen as the product of two $N \times k$ matrices). □

4. Fredholm determinant and asymptotics

Let us denote, for a Borel set A and a symmetric function K on $\mathbb{R} \times \mathbb{R}$, the Fredholm determinant

$$\Delta(A, K) := 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_A \cdots \int_A \det_{i,j=1}^{k} K(x_i, x_j) \prod_{i=1}^{k} dx_i$$

Then, we claim (and leave as an exercise, see below) that

Lemma 2.6. *For any compact set A, if K_n is a sequence that converges uniformly towards K on A as n goes to infinity, $\Delta(A, K_n)$ converges towards $\Delta(A, K)$.*

As a consequence, if we take $A = N^{-\frac{1}{2}}B$ with a compact set B, we see that the spacing distribution in the bulk announced in (2) is a consequence of the (uniform on compact) convergence

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} K^{(N)}(\frac{x}{\sqrt{N}}, \frac{y}{\sqrt{N}}) = \frac{\sin(x-y)}{\pi(x-y)}. \tag{27}$$

Similarly, if we take $A = 2 + N^{-\frac{2}{3}}[t, t']$, we find that the probability (3) that there is no eigenvalue in A can be obtained from the asymptotics

$$\lim_{N\to\infty} \frac{1}{N^{\frac{1}{6}}} K^{(N)}(2\sqrt{N} + \frac{x}{N^{\frac{1}{6}}}, 2\sqrt{N} + \frac{y}{N^{\frac{1}{6}}}) = \frac{Ai(x)Ai'(y) - Ai'(x)Ai(y)}{x-y}$$

if Ai denotes the Airy function.

Such asymptotics are obtained thanks to the formula (left as an exercise)

$$K^{(N)}(x, y) = \sqrt{N} \frac{\psi_N(x)\psi_{N-1}(y) - \psi_{N-1}(x)\psi_N(y)}{x-y} \tag{28}$$

and the associated asymptotics of the harmonic functions ψ_N and ψ_{N-1}. We propose below to derive the asymptotics (27) as an exercise. The limit at the edge and the Airy kernel is a more challenging exercise that requires saddle point analysis that we did not dare to leave as an exercise. The interested reader can find a full treatment in [**3**].

Exercise 2.7. *(Proof of Lemma 2.6) Let A be a compact subset of $\mathbb{R}$ and denote $\|K\|_A = \sup_{(x,y)\in A\times A} |K(x, y)| < \infty$. Let K_i be functions on $A \times A$.*

-Prove that for any $x_i, y_i \in A$, any $n \in \mathbb{N}$,

$$\left| \det_{i,j=1}^{n} K_i(x_i, y_j) \right| \le n^{n/2} \prod_{i=1}^{n} \|K_i\|_A. \tag{29}$$

Hint: use Hadamard inequality: For any column vectors $v_1, \ldots, v_n$ of length n with complex entries, it holds that

$$\det \left[\begin{array}{ccc} v_1 & \ldots & v_n \end{array} \right] \le \prod_{i=1}^{n} \sqrt{\bar{v_i}^T v_i}$$

-Prove that for any $x_i, y_i \in A$,

$$\left| \det_{i,j=1}^{n} K_1(x_i, y_j) - \det_{i,j=1}^{n} K_2(x_i, y_j) \right|$$

$$\le \quad n^{1+n/2} \|K_1 - K_2\|_A \cdot \max(\|K_1\|_A, \|K_2\|_A)^{n-1}$$

-Conclude that $K \to \Delta(A, K)$ is Lipschitz for $\|\cdot\|_A$ on the set of functions $\{K : \|K\|_A \le M\}$.

Exercise 2.8. *(Proof of (28)) Prove that*

$$\sum_{k=0}^{N-1} \frac{1}{k!} \mathrm{h}_k(x)\mathrm{h}_k(y) = \sqrt{N} \frac{\mathrm{h}_N(x)\mathrm{h}_{N-1}(y) - \mathrm{h}_{N-1}(x)\mathrm{h}_N(y)}{(N-1)!(x-y)}$$

Hint - Multiply both sides by $(x-y)F(x, y)$ and integrate with respect to x, y on both sides.

- To prove the equality of the two sides for $F(x, y) = \mathrm{h}_p(x)\mathrm{h}_\ell(y)$, show that $\mathrm{h}_{p+1}(x) = x\mathrm{h}_p(x) - \mathrm{h}'_p(x)$, $\mathrm{h}'_p(x) = p\mathrm{h}_{p-1}(x)$, $\int f\mathrm{h}_k(x)dx = 0$ for all polynomial f of degree $< k$, $\int \mathrm{h}_k(x)\mathrm{h}_p(x)e^{-\frac{x^2}{2}}dx = \sqrt{2\pi}1_{k=p}k!$.

-Conclude.

Exercise 2.9. *(Proof of (27)) It is enough to obtain the asymptotics of $\Psi_\nu(t) := N^{\frac{1}{4}}\psi_\nu(\frac{t}{\sqrt{N}})$ for $\nu - N$ finite. The goal of this exercise is to show that*

$$\lim_{N\to\infty} \Psi_\nu(t) = \frac{1}{\sqrt{\pi}} \cos(t - \frac{\pi\nu}{2}).$$

-Prove that

$$\Psi_\nu(t) \quad = \quad \frac{(2\pi)^{1/4} C_{\nu,N} e^{t^2/(4N)} N^{1/4+\nu/2}}{\sqrt{\nu!}} \int (\xi e^{-\xi^2/2})^N e^{-i\xi t} \xi^{\nu-N} d\xi$$

with $C_{\nu,n} = i^\nu \sqrt{n}/(2\pi)$. *Hint: observe that* $\partial_x^N e^{-\frac{x^2}{2}} = \partial_x^N \int e^{i\xi x} e^{-\frac{\xi^2}{2}} d\xi$.

-Use Laplace method (observe here that the complex part is of order one and so the integral will concentrate on the optimizers of $|\xi| e^{-\xi^2/2}$*).*

Exercise 2.10. *Let* $X^{2k}_{\cdot}$ *be a random walk starting at* $X_0^{2k} = 2k$ *such that* $\Delta X_n^{2k} := X_{n+1}^{2k} - X_n^{2k}$ *are independent equidistributed Bernoulli variables, equal to* $+1$ *with probability* p *and* -1 *with probability* $1-p$. *We let* K_T *be the associated transition kernel (*$K_T(x,y) = P(X_T = y \cap X_0 = y)$*).*

We consider N *random walks* $X^{2k}, 1 \le k \le N$.

Show that for any sequence $x_1 < x_2 \cdots < x_N$,

$$P(\{X_T^{2k} = x_k \forall 1 \le k \le N\} \cap X_n^{2k} < X_n^{2k+2} \forall\, 0 \le n \le T, \forall\, k) = \det(K_T(2k, x_l)_{k,l})$$

Hint: Expand the right hand side has a sum over permutations and paths. Show that intersecting paths have a null contribution (use the fact that paths that cross intersect and that intersecting paths come by pairs, with final data exchanged by a permutation)(i.e., the reflexion principle).

3

Wigner matrices and concentration inequalities

Concentration inequalities came up to be a very powerful tool in probability theory. They provide a general framework to control the probability of deviations of smooth functions of random variables from their mean or their median. We begin this section by providing some general framework where concentration inequalities can be obtained. We first consider the case where the underlying measure satisfies a log-Sobolev inequality; we show how to prove this inequality in a simple context and then how it implies concentration inequalities. We then review a few other contexts where concentration inequalities hold. To apply these techniques to random matrices, we show that certain functions of the eigenvalues of matrices, such as $\int f(x)dL_{\mathbf{X}^N}(x) = \frac{1}{N}\mathrm{Tr}(f(\mathbf{X}^N))$ with f Lipschitz, are smooth functions of the entries of the matrix so that concentration inequalities hold as soon as the joint law of the entries satisfies one of the conditions seen in the first two sections of this lecture. As a consequence, we will see that if the entries of $\mathbf{X}^N$ satisfy a log-Sobolev inequality, $Z_f^N := N(\int f(x)dL_{\mathbf{X}^N}(x) - \mathbb{E}[\int f(x)dL_{\mathbf{X}^N}(x)])$ has a sub-Gaussian tail for all N and for all Lipschitz function f. Another useful *a priori* control is provided by Brascamp-Lieb inequalities; we shall apply them in the context of random matrices at the end of this lecture. The interest of such inequalities is that they provide bounds on probabilities of deviations from the mean that do not depend on the dimension. They can be used to show laws of large numbers (reducing the proof of the almost sure convergence to the prove of the convergence in expectation) or to ease the proof of central limit theorems (since it implies that Z_f^N has a sub-Gaussian tail independent of N, and thus provides tightness arguments for free).

In this section, we denote by $\langle \cdot, \cdot \rangle$ the Euclidean scalar product on $\mathbb{R}^M$ (or $\mathbb{C}^M$), $\langle x, y \rangle = \sum_{i=1}^M x_i y_i$ ($\langle x, y \rangle := \sum_{i=1}^M x_i y_i^*$), and by $||\cdot||_2$ the associated norm $||x||_2^2 := \langle x, x \rangle$.

1. Concentration inequalities and logarithmic Sobolev inequalities

We first derive concentration inequalities based on the logarithmic Sobolev inequality and then give some generic and classical examples of laws that satisfy this inequality.

1.10. Concentration inequalities for laws that satisfy logarithmic Sobolev inequalities

Throughout this section an integer number N will be fixed.

Definition 3.1. *A probability measure P on $\mathbb{R}^N$ is said to satisfy the logarithmic Sobolev inequality (LSI) with constant c if, for any differentiable function $f : \mathbb{R}^N \to \mathbb{R}$,*

$$\int f^2 \log \frac{f^2}{\int f^2 dP} dP \leq 2c \int \|\nabla f\|_2^2 dP. \tag{30}$$

Here, $\|\nabla f\|_2^2 = \sum_{i=1}^N (\partial_{x_i} f)^2$

The interest in the logarithmic Sobolev inequality, in the context of concentration inequalities, lies in the following argument, that among other things, shows that LSI implies sub-Gaussian tails. This fact and a general study of logarithmic Sobolev inequalities may be found in [**54**] or [**74**]. The Gaussian law, and any probability measure ν absolutely continuous with respect to the Lebesgue measure satisfying the Bobkov and Götze [**18**] condition (including $\nu(dx) = Z^{-1}e^{-|x|^\alpha}dx$ for $\alpha \geq 2$, where $Z = \int e^{-|x|^\alpha}dx$), as well as any distribution absolutely continuous with respect to such laws possessing a bounded above and below density, satisfies the LSI [**74**], [**54**, Property 4.6].

Lemma 3.2 (Herbst). *Assume that P satisfies the LSI on $\mathbb{R}^N$ with constant c. Let G be a Lipschitz function on $\mathbb{R}^N$, with Lipschitz constant $|G|_{\mathcal{L}}$. Then, for all $\lambda \in \mathbb{R}$, we have*

$$E_P[e^{\lambda(G - E_P(G))}] \leq e^{c\lambda^2|G|_{\mathcal{L}}^2/2}, \tag{31}$$

and so for all $\delta > 0$

$$P\left(|G - E_P(G)| \geq \delta\right) \leq 2e^{-\delta^2/2c|G|_{\mathcal{L}}^2}. \tag{32}$$

Proof of Lemma 3.2. Note first that (32) follows from (31). Indeed, by Chebychev's inequality, for any $\lambda > 0$,

$$\begin{aligned}
P\left(|G - E_P G| \geq \delta\right) &\leq e^{-\lambda\delta} E_P[e^{\lambda|G - E_P G|}] \\
&\leq e^{-\lambda\delta}(E_P[e^{\lambda(G - E_P G)}] + E_P[e^{-\lambda(G - E_P G)}]) \\
&\leq 2e^{-\lambda\delta} e^{c|G|_{\mathcal{L}}^2\lambda^2/2}.
\end{aligned}$$

Optimizing with respect to λ (by taking $\lambda = \delta/c|G|_{\mathcal{L}}^2$) yields the bound (32).

Turning to the proof of (31), we assume that G is a bounded differentiable function such that

$$|| \, ||\nabla G||_2^2 ||_\infty := \sup_{x \in \mathbb{R}^N} \sum_{i=1}^N (\partial_{x_i} G(x))^2 < \infty.$$

We leave the generalization to the reader (see also [**3**] or [**5**]). Define

$$X_\lambda = \log E_P e^{2\lambda(G - E_P G)}.$$

Then, taking $f = e^{\lambda(G - E_P G)}$ in (30), some algebra reveals that for $\lambda > 0$,

$$\frac{d}{d\lambda}\left(\frac{X_\lambda}{\lambda}\right) \leq 2c|| \, ||\nabla G||_2^2 ||_\infty .$$

Now, because $G - E_P(G)$ is centered,

$$\lim_{\lambda \to 0^+} \frac{X_\lambda}{\lambda} = 0$$

and hence integrating with respect to λ yields $X_\lambda \leq 2c|| \, ||\nabla G||_2^2 ||_\infty \lambda^2$, first for $\lambda \geq 0$ and then for any $\lambda \in \mathbb{R}$ by considering the function $-G$ instead of G. This completes the proof of (31) in case G is bounded and differentiable. □

1.11. A few laws that satisfy a log-Sobolev inequality

In the sequel, we shall be interested in laws of variables that are either independent or in interaction via a potential. We shall give sufficient conditions to ensure that a log-Sobolev inequality is satisfied.

- *Laws of independent variables.*

  One of the most important properties of the log-Sobolev inequality is the product property (we leave again the proof as an exercise)

  Lemma 3.3. *Let $(\mu_i)_{i=1,2}$ be two probability measures on $\mathbb{R}^N$ and $\mathbb{R}^M$, respectively, satisfying the logarithmic Sobolev inequalities with coefficients $(c_i)_{i=1,2}$. Then, the product probability measure $\mu_1 \otimes \mu_2$ on $\mathbb{R}^{M+N}$ satisfies the logarithmic Sobolev inequality with coefficient* $\max(c_1, c_2)$.

  Consequently, if μ is a probability measure on $\mathbb{R}^M$ satisfying a logarithmic Sobolev inequality with a coefficient $c < \infty$, then the product probability measure $\mu^{\otimes n}$ satisfies the logarithmic Sobolev inequality with the same coefficient c for any integer number n.

- *Log-Sobolev inequalities for variables in convex interaction.*

  We follow below [**5**] chapter 5 and [**54**] chapter 4, which we recommend for more details. Let dx denote the Lebesgue measure on $\mathbb{R}$ and Φ be a smooth function (at least twice continuously differentiable) from $\mathbb{R}^N$ into $\mathbb{R}$ going to infinity fast enough so that the probability measure

  $$\mu_\Phi(dx) := \frac{1}{Z} e^{-\Phi(x_1, \cdots, x_N)} dx_1 \cdots dx_N$$

  is well defined. Then, Bakry and Emery showed that if Φ is strictly convex, μ_Φ satisfies a log-Sobolev inequality. Namely,

  Theorem 3.4. *Let I denote the identity in the space of $N \times N$ matrices. If for all $x \in \mathbb{R}^N$,*

  $$\mathrm{Hess}(\Phi)(x) = \big((\partial_{x_i}\partial_{x_j}\Phi)(x)\big)_{1 \leq i,j \leq N} \geq \frac{1}{c} I$$

  in the sense of the partial order on self-adjoint operators, then (BE) is satisfied and μ_Φ satisfies the logarithmic Sobolev inequality with constant c.

  In particular, if μ is the law of N independent Gaussian variables with covariance bounded above by c, then μ satisfies the logarithmic Sobolev inequality with constant c.

1.12. Local concentration inequalities

In many instances we shall encounter later, we will need to control the concentration of functions that are only locally Lipschitz, for instance polynomial functions. To this end we state (and prove) the following lemma. Let (X,d) be a metric space and set for $f : X \to \mathbb{R}$

$$|f|_{\mathcal{L}} := \sup_{x,y\in X} \frac{|f(x)-f(y)|}{d(x,y)}.$$

Denote, for a subset B of X, $d(x,B) = \inf_{y\in B} d(x,y)$. Then

Lemma 3.5. *Assume that a probability measure μ on (X,d) satisfies a concentration inequality; for all $\delta > 0$, for all $f : X \to \mathbb{R}$,*

$$\mu(|f-\mu(f)| \geq \delta) \leq e^{-g(\frac{\delta}{|f|_{\mathcal{L}}})}$$

for some increasing function g on $\mathbb{R}^+$. Let B be a subset of X and let $f : B \to \mathbb{R}$ such that

$$|f|^B_{\mathcal{L}} := \sup_{x,y\in B} \frac{|f(x)-f(y)|}{d(x,y)}$$

is finite. Then, with $\delta(f) := \mu\left(1_{B^c}(\sup_{x\in B}|f(x)| + |f|^B_{\mathcal{L}} d(x,B))\right)$, we have

$$\mu(\{|f-\mu(f1_B)| \geq \delta + \delta(f)\} \cap B) \leq e^{-g(\frac{\delta}{|f|^B_{\mathcal{L}}})}$$

Proof. It is enough to define a Lipschitz function $\tilde{f}$ on X, whose Lipschitz constant $|\tilde{f}|_{\mathcal{L}}$ is bounded above by $|f|^B_{\mathcal{L}}$ and so that $\tilde{f} = f$ on B. We just set

$$\tilde{f}(x) = \sup_{y\in B}\{f(y) - |f|^B_{\mathcal{L}} d(x,y)\}.$$

Note that, if $x \in B$, since $f(y) - f(x) - |f|^B_{\mathcal{L}} d(x,y) \leq 0$, the above supremum is taken at $y = x$ and $\tilde{f}(x) = f(x)$. Applying the concentration inequality to $\tilde{f}$ yields the result (the constant $\delta(f)$ accounts for the centering with respect to $\mu(\tilde{f})$ rather than $\mu(f1_B)$). ☐

Exercise 3.6. *The goal of this exercise is to obtain a concentration of measure inequality under a measure μ on $\mathbb{R}^N$ that satisfies the spectral gap inequality*

$$\mu(f^2) - (\mu(f))^2 \leq \frac{1}{m}\mu(\|\nabla f\|_2^2)$$

for all differentiable function f.

(1) *Take $u(t) = \mu(e^{t(f-\mu(f))})$ with a bounded differentiable function f. Show that*

$$u(2t) \leq u(t)^2 + \frac{t^2}{m}\|\|\nabla f\|_2^2\|_\infty u(2t).$$

Conclude that for $t^2 = m/2\|\|\nabla f\|_2^2\|_\infty$,

$$u(2t) \leq 2u(t)^2.$$

Iterating, deduce that for this same t,

$$u(2t) \leq \prod_{i=0}^{\infty}(1 - \frac{1}{2}\frac{1}{4^i})^{-1} := K$$

(2) *Deduce that*

$$\mu(|f-\mu(f)| \geq \delta) \leq 2Ke^{-\frac{m\delta}{\|\|\nabla f\|_2\|_\infty}}.$$

Show that in particular μ must have sub-exponential tail.

(3) *Show that if μ satisfies a LSI with constant c then μ satisfies a spectral gap inequality with constant $m = 1/c$. Hint: take $f = 1 + \epsilon g$ in the LSI with ϵ going to zero.*

Exercise 3.7. *The goal of this exercise is to obtain a concentration of measure inequality under a measure μ on $\mathbb{R}^N$ that has only a second moment. We take P a measure on $\mathbb{R}^n$ such that $L := \max_i \int (x_i - \int x_i dP)^2 dP$. Show that for any Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$,*

$$\mathbb{E}[(f-\mathbb{E}(f))^2] \leq L\sum_{i=1}^{n}\|\partial_i f\|_\infty^2.$$

Apply this result with $f(x_1,\cdots,x_n)=\frac{1}{n}\sum_{i=1}^n f(x_i)$ *for* f *Lipschitz.*

Hint: Write the martingale decomposition $f-\mathbb{E}[f]=\sum_{i=1}^n(E[f|\mathcal{F}_i]-E[f|\mathcal{F}_{i-1}])$ *with* $\mathcal{F}_i$ *the* σ *algebra generated by* $(x_1,\cdots,x_i)$ *and observe that* $\mathbb{E}[(E[f|\mathcal{F}_i]-E[f|\mathcal{F}_{i-1}])^2]\le L\|\partial_i f\|_\infty^2$.

2. Smoothness and convexity of the eigenvalues of a matrix and of traces of matrices

We shall not follow [**55**] where smoothness and convexity were mainly proved by hand for smooth functions of the empirical measure and for the largest eigenvalue. We will rather, as in [**3**], rely on Weyl and Lidskii inequalities that we now recall. We recall that we will denote, for $\mathbf{B}\in\mathcal{M}_N(\mathbb{C})$, $\|\mathbf{B}\|_2$ its Euclidean norm;

$$\|\mathbf{B}\|_2 := \left(\sum_{i,j=1}^N |B_{ij}|^2\right)^{\frac{1}{2}}.$$

Theorem 3.8 (Lidskii)**.** *Let* $A\in\mathcal{H}_N^{(2)}$ *and* $z\in\mathbb{C}^N$. *We order the eigenvalues of* $A\pm zz^*$ *in increasing order. Then*

$$\lambda_k(A\pm zz^*)\le\lambda_{k+1}(A)\le\lambda_{k+2}(A\pm zz^*).$$

Theorem 3.9 (Weyl)**.** *Let* $A,E\in\mathcal{H}_N^{(2)}$. *Then,*

$$\sum_{k=1}^N|\lambda_k(A+E)-\lambda_k(A)|^2\le\sum_{k=1}^N\lambda_k(E)^2\,. \tag{33}$$

We denote $\lambda_1(\mathbf{A})\le\lambda_2(\mathbf{A})\le\cdots\le\lambda_N(\mathbf{A})$ *the eigenvalues of* $\mathbf{A}\in\mathcal{H}_N^{(2)}$. *Then for all* $k\in\{1,\cdots,N\}$,

$$|\lambda_k(A+E)-\lambda_k(A)|\le\|E\|_2.$$

In other words, for all $k\in\{1,\cdots,N\}$,

$$(A_{ij})_{1\le i\le j\le N}\in\mathbb{C}^{N(N+1)/2}\to\lambda_k(\mathbf{A})$$

is Lipschitz with constant one. The same holds for the spectral radius $\lambda_{\max}(\mathbf{A})=\max_{1\le i\le N}|\lambda_i(\mathbf{A})|$.

From Theorem 3.9, we deduce the following.

Lemma 3.10. *For all Lipschitz functions* f *with Lipschitz constant* $|f|_{\mathcal{L}}$, *the function*

$$(A_{ij})_{1\le i\le j\le N}\in\mathbb{C}^{N(N+1)/2}\to\sum_{k=1}^N f(\lambda_k(\mathbf{A}))$$

is Lipschitz with respect to the Euclidean norm with a constant bounded above by $\sqrt{N}|f|_{\mathcal{L}}$. *When* f *is continuously differentiable we have*

$$\lim_{\epsilon\to0}\epsilon^{-1}\left(\sum_{k=1}^N f(\lambda_k(\mathbf{A}+\epsilon\mathbf{B}))-\sum_{k=1}^N f(\lambda_k(\mathbf{A}))\right)=\mathrm{Tr}(f'(\mathbf{A})\mathbf{B}).$$

Proof. The first inequality is a direct consequence of Theorem 3.9 and entails the same control on $\lambda_{\max}(\mathbf{A})$. For the second we only need to use Cauchy-Schwarz's inequality;

$$\begin{aligned}\left|\sum_{i=1}^N f(\lambda_i(\mathbf{A}))-\sum_{i=1}^N f(\lambda_i(\mathbf{A}+\mathbf{B}))\right| &\le |f|_{\mathcal{L}}\sum_{i=1}^N|\lambda_i(\mathbf{A})-\lambda_i(\mathbf{A}+\mathbf{B})|\\ &\le \sqrt{N}|f|_{\mathcal{L}}\left(\sum_{i=1}^N|\lambda_i(\mathbf{A})-\lambda_i(\mathbf{A}+\mathbf{B})|^2\right)^{\frac{1}{2}}\\ &\le \sqrt{N}|f|_{\mathcal{L}}\|\mathbf{B}\|_2\end{aligned}$$

where we used Theorem 3.9 in the last line. For the last point, we check it for $f(x)=x^k$ where the result is clear since

$$\mathrm{Tr}((\mathbf{A}+\epsilon\mathbf{B})^k)=\mathrm{Tr}(\mathbf{A}^k)+\epsilon k\mathrm{Tr}(\mathbf{A}^{k-1}\mathbf{B})+O(\epsilon^2) \tag{34}$$

and complete the argument by density of the polynomials. □

We can think of $\sum_{i=1}^N f(\lambda_i(\mathbf{A}))$ as $\mathrm{Tr}(f(\mathbf{A}))$. Then, the second part of the previous Lemma can be extended to several matrices as follows.

Lemma 3.11. *Let P be a polynomial in m-non commutative indeterminates. For $1 \le i \le m$, we denote D_i the cyclic derivative with respect to the i^{th} variable given, if P is a monomial, by*

$$D_i P(X_1, \cdots, X_m) = \sum_{P = P_1 X_i P_2} P_2(X_1, \cdots, X_m) P_1(X_1, \cdots, X_m)$$

where the sum runs over all decompositions of P into $P_1 X_i P_2$ for some monomials P_1 and P_2. D_i is extended linearly to polynomials. Then, for all $(\mathbf{A}_1, \cdots, \mathbf{A}_m)$ and $(\mathbf{B}_1, \cdots, \mathbf{B}_m) \in \mathcal{H}_N^{(2)}$,

$$\begin{aligned}\lim_{\epsilon \to 0} \epsilon^{-1} \left(\mathrm{Tr}(P(\mathbf{A}_1 + \epsilon \mathbf{B}_1, \cdots, \mathbf{A}_m + \epsilon \mathbf{B}_m)) - \mathrm{Tr}(P(\mathbf{A}_1, \cdots, \mathbf{A}_m)) \right) \\ = \sum_{i=1}^m \mathrm{Tr}(D_i P(\mathbf{A}_1, \cdots, \mathbf{A}_m) \mathbf{B}_i).\end{aligned}$$

In particular, if $(\mathbf{A}_1, \cdots, \mathbf{A}_m)$ belong to the subset Λ_M^N of elements of $\mathcal{H}_N^{(2)}$ with spectral radius bounded by $M < \infty$,

$$((A_k)ij)_{\substack{1 \le i \le j \le N \\ 1 \le k \le m}} \in \mathbb{C}^{N(N+1)m/2}, \mathbf{A}_k \in \mathcal{H}_N^{(2)} \cap \Lambda_M^N \to \mathrm{Tr}(P(\mathbf{A}_1, \cdots, \mathbf{A}_m))$$

is Lipschitz with a Lipschitz norm bounded by $\sqrt{N} C(P, M)$ for a constant $C(P, M)$ that depends only on M and P. If P is a monomial of degree d, one can take $C(P, M) = d M^{d-1}$.

Proof. We can assume without loss of generality that P is a monomial. The first equality is due to the simple expansion

$$\begin{aligned}&\mathrm{Tr}(P(\mathbf{A}_1 + \epsilon \mathbf{B}_1, \cdots, \mathbf{A}_m + \epsilon \mathbf{B}_m)) - \mathrm{Tr}(P(\mathbf{A}_1, \cdots, \mathbf{A}_m)) \\ &= \epsilon \sum_{i=1}^m \sum_{P = P_1 X_i P_2} \mathrm{Tr}(P_1(\mathbf{A}_1, \cdots, \mathbf{A}_m) \mathbf{B}_i P_2(\mathbf{A}_1, \cdots, \mathbf{A}_m)) + O(\epsilon^2)\end{aligned}$$

together with the trace property $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$.

For the estimate on the Lipschitz norm, observe that if P is a monomial containing d_i times the letter X_i, $\sum_{i=1}^m d_i = d$ and $D_i P$ is the sum of exactly d_i monomials of degree $d-1$. Hence, $D_i P(\mathbf{A}_1, \cdots, \mathbf{A}_m)$ has spectral radius bounded by $d_i M^{d-1}$ when $(\mathbf{A}_1, \cdots, \mathbf{A}_m)$ are Hermitian matrices in Λ_M^N. Hence, by Cauchy-Schwarz's inequality, we obtain

$$\begin{aligned}|\sum_{i=1}^m \mathrm{Tr}(D_i P(\mathbf{A}_1, \cdots, \mathbf{A}_m) \mathbf{B}_i)| &\le \left(\sum_{i=1}^m \mathrm{Tr}(|D_i P(\mathbf{A}_1, \cdots, \mathbf{A}_m)|^2) \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \mathrm{Tr}(\mathbf{B}_i^2) \right)^{\frac{1}{2}} \\ &\le \left(N \sum_{i=1}^m d_i^2 M^{2(d-1)} \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|\mathbf{B}_i\|_2^2 \right)^{\frac{1}{2}} \\ &\le \sqrt{N} d M^{d-1} \left(\sum_{i=1}^m \|\mathbf{B}_i\|_2^2 \right)^{\frac{1}{2}}.\end{aligned}$$

□

Exercise 3.12. *Prove that when $m = 1$, $D_1 P(x) = P'(x)$.*

We now prove the following result originally due to Klein

Lemma 3.13 (Klein's lemma)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function. Then, if $\mathbf{A}$ is the $N \times N$ Hermitian matrix with entries $(A_{ij})_{1 \le i \le j \le N}$ on and above the diagonal,*

$$\psi_f : (A_{ij})_{1 \le i \le j \le N} \in \mathbb{C}^N \to \sum_{i=1}^N f(\lambda_i(\mathbf{A}))$$

is convex. Moreover, if f is twice continuously differentiable with $f''(x) \ge c$ for all x, ψ_f is twice continuously differentiable with Hessian bounded below by cI.

Proof. We give a proof below, that also provides a lower bound of the Hessian of ψ_f. The smoothness of ψ_f is clear when f is a polynomial since then $\psi_f((A_{ij})_{1\le i\le j\le N})$ is a polynomial function in the entries. Let us compute its second derivative when $f(x)=x^p$. Expanding (34) one step further gives

$$\begin{aligned}\mathrm{Tr}((\mathbf{A}+\epsilon\mathbf{B})^k) &= \mathrm{Tr}(\mathbf{A}^k)+\epsilon\sum_{k=0}^{p-1}\mathrm{Tr}(\mathbf{A}^k\mathbf{B}\mathbf{A}^{p-1-k})\\ &\quad+\epsilon^2\sum_{0\le k+l\le p-2}\mathrm{Tr}(\mathbf{A}^k\mathbf{B}A^l\mathbf{B}A^{p-2-k-l})+O(\epsilon^3)\\ &= \mathrm{Tr}(A^k)+\epsilon p\mathrm{Tr}(A^{p-1}\mathbf{B})+\frac{\epsilon^2}{2}p\sum_{0\le l\le p-2}\mathrm{Tr}(A^l\mathbf{B}A^{p-2-l}\mathbf{B})+O(\epsilon^3).\end{aligned}\tag{35}$$

A compact way to write this formula is by defining, for two real numbers x,y,

$$g_f(x,y):=\frac{f'(x)-f'(y)}{x-y}$$

and setting for a matrix $\mathbf{A}$ with eigenvalues $\lambda_i(A)$ and eigenvector e_i, $1\le i\le N$,

$$g_f(\mathbf{A},\mathbf{A})=\sum_{i,j=1}^N g_f(\lambda_i(\mathbf{A}),\lambda_j(\mathbf{A}))e_ie_i^*\otimes e_je_j^*.$$

Since $g_{x^p}(x,y)=p\sum_{r=0}^{p-1}x^ry^{p-1-r}$, the last term in the r.h.s. of (35) reads

$$p\sum_{0\le l\le p-1}\mathrm{Tr}(\mathbf{A}^l\mathbf{B}\mathbf{A}^{p-2-l}\mathbf{B})=\langle g_{x^p}(\mathbf{A},\mathbf{A}),\mathbf{B}\otimes\mathbf{B}\rangle\tag{36}$$

where for $\mathbf{B},\mathbf{C},\mathbf{D},\mathbf{E}\in M_N(\mathbb{C})$, $\langle\mathbf{B}\otimes\mathbf{C},\mathbf{D}\otimes\mathbf{E}\rangle:=\langle\mathbf{B},\mathbf{D}\rangle_2\langle\mathbf{C},\mathbf{E}\rangle_2$ with $\langle\mathbf{B},\mathbf{D}\rangle_2=\sum_{i,j=1}^N B_{ij}\bar{D}_{ij}$. In particular, $\langle e_ie_i^*\otimes e_je_j^*,\mathbf{B}\otimes\mathbf{B}\rangle=|<e_i,Be_j>|^2$ with $<u,Bv>=\sum_{i,j=1}^N u_i\bar{v}_jB_{ij}$. By (35) and (36), for any Hermitian matrix $\mathbf{X}$,

$$\begin{aligned}\mathrm{Hess}(\mathrm{Tr}(\mathbf{A}^p))[X,X] &= \langle g_{x^p}(\mathbf{A},\mathbf{A}),X\otimes X\rangle\\ &= \sum_{r,m=1}^N g_{x^p}(\lambda_r(\mathbf{A}),\lambda_m(\mathbf{A}))|<e_r,Xe_m>|^2\end{aligned}$$

Now $g_f(\mathbf{A},\mathbf{A})$ makes sense for any twice continuously differentiable function f and by density of the polynomials in the set of twice continuously differentiable function f, we can conclude that ψ_f is twice continuously differentiable too. Moreover, for any twice continuously differentiable function f,

$$\mathrm{Hess}(\mathrm{Tr}(f(\mathbf{A})))[X,X] = \sum_{r,m=1}^N g_f(\lambda_r(\mathbf{A}),\lambda_m(\mathbf{A}))|<e_r,Xe_m>|^2.$$

Since $g_f\ge c$ when $f''\ge c$ we finally have proved

$$\mathrm{Hess}(\mathrm{Tr}(f(\mathbf{A})))[\mathbf{X},\mathbf{X}] \ge c\mathrm{Tr}(\mathbf{X}\mathbf{X}^*).$$

The proof is thus complete. □

Let us also notice that

Lemma 3.14. *Assume $\lambda_1(\mathbf{A})\le\lambda_2(\mathbf{A})\cdots\le\lambda_N(\mathbf{A})$. The functions*

$$\mathbf{A}\in\mathcal{H}_N^{(2)}\to\lambda_1(\mathbf{A})\text{ and }\mathbf{A}\in\mathcal{H}_N^{(2)}\to\lambda_N(\mathbf{A})$$

are convex. For any norm $\|\cdot\|$ on $\mathcal{M}_N^{(2)}$, $(A_{ij})_{1\le i,j\le N}\to\|\mathbf{A}\|$ is convex.

Proof. The first result is clear since we have already seen that $\lambda_N(\mathbf{A}+\mathbf{B}) \le \lambda_N(\mathbf{A})+\lambda_N(\mathbf{B})$. Since for $\alpha \in \mathbb{R}$, $\lambda_i(\alpha\mathbf{A}) = \alpha\lambda_i(\mathbf{A})$, we conclude that $\mathbf{A} \to \lambda_N(\mathbf{A})$ is convex. The same result holds for λ_1 (by changing the sign $\mathbf{A} \to -\mathbf{A}$). The convexity of $(A_{ij})_{1\le i,j\le N} \to \|\mathbf{A}\|$ is due to the definition of the norm. □

3. Concentration inequalities for random matrices

3.13. Concentration inequalities for the eigenvalues of random matrices

We consider a Hermitian random matrix $\mathbf{A}^N$ whose real or complex entries have joint law μ^N. We can now state the following theorems.

Theorem 3.15. *Suppose there exists $c > 0$ so that either*

• *(H1) there exists a strictly convex twice continuously differentiable function $V : \mathbb{R} \to \mathbb{R}$, $V''(x) \ge \frac{1}{c} > 0$, so that*

$$\mu^N(d\mathbf{A}^N) = Z_N^{-1} e^{-N\mathrm{Tr}(V(\mathbf{A}^N))} d\mathbf{A}^N$$

with $d\mathbf{A}^N = \prod_{1\le i\le j\le N} d\Re(A_{ij}) \prod_{1\le i<j\le N} d\Im(A_{ij})$ for complex entries or $d\mathbf{A}^N = \prod_{1\le i\le j\le N} dA_{ij}$ for real entries.

• *(H2) $\mathbf{A}^N = \mathbf{X}^N/\sqrt{N}$ with $(\mathbf{X}^N_{ij}, 1 \le i \le j \le N)$ independent, $\mathbf{X}^N_{ij}$ with law μ^N_{ij}, that are probability measures on $\mathbb{C}$ or on $\mathbb{R}$, all of them satisfying the log Sobolev inequality with constant $c < \infty$. Then:*

(1) *For any Lipschitz function f on $\mathbb{R}$, for any $\delta > 0$,*

$$\mu^N\left(|L_{\mathbf{A}^N}(f) - \mu^N[L_{\mathbf{A}^N}(f)]| \ge \delta\right) \le 2e^{-\frac{1}{4c|f|^2_{\mathcal{L}}}N^2\delta^2}.$$

(2) *For any $k \in \{1, \cdots, N\}$,*

$$\mu^N\left(|\lambda_k(\mathbf{A}^N) - \mu^N(\lambda_k(\mathbf{A}^N))| \ge \delta\right) \le 2e^{-\frac{1}{4c}N\delta^2}.$$

The same bound holds for the spectral radius $\lambda_{\max}(\mathbf{A}^N)$.

In particular, these results hold when the $\mathbf{X}_{ij}$ are independent Gaussian variables with bounded above variances.

Note that the same result holds when the entries of $\mathbf{X}$ are real.

Proof of Theorem 3.15. For (H1), the assumption $V''(x) \ge \frac{1}{c}$ implies, by Lemma 3.13, that $\mathbf{A}^N \to N\mathrm{Tr}(V(\mathbf{A}^N))$ is twice continuously differentiable with Hessian bounded below by $\frac{N}{c}$. The second case uses the product property of Lemma 3.3 which implies that $\otimes_{i\le j}\mu^N_{ij}$ satisfies the log Sobolev inequality with constant c. Hence the law μ^N of $\mathbf{A} = \mathbf{X}/\sqrt{N}$ satisfies the log Sobolev inequality with constant c/N.

Thus, to complete the proof of the first result of the theorem, we only need to recall that by Lemma 3.10, $G(A^N_{ij}, 1 \le i \le j \le N) = \mathrm{Tr}(f(\mathbf{A}^N))$ is Lipschitz with constant bounded by $\sqrt{N}|f|_{\mathcal{L}}$ whereas $A^N_{ij}, 1 \le i \le j \le N \to \lambda_k(\mathbf{A})$ is Lipschitz with constant one. For the second, we use Lemma 3.14. □

Exercise 3.16. *State the concentration result when the μ^N_{ij} only satisfy Poincaré inequality.*

Exercise 3.17. *If $\mathbf{A}$ is not Hermitian but have all entries with a joint law of type μ^N as above, show that the law of the spectral radius of $\mathbf{A}$ concentrates.*

Observe that the speed of the concentration we obtained for $\mathrm{Tr}(f(\mathbf{X}^N))$ is optimal (since it agrees with the speed of the central limit theorem). It is also optimal in view of the large deviation principle [**10**] which proves that indeed deviations probabilities are of order e^{-N^2}. However, it does not capture the true scale of the fluctuations of $\lambda_{\max}(\mathbf{A}^N)$ that are of order $N^{-\frac{1}{3}}$. Improvements of concentration inequalities in that direction were obtained by M. Ledoux [**75**].

We emphasize that Theorem 3.15 applies also when the variance of X^N_{ij} depends on i, j. For instance, it includes the case where $X^N_{ij} = a^N_{ij} Y^N_{ij}$ with Y^N_{ij} i.i.d. with law P satisfying the log-Sobolev inequality and a_{ij} uniformly bounded (since if P satisfies the log-Sobolev inequality with constant c, the law of ax under P satisfies it also with a constant bounded by a^2c).

3.14. Concentration inequalities for traces of several random matrices

The previous Theorems also extend to the setting of several random matrices. If we wish to consider polynomial functions of these matrices, we can use local concentration results (see Lemma 3.5). We do not need to assume the random matrices independent if they interact *via* a convex potential.

Let V be a polynomial in m non-commutative indeterminates. Assume that for any $N \in \mathbb{N}$,

$$\phi_V^N : ((A_k)_{ij})_{\substack{i\le j\\ 1\le k\le m}}, \mathbf{A}_1, \cdots, \mathbf{A}_m \in \mathcal{H}_N^{(2)} \to \mathrm{Tr} V(\mathbf{A}_1, \cdots, A_m)$$

is real valued and convex. Let c be a positive real.

$$d\mu_V^{N,\beta}(A_1, \cdots, A_m) := \frac{1}{Z_V^N} e^{-N\mathrm{Tr}(V(A_1,\cdots,A_m))} d\mu_c^{N,\beta}(A_1)\cdots d\mu_c^{N,\beta}(A_m)$$

with $\mu_c^{N,\beta}$ the law of a $N \times N$ Wigner matrix with complex ($\beta = 2$) or real ($\beta = 1$) Gaussian entries with covariance $1/cN$, that is the law of the self-adjoint $N \times N$ matrix A with entries with law

$$\mu_c^{N,2}(dA) = \frac{1}{Z_N^c} e^{-\frac{cN}{2}\sum_{i,j=1}^N |A_{ij}|^2} \prod_{i\le j} d\Re A_{ij} \prod_{i\le j} d\Im A_{ij}$$

and

$$\mu^{N,1}(dA) = \frac{1}{Z_N^c} e^{-\frac{cN}{4}\sum_{i,j=1}^N A_{ij}^2} \prod_{i\le j} dA_{ij} \prod_{i\le j} d\Im A_{ij}.$$

We then have the following corollary.

Corollary 3.18. *Let $\mu_V^{N,\beta}$ be as above. Then*

(1) *For any Lipschitz function f of the entries of the matrices $A_i, 1 \le i \le m$, for any $\delta > 0$,*

$$\mu_V^{N,\beta}(|f - \mu_V^{N,\beta}(f)| > \delta) \le 2e^{-\frac{Nc\delta}{2|f|_{\mathcal{L}}}}.$$

(2) *Let M be a positive real, denote $\Lambda_M^N = \{A_i \in \mathcal{H}_N^{(2)}; \max_{1\le i\le m}\lambda_{\max}(A_i) \le M\}$ and P be a monomial of degree $d \in \mathbb{N}$. Then, for any $\delta > 0$*

$$\mu_V^{N,\beta}\left(\{|\mathrm{Tr}(P(\{X_i\}_{1\le i\le m})) - \mu_V^{N,\beta}(\mathrm{Tr}(P(\{X_i\}_{1\le i\le m})1_{\Lambda_M^N}))| > \delta + \delta(M,N)\} \cap \Lambda_M^N\right)$$
$$\le 2e^{-\frac{c\delta^2}{d^2 M^{2(d-1)}}}$$

with

$$\delta(M,N) \le M^d \mu_V^{N,\beta}\left((1 + d\|A\|_2)1_{(\Lambda_M^N)^c}\right).$$

Proof. By our assumption, the law $\mu_V^{N,\beta}$ of the entries of $(X_1, \cdots, X_m)$ is absolutly continuous with respect to Lebesgue measure. The Hessian of the logarithm of the density is bounded above by $-NcI$. Hence, by Corollary 3.4, $\mu_V^{N,\beta}$ satisfies a log Sobolev inequality with constant $1/Nc$ and thus by Lemma 3.2 we find that $\mu_V^{N,\beta}$ satisfies the first statement of the Corollary. We finally conclude by using Lemma 3.5 and the fact that $X_1, \cdots, X_m \to \mathrm{Tr}(P(X_1, \cdots, X_m))$ is locally Lipschitz by Lemma 3.11. □

4. Brascamp-Lieb inequalities; Applications to random matrices

We introduce Brascamp-Lieb inequalities and show how they can be used to obtain a priori controls for random matrices quantities such as the spectral radius. Such controls will be particularly useful in the next lecture.

4.15. Brascamp-Lieb inequalities

The Brascamp-Lieb inequalities we shall be interested in allow to compare the expectation of convex functions under a Gaussian law and under a law with a log-concave density with respect to this Gaussian law. It states as follows.

Theorem 3.19. *[Brascamp-Lieb[**25**], Hargé [**61**], Theorem 1.1]Let $n \in \mathbb{N}$. Let g be a convex function on $\mathbb{R}^n$ and f a log-concave function on $\mathbb{R}^n$. Let γ be a Gaussian measure on $\mathbb{R}^n$. We suppose that all the following integrals are well defined, then:*

$$\int g(x + l - m)\frac{f(x)d\gamma(x)}{\int f d\gamma} \le \int g d\gamma$$

where

$$l = \int x d\gamma, \, m = \int x \frac{f(x)d\gamma(x)}{\int f d\gamma}.$$

This theorem was proved by Brascamp and Lieb [**25**] (Theorem 7, $g(x) = |x_1|^\alpha$), by Caffarelli [**29**] (Corollary 6, $g(x) = g(x_1)$) and then for a general convex function g by Hargé [**61**].

4.16. Applications of Brascamp-Lieb inequalities to random matrices

We apply now Brascamp-Lieb inequalities to the setting of random matrices. To this end, we must restrict ourselves to random matrices with entries following a law that is absolutely continuous with respect to Lebesgue measure and with strictly log-concave density. We restrict ourselves to the case of m $N \times N$ Hermitian (or symmetric) random matrices with entries following the law

$$d\mu_V^{N,\beta}(A_1, \cdots, A_m) := \frac{1}{Z_V^N} e^{-N\mathrm{Tr}(V(A_1,\cdots,A_m))} d\mu_c^{N,\beta}(A_1) \cdots d\mu_c^{N,\beta}(A_m)$$

with $\mu_c^{N,\beta}$ the law of a $N \times N$ Wigner matrix with complex ($\beta = 2$) or real ($\beta = 1$) Gaussian entries with covariance $1/cN$. We assume that V is convex in the sense that for any $N \in \mathbb{N}$,

$$\{\Re(A_k)_{ij}, i \le j, \Im(A_k)_{ij}, i < j\}_{1 \le k \le m} \to \mathrm{Tr}(V(A_1, \cdots, A_m))$$

is real valued and convex.

This hypothesis covers the case where $V(A_1, \cdots, A_m) = \sum_{i=1}^k V_i(\sum_{j=1}^m \alpha_j^i A_j)$ when α_j^i are real variables and V_i are convex functions on $\mathbb{R}$ by Klein's Lemma 3.13.

Theorem 3.19 implies that for all convex function g on $(\mathbb{R})^{\beta m N(N-1)/2+mN}$,

$$\int g(\mathbf{A} - \mathbf{M}) d\mu_V^{N,\beta}(\mathbf{A}) \le \int g(\mathbf{A}) \prod_{i=1}^m d\mu_c^{N,\beta}(A_i) \tag{37}$$

where $\mathbf{M} = \int \mathbf{A} d\mu_V^{N,\beta}(\mathbf{A})$ is the m-tuple of deterministic matrices $(\mathbf{M}_k)_{ij} = \int (\mathbf{A}_k)_{ij} d\mu_V^{N,\beta}(\mathbf{A})$. In (37), $g(\mathbf{A})$ is a shorthand for a function of the (real and imaginary parts of the) entries of the matrices $\mathbf{A} = (A_1, \cdots, A_m)$.

By different choices of the function g we shall now obtain some a priori bounds on the random matrices $(A_1, \cdots, A_m)$ with law $\mu_c^{N,\beta}$.

Lemma 3.20. *For $c > 0$, there exists $C_0 = C_0(c, V(0), D_i V(0), c(V), d)$ finite such that for all $i \in \{1, \cdots, m\}$, all $n \in \mathbb{N}$,*

$$\limsup_N \mu_V^N(\frac{1}{N}\mathrm{Tr}(X_i^{2n})) \le C_0^n.$$

Moreover, C_0 depends continuously on $V(0), D_i V(0), \sigma_m(V(\frac{X_1}{\sqrt{c}}, \cdots, \frac{X_m}{\sqrt{c}})$ and in particular is uniformly bounded when these quantities are.

Note that this lemma shows that, for $i \in \{1, \cdots, m\}$, the spectral measure of A_i is asymptotically contained in the compact set $[-\sqrt{C_0}, \sqrt{C_0}]$. σ_m is the law of m semi-circular variables as already met in Theorem 1.17. Observe that since for any monomial $q = A_{\ell_1} \cdots A_{\ell k}$, $\sigma^m(q(X/\sqrt{c}))$ is bounded by $(2/\sqrt{c})^k$, $\sigma_m(V(\frac{X_1}{\sqrt{c}}, \cdots, \frac{X_m}{\sqrt{c}}))$ is finite for all polynomial V (and locally bounded in the parameters of V).

Proof. Let k be in $\{1, \cdots, m\}$. As $A \to \mathrm{Tr}(A_k^{4d})$ is convex by Klein's lemma 3.13, Brascamp-Lieb inequality (37) implies that

$$\mu_V^N(\frac{1}{N}\mathrm{Tr}(A_k - M_k)^{4d}) \le \mu_c^{N,\beta}(\frac{1}{N}\mathrm{Tr}(A_k)^{4d}) = \mu_c^{N,\beta}(\mathbf{L}_{A_k}(x^{4d})) \tag{38}$$

where $M_k = \mu_V^N(A_k)$ stands for the matrix with entries $\int (A_k)_{ij} d\mu_V^N(d\mathbf{A})$. Thus, since $\mu_c^{N,\beta}(\mathbf{L}_{A_k}(x^{4d}))$ converges by Wigner theorem 1.1 towards $c^{-2d} C_{2d} \le (c^{-1}4)^{2d}$ with C_{2d} the Catalan number, we only need to control $M_k := \mu_V^N(A_k)$. First observe that for all k the law of A_k is invariant under the multiplication by unitary matrices so that for all unitary matrices U,

$$M_k = \mu_V^N[A_k] = U \mu_V^N[A_k] U^* \Rightarrow M_k = \mu_V^N(\frac{1}{N}\mathrm{Tr}(A_k))I. \tag{39}$$

Let us bound $\mu_V^N(\frac{1}{N}\mathrm{Tr}(A_k))$. Jensen's inequality implies

$$Z_N^V \ge e^{-N^2 \mu_c^{N,\beta}(\frac{1}{N}\mathrm{Tr}(V))}$$

and so

$$\liminf_{N\to\infty}\frac{1}{N^2}\log Z_N^V \geq -\limsup_{N\to\infty}\mu_c^{N,\beta}(\frac{1}{N}\mathrm{Tr}(V)).$$

According to Theorem 1.17, $\mu_c^{N,\beta}(\frac{1}{N}\mathrm{Tr}(V))$ converges as N goes to infinity for any polynomial function V. Thus, for N sufficiently large, if $C(V) = 2|\sigma^m(q(X/\sqrt{c}))|$, $Z_N^V \geq e^{-N^2 C(V)}$.

We now use the convexity of V, to find that for all N,

$$\mathrm{Tr}\,(V(\mathbf{A})) \geq \mathrm{Tr}(V(0) + \sum_{i=1}^{m} D_i V(0) A_i)$$

with D_i the cyclic derivative introduced in Lemma 3.11. By Chebychev's inequality, we therefore obtain, for all $\lambda \geq 0$,

$$\begin{aligned} & \mu_V^N(|\mathbf{L}_{A_k}(x)| > y) \leq \mu_V^N(\frac{1}{N}\mathrm{Tr}(A_k) \geq y) + \mu_V^N(-\frac{1}{N}\mathrm{Tr}(A_k) \geq y) \\ \leq\ & e^{N^2(C(V)-V(0)-\lambda y)}\left(\mu_c^{N,\beta}(e^{-N\mathrm{Tr}(\sum_{i=1}^m D_iV(0)A_i - \lambda A_k)}) + \mu_c^{N,\beta}(e^{-N\mathrm{Tr}(\sum_{i=1}^m D_iV(0)A_i + \lambda A_k)})\right) \\ =\ & e^{N^2(C(V)-V(0)-\lambda y)} e^{\frac{N}{2c}\sum_{\ell\neq k}\mathrm{Tr}(D_iV(0)^2)}\left(e^{\frac{N}{2c}\mathrm{Tr}((D_kV(0)-\lambda)^2)} + e^{\frac{N}{2c}\mathrm{Tr}((D_kV(0)+\lambda)^2)}\right). \end{aligned}$$

Optimizing with respect to λ shows that there exists $B = B(V)$

$$\mu_V^N\,(|\mathbf{L}_{A_k}(x)| \geq y) \leq e^{BN^2 - \frac{N^2 c}{4}y^2}$$

so that for N large enough,

$$\begin{aligned} \mu_V^N(|\mathbf{L}_{A_k}(x)|) &= \int \mu_V^N\,(|\mathbf{L}_{A_k}(x)| \geq y)\,dy \\ &\leq 4\sqrt{c^{-1}B} + \int_{y\geq 4\sqrt{c^{-1}B}} e^{-\frac{N^2c}{4}(y^2 - 4\frac{B}{c})}dy \leq 8\sqrt{Bc^{-1}}. \end{aligned} \tag{40}$$

This completes the proof with (38). ☐

Let us derive some other useful properties due to Brascamp-Lieb inequality. We first obtain an estimate on the spectral radius $\lambda_{\max}^N(\mathbf{A})$, defined as the maximum of the spectral radius of $A_1,\ldots,A_m$ under the law μ_V^N.

Lemma 3.21. *Under the same hypothesis than in the previous lemma, there exists $\alpha = \alpha(c) > 0$ and $M_0 = M_0(V) < \infty$ such that for all $M \geq M_0$ and all integer N,*

$$\mu_V^N(\lambda_{max}^N(\mathbf{A}) > M) \leq e^{-\alpha MN}.$$

Moreover, $M_0(V)$ is uniformly bounded when $V(0)$, $D_iV(0)$ and $c(V)$ are.

Proof. Since the spectral radius $\lambda_{\max}^N(\mathbf{A})$ is a convex function of the entries, we can apply Brascamp-Lieb inequality (37) and Theorem 3.15 (applied with a quadratic potential V) to obtain an exponential bound on $\lambda_{\max}^N(\mathbf{A}-\mathbf{M})$ with $\mathbf{M} = \mathbb{E}[\mathbf{A}]$. But, by (39) and (40), $\lambda_{\max}^N(\mathbf{M})$ is bounded independently of N and therefore, $\lambda_{\max}^N(\mathbf{A}) \in [\lambda_{\max}^N(\mathbf{A}) - |\lambda_{\max}^N(\mathbf{M})|, \lambda_{\max}^N(\mathbf{A}) + |\lambda_{\max}^N(\mathbf{M})|]$ also satisfies an exponential bound. ☐

Lemma 3.22. *If $c > 0$, $\epsilon \in]0, \frac{1}{2}[$, then there exists $C = C(c,\epsilon) < \infty$ such that for all $d \leq N^{\frac{1}{2}-\epsilon}$,*

$$\mu_V^N(|\lambda_{max}^N(\mathbf{A})|^d) \leq C^d.$$

Note that this control could be generalized to $d \leq N^{2/3-\epsilon}$. by using the refinements obtained by Soshnikov, Theorem 2 p.17 in [**92**] but we shall not need it here.

Proof. Since $\mathbf{A} \to \lambda_{\max}^N(\mathbf{A})$ is convex, we can again use Brascamp-Lieb inequalities as well as the uniform bound on $\mathbf{M} = \mu_V^N(\mathbf{A})$ to conclude with Theorem 1.20. ☐

4.17. Coupling concentration inequalities and Brascamp-Lieb inequalities

We next turn to concentration inequalities for trace of polynomials on the set

$$\Lambda_M^N = \{\mathbf{A} \in \mathcal{H}_N^m : \lambda_{\max}^N(\mathbf{A}) = \max_{1\le i\le m}(\lambda_{\max}^N(A_i)) \le M\} \subset \mathbb{R}^{N^2 m}.$$

We let

$$\tilde\delta^N(P) := \mathrm{Tr}(P(A_1,\cdots,A_m)) - \mu_V^N\left(\mathrm{Tr}(P(A_1,\cdots,A_m))\right).$$

Then, we have the following by Corollary 3.18.

Lemma 3.23. *For all N in $\mathbb{N}$, all $M>0$, there exists a finite constant $C(P,M)$ and $\epsilon(P,M,N)$ such that for any $\epsilon>0$,*

$$\mu_V^N\left(\{|\tilde\delta^N(P)|\ge\epsilon+\epsilon(P,M,N)\}\cap\Lambda_M^N\right)\le 2e^{-\frac{c\epsilon^2}{2C(P,M)}}.$$

If P is a monomial of degree d we can choose

$$C(P,M)\le d^2M^{2(d-1)}$$

and there exists $M_0<\infty$ so that for $M\ge M_0$, all $\epsilon\in]0,\frac{1}{2}[$, and all monomial P of degree smaller than $N^{1/2-\epsilon}$,

$$\epsilon(P,M,N)\le 3dN(CM)^{d+1}e^{-\frac{\alpha}{2}NM}$$

with C the constant of Lemma 3.22.

For later purposes, we give a control on the variance of $\mathbf{L}$, that can be easily derived from the previous lemma and the estimate on $\mu_V^N((\Lambda_M^N)^c)$ given in Lemma 3.21.

Lemma 3.24. *For any $c>0$ and $\epsilon\in]0,\frac{1}{2}[$, there exists $B,C,M_0>0$ such that for all $\mathbf{t}\in\mathbf{B}_{\eta,\mathbf{c}}$, all $M\ge M_0$, and monomial P of degree less than $N^{\frac{1}{2}-\epsilon}$,*

$$\mu_V^N\left((\tilde\delta^N(P))^2\right)\le BC(P,M)+C^{2d}N^4e^{-\frac{\alpha MN}{2}}. \tag{41}$$

Moreover, the constants C,M_0,B depend continuously on $V(0),D_iV(0)$ and $c(V)$.

Exercise 3.25. *The goal of this exercise is to give a new proof of Wigner's theorem in the case where the entries are independent Gaussian variables, by using concentration inequalities. This can be viewed as a warm up towards the next lecture. So we let X_N be a $N\times N$ symmetric matrix such that $X_N(ij), i<j$ are i.i.d $N(0,1/N)$(real centered gaussian with covariance $1/N$) and $X_N(ii), 1\le i\le N$ are independent, independent from $X_N(ij), i<j$ and $N(0,2/N)$ distributed.*

(1) *Show that for every differentiable function f that goes to infinity more slowly than $e^{N\frac{x^2}{2}}$,*

$$\int_{-\infty}^{\infty} xf(x)e^{-N\frac{x^2}{2}}dx = N^{-1}\int_{-\infty}^{\infty} f'(x)e^{-N\frac{x^2}{2}}dx.$$

This is also known as Stein 's Lemma.

(2) *Show that for every $z\in\mathbb{C}^+=\{z:\Im(z)>0\}$, any indices i,j,k,l*

$$\partial_{X_{ij}}[(z-X_N)^{-1}]_{kl} = [(z-X_N)^{-1}\Delta_{ij}(z-X_N)^{-1}]_{kl}$$

with Δ_{ij} the matrix with null entries except at (ij) and (ji) where the entries equal one.

(3) *Let $z\in\mathbb{C}^+=\{z:\Im(z)>0\}$. Show that*

$$\mathbb{E}[\mathrm{Tr}(\frac{X_N}{z-X_N})] = \frac{1}{N}\mathbb{E}[\mathrm{Tr}((z-X_N)^{-1})\mathrm{Tr}((z-X_N)^{-1}) + \mathrm{Tr}((z-X_N)^{-2})].$$

Hint: Write

$$\mathbb{E}[\mathrm{Tr}(\frac{X_N}{z-X_N})] = \sum_{ij}\mathbb{E}[X(ij)[(z-X_N)^{-1}]_{ji}]$$

apply Stein's Lemma and the previous formula.

(4) *Using*

$$\mathbb{E}[\mathrm{Tr}(\frac{z}{z-X_N})] = N + \mathbb{E}[\mathrm{Tr}(\frac{X_N}{z-X_N})]$$

deduce that

$$z\mathbb{E}[\frac{1}{N}\mathrm{Tr}(\frac{1}{z-X_N})] = 1 + \mathbb{E}[(\frac{1}{N}\mathrm{Tr}((z-X_N)^{-1}))^2 + \frac{1}{N^2}\mathrm{Tr}((z-X_N)^{-2})].$$

(5) *Let z with positive imaginary part. Deduce from the fact that X_N is symmetric that*
$$|\mathrm{Tr}((z-X_N)^{-2})| \le N/|\Im z|^2$$
and that $X(ij), 1 \le i \le j \le N \to \mathrm{Tr}((z-X_N)^{-1})$ is Lipschitz. Deduce by concentration inequalities that
$$|\mathbb{E}[(\frac{1}{N}\mathrm{Tr}((z-X_N)^{-1}))^2] - \mathbb{E}[(\frac{1}{N}\mathrm{Tr}((z-X_N)^{-1}))]^2| \le \frac{constant}{N|\Im(z)|^2}.$$

(6) *Show that $\{\mathbb{E}[(\frac{1}{N}\mathrm{Tr}((z-X_N)^{-1}))], N \in \mathbb{N}\}$ is a tight sequence and that its limit points $G(z)$ satisfy*
$$zG(z) = 1 + G(z)^2.$$

(7) *Arguing that $G(z)$ must be small for $\Im z$ large, prove that*
$$G(z) = \frac{1}{2}(z - \sqrt{z^2-4}).$$
The rest of the exercise shows that this is sufficient to show that the empirical measure L_{X_N} converges almost surely and in expectation, that the Stieljes transform of the limit is given by $G(z)$ and that it is indeed the semi-circular law.

(8) *We first show that for all $z \in \mathbb{C}^+$, $G_N(z) := \mathbb{E}[(\frac{1}{N}\mathrm{Tr}((z-X_N)^{-1}))]$ converges towards $G(z)$ as above. Show by Arzela-Ascoli theorem that for any $\epsilon > 0$, $\{z : \Im z > \epsilon \to G_N(z)\}_{N\in\mathbb{N}}$ are tight as a sequence of bounded continuous functions. Consider a limit point. Argue that it is analytic (observe that the G_N are analytic and uniformly bounded) and so uniquely defined by the previous point. Conclude.*

(9) *Show that*
$$G(z) = \int \frac{1}{z-x} d\sigma(x)$$
with σ the semi-circle distribution with covariance one.

(10) *Use Cauchy formula to show that for any analytic function f in a strip ($\{z : |\Im z| < \epsilon\}$) around the real line which goes to zero at infinity,*
$$\lim_{N\to\infty} \mathbb{E}[\int f(x) dL_{X_N}(x)] = \int f d\sigma(x).$$
Hint: Cauchy formula says that $f(x) = \pi^{-1}\int_\gamma (x-y)^{-1} f(y)dy$ for a contour γ included in the strip of analyticity of f so that x belongs to the interior of the set delimited by γ. Conclude by density that the same convergence holds for any bounded continuous function that vanishes outside a compact.

(11) *Extend the previous result to all bounded continuous functions.*

(12) *Use concentration inequalities to obtain almost sure convergence of L_{X_N}.*

4
Matrix models

In this lecture, we shall be interested in laws of interacting matrices of the form

$$d\mu_V^N(X_1,\cdots,X_m) := \frac{1}{Z_V^N} e^{-N\mathrm{Tr}(V(X_1,\cdots,X_m))} d\mu^N(X_1)\cdots d\mu^N(X_m)$$

where Z_V^N is the normalizing constant

$$Z_V^N = \int e^{-N\mathrm{Tr}(V(X_1,\cdots,X_m))} d\mu^N(X_1)\cdots d\mu^N(X_m)$$

and V is a polynomial in m non-commutative indeterminates;

$$V(X_1,\cdots,X_m) = \sum_{i=1}^n t_i q_i(X_1,\cdots,X_m)$$

with q_i non-commutative monomials;

$$q_i(X_1,\cdots,X_m) = X_{j_1^i}\cdots X_{j_{r_i}^i}$$

for some $j_l^k \in \{1,\cdots,m\}$, $r_i \geq 1$. Moreover, $d\mu^N(X)$ denotes the standard law of the **GUE**, i.e under $d\mu^N(X)$, X is a $N\times N$ Hermitian matrix such that

$$X_{kl} = \bar{X}_{lk} = \frac{g_{kl} + i\tilde{g}_{kl}}{\sqrt{2N}},\ k<l, \quad X_{kk} = \frac{g_{kk}}{\sqrt{N}}$$

with independent centered standard Gaussian variables $(g_{kl},\tilde{g}_{kl})_{k\leq l}$. In other words

$$d\mu^N(X) = Z_N^{-1} 1_{X\in\mathcal{H}_N^{(2)}} e^{-\frac{N}{2}\mathrm{Tr}(X^2)} \prod_{1\leq i\leq j\leq N} d\Re(X_{ij}) \prod_{1\leq i<j\leq N} d\Im(X_{ij}).$$

Let us denote $\mathbb{C}\langle X_1,\cdots,X_m\rangle$ the set of polynomials in m non-commutative indeterminates and, for $P\in\mathbb{C}\langle X_1,\cdots,X_m\rangle$,

$$\hat{\mathbf{L}}^N(P) := \mathbf{L}_{X_1,\cdots,X_m}(P) = \frac{1}{N}\mathrm{Tr}\left(P(X_1,\cdots,X_m)\right)$$

When V is null, we have seen in Lecture 2 that for all $P\in\mathbb{C}\langle X_1,\cdots,X_m\rangle$, $\hat{\mathbf{L}}^N(P)$ converges as N goes to infinity. Moreover the limit $\sigma^m(P)$ is such that if P is a monomial, $\sigma^m(P)$ is the number of non crossing partitions of a set of points with m colors, or equivalently the number of planar map with one star of type P. In this part, we shall generalize such a type of result to the case where V is not null but 'small' and 'nice' in a sense to precise.

This lecture is motivated by the work of Brézin, Itzykson, Parisi and Zuber [**26**] and large developments that occurred thereafter in theoretical physics [**35**]. They in fact noticed that if $V = \sum_{i=1}^n t_i q_i$ with fixed monomials q_i of m non-commutative indeterminates, and if we see $Z_V^N = Z_{\mathbf{t}}^N$ as a function of $\mathbf{t} = (t_1,\cdots,t_n)$

$$\log Z_{\mathbf{t}}^N := \sum_{g\geq 0} N^{2-2g} F_g(\mathbf{t}) \tag{42}$$

where

$$F_g(\mathbf{t}) := \sum_{k_1,\cdots,k_n\in\mathbb{N}^k\setminus\{0,\cdots,0\}} \prod_{i=1}^k \frac{(-t_i)^{k_i}}{k_i!} \mathcal{M}_g((q_i,k_i)_{1\leq i\leq k})$$

is a generating of integer numbers $\mathcal{M}_g((q_i,k_i)_{1\leq i\leq k})$ that count certain graphs called maps. A map is a connected oriented diagram that is embedded into a surface. Its genus g is by definition the genus of a surface in which it can be embedded in such a way that edges do not cross and the faces of the graph (that are defined by following the boundary of the graph) are homeomorphic to a disc. The vertices of the maps we shall consider will have the structure of a star; a star of type q, for some monomial $q = X_{\ell_1}\cdots X_{\ell_k}$, is a vertex with valence $\deg(q)$ and oriented colored half-edges with one marked half edge of color ℓ_1, the second of color ℓ_2 etc until the last one of color ℓ_k. $\mathcal{M}_g((q_i,k_i)_{1\leq i\leq k})$ is then the number of maps with k_i stars of type q_i, $1\leq i\leq n$. Observe that a star of type q is in bijection with the set $S(q)$ of ordered colored points introduced in Definition 1.16. When $V\neq 0$, the numbers of interest will depend on several such sets that will have matching among each other (the total graph being connected). To describe this global picture, it is thus prettier to draw the set $S(q)$ as ordered dots on a circle or, as we do here, as end points of half-edges of a vertex.

The equality (42) obtained by 't Hooft [**94**] (in a more general context) and then by Brezin, Parisi, Itzykson and Zuber [**26**] was only formal, i.e means that all the derivatives on both sides of the equality coincide at $\mathbf{t} = 0$. This result can then be deduced from Wick formula which gives the expression of arbitrary moments of Gaussian variables (see section 2).

Adding to V a term $t\,q$ for some monomial q and identifying the first order derivative with respect to t at $t = 0$ we derive from (42)

$$\mu_V^N(\hat{\mathbf{L}}^N(q)) = \sum_{g\geq 0} N^{-2g} \sum_{k_1,\cdots,k_n\in\mathbb{N}^k} \prod_{i=1}^k \frac{(-t_i)^{k_i}}{k_i!}\mathcal{M}_g((q_i,k_i)_{1\leq i\leq k},(q,1)). \tag{43}$$

Expansions such as (42) and (43) where first introduced by 't Hooft [**94**] to compute integrals such as the one in the left hand side of (43). When a few years later, Brezin, Parisi, Itzykson and Zuber [**26**] specialized the work of 't Hooft to matrix integrals as in (43) they already study the the natural reverse question of computing the numbers $\mathcal{M}_g((q_i,k_i)_{1\leq i\leq k})$ by studying the associated integrals over matrices encountered a large success in theoretical physics (see e.g. the review papers [**35**]). In the course of doing so, one would like for instance to compute $\lim_{N\to\infty} N^{-2}\log Z_{\mathbf{t}}^N$ and claim that this limit has to be equal to $F_0(\mathbf{t})$. There is here the belief that one can interchange derivatives and limit, a claim that we shall study in this lecture.

In fact, the formal limit can be straightened into a large N expansion in the sense that for all integer number n, for sufficiently small t_i's , (43) can be turned into the large N expansion

$$\mu_V^N[\hat{\mathbf{L}}_N(P)] = \sum_{p=0}^n \frac{1}{N^{2p}}\sigma_p^V(P) + o(N^{-2n}) \tag{44}$$

where $\sigma_g^V(q) = \sum_{k_1,\cdots,k_n\in\mathbb{N}^k}\prod_{i=1}^k \frac{(-t_i)^{k_i}}{k_i!}\mathcal{M}_g((q_i,k_i)_{1\leq i\leq k},(q,1))$ for monomial functions q.

This requires of course that V satisfies some additional hypothesis, for instance insuring that Z_N^V is finite. A natural hypothesis is to assume that $\mathrm{Tr}(V(X_1,\cdots,X_m))$ is a convex function of the entries. Such an assumption can be relaxed by adding a cutoff to the integral but we shall not consider this issue in these notes.

Observe that (44) can be compared to cluster type expansion; recall for instance that Dobrushin [**37**] proved that if we consider the Ising model $V(s) = \sum_{i\approx j} s_i s_j$ with i,j in some box Λ, $s_i \in \{+1,-1\}$, and $i\approx j$ meaning that i and j are nearest neighbours on the lattice,

$$Z_\Lambda^\beta = \frac{1}{|\Lambda|}\log \sum_{\substack{s_i=\{+1,-1\}\\ i\in\Lambda}} e^{\beta V(s)}$$

expands analytically as a function of β in the vicinity of the origin. Moreover, the radius of convergence does not depend on Λ. The main difference with (44) is that in the case of large random matrices, the expansion is analytic with a radius of convergence that does not depend on N only if we cut the expansion at some n (in particular the full series diverges in general).

(44)'s type of expansion have been first derived in the context of one matrix in [**2, 1, 40**]. The methods however used orthogonal polynomials techniques, which are not available in general in the context of several matrices. For several matrices, it was proved in the series of papers [**50, 52, 77**] (the expansion up to $n = 0$ being derived in [**50**], up to $n = 1$ in [**52**] and the full expansion in [**77**]). This lecture summarizes the results from [**50**] that concerns only the first order expansion.

1. Combinatorics of maps and non-commutative polynomials

In this section, we introduce the set up of this lecture, namely non-commutative polynomials and non-commutative laws such as the 'empirical distribution' of matrices $A_1,\cdots,A_m$ simply given as the complex valued linear functional on the set of polynomials which associates to a polynomial the normalized trace of the polynomial evaluated at $A_1,\cdots,A_m$. We will then describe precisely the combinatorial objects related with matrix integrals. Introducing the bijection between non-commutative monomials and graphical objects such as 'stars' or ordered sets of colored point, we will show how natural operations such as derivatives on monomials have their graphical interpretation. This will be our basis to show that some differential equations for non-commutative laws can be interpreted in terms of induction relations for map enumeration.

1.18. Non-commutative polynomials

We denote by $\mathbb{C}\langle X_1, \cdots, X_m\rangle$ the set of complex polynomials in the non-commutative unknowns $X_1,\ldots,X_m$. Let $*$ denote the linear involution such that for all complex z and all monomials

$$(zX_{i_1}\ldots X_{i_p})^* = \overline{z}X_{i_p}\ldots X_{i_1}. \tag{45}$$

We will say that a polynomial P is self-adjoint if $P = P^*$ and denote $\mathbb{C}\langle X_1, \cdots, X_m\rangle_{sa}$ the set of self-adjoint elements of $\mathbb{C}\langle X_1, \cdots, X_m\rangle$.

The potential V will be later on assumed to be self-adjoint. This means that

$$V(\mathbf{A}) = \sum_{j=1}^{n} t_j q_i = \sum_{j=1}^{n} \bar{t}_j q_j^* = \sum_{j=1}^{n} \Re(t_j)\frac{q_j + q_j^*}{2} + \sum_{j=1}^{n} \Im(t_j)\frac{q_j - q_j^*}{2i}.$$

Note that the parameters $(t_j = \Re(t_j)+i\Im(t_j), 1 \le j \le n)$ may a priori be complex. This hypothesis guarantees that $\mathrm{Tr}(V(\mathbf{A}))$ is real for all $\mathbf{A} = (A_1, \cdots, A_m)$ in the set $\mathcal{H}_N^{(2)}$ of $N \times N$ Hermitian matrices.

In the sequel, the monomials $(q_i)_{1\le i\le n}$ will be fixed and we will consider $V = V_{\mathbf{t}} = \sum_{i=1}^n t_i q_i$ as the parameters t_i vary in such a way that V stays self-adjoint.

1.18.1. *Convexity.* We shall assume also that V satisfies some convexity property. Namely, we will say that V is convex if V is self-adjoint and for any $N \in \mathbb{N}$

$$\begin{array}{llll} \phi_V^N & : (\mathcal{H}_N^{(2)})^m \simeq \mathbb{R}^{N^2 m} & \longrightarrow & \mathbb{R} \\ & (A_1, \cdots, A_m) & \longrightarrow & \mathrm{Tr}(V(A_1, \cdots, A_m)) \end{array}$$

is a convex function of the entries of the Hermitian matrices $A_1, \cdots, A_m$.

While it may not be the optimal hypothesis, convexity provides many simple arguments. Note that as we add a Gaussian potential $\frac{1}{2}\sum_{i=1}^m X_i^2$ to V we can relax the hypothesis a little.

Definition 4.1. *We say that V is c-convex if $c > 0$ and $V + \frac{1-c}{2}\sum_1^m X_i^2$ is convex. In other words, the Hessian of*

$$\begin{array}{llll} \phi_V^{N,c} & : (\mathbb{R})^{N^2 m} & \longrightarrow & \mathbb{R} \\ & (\Re(A_k(ij)), \Im(A_k(ij)))_{1\le i\le j\le N}^{1\le k\le m} & \longrightarrow & \mathrm{Tr}(V(A_1, \cdots, A_m) + \frac{1-c}{2}\sum_{k=1}^m A_i^2) \end{array}$$

is non-negative. Here, for $k \in \{1, \cdots, m\}$, A_k is the Hermitian matrix with entries $\sqrt{2}^{-1}(A_k(pq)+ iA_k(qp))$ above the diagonal and A_{ii} on the diagonal.

An example is

$$V = \sum_{i=1}^{n} P_i\left(\sum_{k=1}^{m} \alpha_k^i A_k\right) + \sum_{k,l=1}^{n} \beta_{k,l} A_k A_l$$

with convex real polynomials P_i in one unknown, real parameters α_k^i and, for all l, $\sum_k |\beta_{k,l}| \le (1-c)$. This is due to Klein's Lemma 3.13.

Note that when V is c-convex, μ_V^N has a log-concave density with respect to Lebesgue measure so that many results from the previous lecture will apply, in particular concentration inequalities and Brascamp-Lieb inequalities.

In the rest of this lecture, we shall assume that V is c-convex for some $c > 0$ fixed. Arbitrary potentials could be considered as far as first order asymptotics are considered in [**51**], at the price of adding a cutoff. In fact, adding a cutoff and choosing the parameters t_i's small enough (depending possibly on this cutoff), forces the interaction to be convex so that most of the machinery we are going to describe will apply also in this context. We choose here to restrict ourselves to convex potentials. Since $V = V_{\mathbf{t}}$ with $\mathbf{t}$ varying but fixed monomials, we will let $U_c = \{\mathbf{t} : V_{\mathbf{t}} \text{ is c-convex}\} \subset \mathbb{C}^n$.

1.18.2. *Non-commutative derivatives.* First, for $1\le i\le m$, let us define the non-commutative derivatives ∂_i with respect to the variable X_i. They are linear maps from $\mathbb{C}\langle X_1,\cdots,X_m\rangle$ to $\mathbb{C}\langle X_1,\cdots,X_m\rangle^{\otimes 2}$ given by the Leibniz rule

$$\partial_i PQ = \partial_i P\times(1\otimes Q)+(P\otimes 1)\times\partial_i Q$$

and $\partial_i X_j = \mathbf{1}_{i=j}1\otimes 1$. Here, $\times$ is the multiplication on $\mathbb{C}\langle X_1,\cdots,X_m\rangle^{\otimes 2}$; $P\otimes Q\times R\otimes S = PR\otimes QS$. So, for a monomial P, the following holds

$$\partial_i P = \sum_{P=RX_iS} R\otimes S$$

where the sum runs over all possible monomials R,S so that P decomposes into RX_iS.

Exercise 4.2. *Show that when $m=1$,*

$$\partial_1 X^k = \sum_{j=0}^{k-1} X^j\otimes X^{k-1-j}.$$

Identifying $\mathbb{C}[X]\otimes\mathbb{C}[X]$ with $\mathbb{C}[X,Y]$ we thus have

$$\partial_1 P(X,Y) = \frac{P(X)-P(Y)}{X-Y}.$$

Notice that ∂_i arises naturally when considering derivatives of polynomials in matrices since for any $N\times N$ Hermitian matrices $(X_1,\cdots,X_m)$, any polynomial P, $P(X_1,\cdots,X_m)$ is an $N\times N$ matrix and for any indices $p\in\{1,\cdots,m\}$, $i,j.k,\ell\in\{1,\cdots,N\}$

$$\partial_{X_p(ij)}(P(X_1,\cdots,X_m))_{k\ell} = [\partial_p P]_{kj,i\ell}$$

where $(A\otimes B)_{kj,i\ell} = A_{kj}B_{i\ell}$.

We can iterate the non-commutative derivatives; for instance

$$\partial_i^2:\mathbb{C}\langle X_1,\cdots,X_m\rangle\to\mathbb{C}\langle X_1,\cdots,X_m\rangle\otimes\mathbb{C}\langle X_1,\cdots,X_m\rangle\otimes\mathbb{C}\langle X_1,\cdots,X_m\rangle$$

is given for a monomial function P by

$$\partial_i^2 P = 2\sum_{P=RX_iSX_iQ} R\otimes S\otimes Q.$$

We denote by $\sharp:\mathbb{C}\langle X_1,\cdots,X_m\rangle^{\otimes 2}\times\mathbb{C}\langle X_1,\cdots,X_m\rangle\to\mathbb{C}\langle X_1,\cdots,X_m\rangle$ the map $P\otimes Q\sharp R = PRQ$ and generalize this notation to $P\otimes Q\otimes R\sharp(S,V) = PSQVR$. So $\partial_i P\sharp R$ corresponds to the derivative of P with respect to X_i in the direction R, and similarly $2^{-1}[D_i^2P\sharp(R,S)+D_i^2P\sharp(S,R)]$ the second derivative of P with respect to X_i in the directions R,S.

We also define the so-called cyclic derivative D_i. If m is the map $m(A\otimes B)=BA$, let us define $D_i = m\circ\partial_i$. For a monomial P, D_iP can be expressed as

$$D_i P = \sum_{P=RX_iS} SR.$$

Note that we have for any $N\times N$ Hermitian matrices $(X_1,\cdots,X_m)$, any polynomial P, $P(X_1,\cdots,X_m)$ is an $N\times N$ matrix and for any indices $p\in\{1,\cdots,m\}$, $i,j.k,\ell\in\{1,\cdots,N\}$

$$\partial_{X_p(ij)}\mathrm{Tr}(P(X_1,\cdots,X_m)) = [D_pP]_{ij} \tag{46}$$

as was already noticed in Lemma 3.11.

Exercise 4.3. *Show that when $m=1$, $D_1P=P'$.*

1.18.3. *Non-commutative laws.* For $(A_1,\cdots,A_m)\in(\mathcal{H}_N^{(2)})^m$, let us define the linear form $\mathbf{L}_{A_1,\cdots,A_m}$ from $\mathbb{C}\langle X_1,\cdots,X_m\rangle$ into $\mathbb{C}$ by

$$\mathbf{L}_{A_1,\cdots,A_m}(P) = \frac{1}{N}\mathrm{Tr}\left(P(A_1,\cdots,A_m)\right)$$

where Tr is the standard trace $\mathrm{Tr}(A)=\sum_{i=1}^N A_{ii}$. When the matrices $A_1,\cdots,A_m$ are generic and distributed according to μ_V^N, we will drop the subscripts $A_1,\cdots,A_m$ and write in short $\hat{\mathbf{L}}^N = \mathbf{L}_{A_1,\cdots,A_m}$. We denote

$$\bar{\mathbf{L}}_t^N(P) := \mu_{V_{\mathbf{t}}}^N[\hat{\mathbf{L}}^N(P)].$$

$\hat{\mathbf{L}}^N, \bar{\mathbf{L}}_t^N$ will be seen as elements of the algebraic dual $\mathbb{C}\langle X_1, \cdots, X_m\rangle^*$ of $\mathbb{C}\langle X_1, \cdots, X_m\rangle$ equipped with the involution $*$; we shall call in these notes non-commutative laws elements of $\mathbb{C}\langle X_1, \cdots, X_m\rangle^*$; this is a very weak point of view that, however, is sufficient for our purpose. The name 'law' at least is justified when $m = 1$, in which case $\mathbf{L}_A = L_A$ is the spectral measure of the matrix A, and hence a probability measure on $\mathbb{R}$, whereas the non-commutativity is clear when $m \geq 2$. There are much deeper reasons for this name when considering C^*-algebras and positivity, and we refer the reader to [**103**] or [**3**]. $\mathbb{C}\langle X_1, \cdots, X_m\rangle^*$ is equipped with its weak topology.

Definition 4.4. *A sequence $(\mu_n)_{n\in\mathbb{N}}$ in $\mathbb{C}\langle X_1, \cdots, X_m\rangle^*$ converges weakly towards $\mu \in \mathbb{C}\langle X_1, \cdots, X_m\rangle^*$ iff for any $P \in \mathbb{C}\langle X_1, \cdots, X_m\rangle$,*

$$\lim_{n\to\infty} \mu_n(P) = \mu(P).$$

The following two lemmas are trivial; I however remind them to the reader to possibly reduce uneasiness related with the non-commutative setting.

Lemma 4.5. *Let $C(\ell_1, \cdots, \ell_r), \ell_i \in \{1, \cdots, m\}, r \in \mathbb{N}$, be finite non-negative constants and*

$$K(C) = \{\mu \in \mathbb{C}\langle X_1, \cdots, X_m\rangle^*; |\mu(X_{\ell_1} \cdots X_{\ell_r})| \leq C(\ell_1, \cdots, \ell_r)\, \forall \ell_i \in \{1, \cdots, m\}, r \in \mathbb{N}\}.$$

Then, any sequence $(\mu_n)_{n\in\mathbb{N}}$ in $K(C)$ is sequentially compact, i.e. has a subsequence $(\mu_{\phi(n)})_{n\in\mathbb{N}}$ that converges weakly.

Proof. Since $\mu_n(X_{\ell_1} \cdots X_{\ell_r}) \in \mathbb{C}$ is uniformly bounded, it has converging subsequences. By a diagonalization procedure, since the set of monomials is countable, we can ensure that for a subsequence $(\phi(n), n \in \mathbb{N})$, the terms $\mu_{\phi(n)}(X_{\ell_1} \cdots X_{\ell_r}), \ell_i \in \{1, \cdots, m\}, r \in \mathbb{N}$ converge simultaneously. The limit defines an element of $\mathbb{C}\langle X_1, \cdots, X_m\rangle^*$ by linearity. □

Corollary 4.6. *Let $C(\ell_1, \cdots, \ell_r), \ell_i \in \{1, \cdots, m\}, r \in \mathbb{N}$, be finite non negative constants and $(\mu_n)_{n\in\mathbb{N}}$ a sequence in $K(C)$ that has a unique limit point. Then $(\mu_n)_{n\in\mathbb{N}}$ converges towards this limit point.*

Proof. Otherwise we could choose a subsequence that stays at positive distance of this limit point, but extracting again a converging subsequence gives a contradiction. Note as well that any limit point will belong automatically to $\mathbb{C}\langle X_1, \cdots, X_m\rangle^*$. □

Remark 4.7. *The laws $\hat{\mathbf{L}}^N, \bar{\mathbf{L}}_t^N$ are more than only linear forms on $\mathbb{C}\langle X_1, \cdots, X_m\rangle$; they satisfy also the properties that define tracial states, namely*

$$\mu(PP^*) \geq 0, \quad \mu(PQ) = \mu(QP), \mu(1) = 1$$

for all polynomial functions P, Q. Since these conditions are closed for the weak topology, we see that any limit point of $\hat{\mathbf{L}}^N, \bar{\mathbf{L}}_t^N$ will as well satisfy these properties. A linear functional on $\mathbb{C}\langle X_1, \cdots, X_m\rangle$ that satisfies such conditions are called tracial states. This leads to the notion of C^*-algebras and representations of the laws as moments of non-commutative operators on a C^*-algebras. We however do not want to detail this point in these notes.

1.19. Maps and polynomials

In this section, we complete section 2.6 to describe the graphs that shall be enumerated by matrix models. Let $q(X_1, \cdots, X_m) = X_{\ell_1} X_{\ell_2} \cdots X_{\ell_k}$ be a monomial in m non-commutative indeterminates.

A star of type q is a vertex equipped with k colored half-edges, one marked half-edge and an orientation such that the marked half-edge is of color ℓ_1, the second (following the orientation) of color ℓ_2 etc until the last half-edge of color ℓ_k. Maps are obtained by gluing half-edges pairwise. This graph can be embedded into a surface in a unique way (up to homeomorphisms of the graphs) so that the orientation of the stars agree with the orientation of the surface and the faces are homeomorphICC to discs. The genus of a map is the genus of the surface in which it is embedded.

Hereafter monomials $(q_i)_{1\leq i\leq n}$ will be fixed and we will denote in short, for $\mathbf{k} = (k_1, \cdots, k_n)$,

$$\begin{aligned} \mathcal{M}_{\mathbf{k}}^g = \text{card}\{ & \text{ maps with genus } g \\ & \text{and } k_i \text{ stars of type } q_i, 1 \leq i \leq n\} \end{aligned}$$

and for a monomial P

$$\begin{aligned} \mathcal{M}_{\mathbf{k}}^g(P) = \text{card}\{ & \text{ maps with genus } g \\ & k_i \text{ stars of type } q_i, 1 \leq i \leq n \text{ and one of type } P\} \end{aligned}$$

2. Formal expansion of matrix integrals

The expansion obtained by Brezin, Itzykson, Parisi and Zuber [**26**] is based on Feynmann diagrams, or equivalently on Wick Formula that states as follows.

Lemma 4.8. *Let $(G_1, \cdots, G_{2n})$ be a Gaussian vector such that $\mathbb{E}[G_i] = 0$ for $i \in \{1, \cdots, 2n\}$. Then,*

$$\mathbb{E}[G_1 \cdots G_{2n}] = \sum_{\pi \in PP(2n)} \prod_{\substack{(b,b') \text{ block of } \pi, \\ b<b'}} \mathbb{E}[G_b G_{b'}]$$

where the sum runs over all pair-partitions of the ordered set $\{1, \cdots, 2n\}$.

We leave the proof of this formula as an exercise; a proof is based on the fact that $E[G^{2n}] = 2n!!$ for any standard Gaussian G and that linear combinations of Gaussian variables is Gaussian.

We now consider moments of traces of Gaussian Wigner's matrices. Since we shall consider the moments of products of several traces, we shall now use the language of stars. Let us recall that a star of type $q(X) = X_{\ell_1} \cdots X_{\ell_2}$ is a vertex equipped with k colored half-edges, one marked half-edge and an orientation such that the marked half-edge is of color ℓ_1, the second (following the orientation) of color ℓ_2 etc till the last half-edge of color ℓ_k. The graphs we shall enumerate will be obtained by gluing pairwise the half-edges.

Definition 4.9. *Let $r, m \in \mathbb{N}$. Let $q_1, \cdots, q_r$ be r monomials in m non-commutative indeterminates. A map with a star of type q_i for $i \in \{1, \cdots, r\}$ is a connected oriented graph with r vertices so that*

(1) *for $1 \le i \le r$, one of the vertices has degree $deg(q_i)$, and this vertex is equipped with the structure of a star of type q_i (i.e. with the corresponding colored half-edges and orientation).*
(2) *The half-edges of the stars are glued pair-wise and two half-edges can be glued iff they have the same color and the same orientation; thus edges have only one color and only one orientation.*

Because of the imposed agreement in the orientation of the stars, each edge is oriented in agreement with the orientation at the vertex; if we follow the orientation from one edge, we end up making a cycle. The surface inside this cyclic curve is homeomorphic to a disk and called a face. The genus g of a map is such that $2-2g$ is the number of vertices, plus the number of edges minus the number of faces. Equivalently, we can draw the map on a surface of genus g in such a way that edges do not cross, faces are homeomorphic to a disk and the orientation of the stars agrees with the orientation of the surface.

We shall soon encounter the question of counting the number of maps with given numbers of stars of a given type and a given genus. In this counting, stars will be labeled and therefore, since stars are oriented and rooted, all half-edges of the stars are labeled. Thus, two maps will be considered to be the same only if they were constructed by matching (or gluing) half-edges with the same labels.

There is a dual way to consider maps; we can replace a star of type $q(X) = X_{i_1} \cdots X_{i_p}$ by a polygon (of type q) with p faces, a boundary edge of the polygon replacing an edge of the star and taking the same color as the edge, and a marked boundary edge and an orientation. A map is then a tiling of a surface (with the same genus as the map) by polygons of type $q_1, \cdots, q_r$ with colored sides, only sides of the same color being matched together.

Example 4.10. *A triangulation (resp. a quadrangulation) of a surface of genus g by F faces (the number of triangles, resp. squares) is equivalent to a map of genus g with F stars of type $q(X) = X^3$ (resp. $q(X) = X^4$) .*

We will denote for $\mathbf{k} = (k_1, \cdots, k_n)$,

$$\mathcal{M}_g((q_i, k_i), 1 \le i \le n) = \text{card}\{ \text{ maps with genus } g \text{ and } k_i \text{ stars of type } q_i, 1 \le i \le n\}.$$

In this section we shall first encounter possibly non-connected graphs; these graphs will then be (finite) union of maps. We denote by $G_{g,c}((q_i, k_i), 1 \le i \le n)$ the set of graphs that can be described as a union of c maps, the total set of stars to construct these maps being k_i stars of type

q_i, $1 \le i \le n$ and the genus of each connected components summing up to g. When counting these graphs we will also assume that all half-edges are labeled.

We now argue that

Lemma 4.11. *Let $q_1, \cdots, q_n$ be monomials. Then,*

$$\int \prod_{i=1}^{n} (N\mathrm{Tr}(q_i(X_1, \cdots, X_m))) d\mu_N(X_1) \cdots d\mu_N(X_m) = \sum_{g\in\mathbb{N}} \sum_{c\ge 1} \frac{1}{N^{2g-2c}} \sharp\{G_{g,c}((q_i, 1), 1 \le i \le n)\}$$

Here $\sharp\{G_{g,c}((q_i, 1), 1 \le i \le n\}$ is the number of different graphs (up to homeomorphism) of the set $G_{g,c}((q_i, 1), 1 \le i \le n)$. In particular, $\sharp\{G_0((q, 1))\}$ equals $\sigma_m(q)$ as found by Voiculescu, Theorem 1.17.

As a warm up, let us show that

Lemma 4.12. *Let q be a monomial. Then,*

$$\int N^{-1}\mathrm{Tr}(q(X_1, \cdots, X_m)) d\mu_N(X_1) \cdots d\mu_N(X_m)$$

$$= \sum_{g\in\mathbb{N}} \frac{1}{N^{2g}} \sharp\{G_g((q, 1))\}$$

Proof. As usual we expend the trace and write, if $q(X_1, \cdots, X_m) = X_{j_1} \cdots X_{j_k}$,

$$\int \mathrm{Tr}(q(X_1, \cdots, X_m)) d\mu_N(X_1) \cdots d\mu_N(X_m) \tag{47}$$

$$= \sum_{r_1\cdots r_k} \int X_{j_1}(r_1 r_2) \cdots X_{j_k}(r_k r_1) d\mu_N(X_1) \cdots d\mu_N(X_m)$$

$$= \sum_{r_1,\cdots,r_k} \sum_{\pi\in PP(k)} \prod_{\substack{(wv) \text{ block of } \pi \\ w<v}} \mathbb{E}[X_{j_w}(r_w r_{w+1}) X_{j_v}(r_v r_{v+1})]. \tag{48}$$

Note that $\prod_{\substack{(wv) \text{ block of } \pi \\ w<v}} \mathbb{E}[X_{j_w}(r_w r_{w+1}) X_{j_v}(r_v r_{v+1})]$ is either zero or $N^{-k/2}$. It is not zero only when $j_w = j_v$ and $r_w r_{w+1} = r_{v+1} r_v$ for all the blocks (v, w) of π. Hence, if we represent q by the star of type q, we see that all the graph where the half-edges of the star are glued pairwise and colorwise will give a contribution. But how many indices will give the same graph ? To represent the indices on the star, we fatten the half-edges as double half-edges. Thinking that each random variable sits at the end of the half-edges, we can associate to each side of the fat half-edge one of the indices of the entry (see Figure 1). When the fattened half-edges meet at the vertex, observe that each side of the fattened half-edges meets one side of an adjacent half-edge on which sits the same index. Hence, we can say that the index stays constant over the broken line made of the union of the two sides of the fattened half-edges.

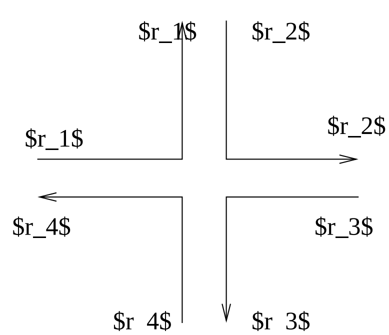


FIGURE 1. Star of type X^4 with prescribed indices

When gluing pairwise the fattened half-edges we see that the condition $r_w r_{w+1} = r_{v+1} r_v$ means that the indices are the same in each side of the half-edge and hence stay constant on the resulting edge. The connected lines made with the sides of the fattened edges can be seen to be the boundaries of the faces of the corresponding graphs. Therefore we have exactly N^F possible choices of indices for a graph with F faces. These graphs are otherwise connected, with one star of type q. (48) thus shows that

$$\begin{aligned}&\int \mathrm{Tr}(q(X_1,\cdots,X_m))d\mu_N(X_1)\cdots d\mu_N(X_m)\\ =\;& \sum_{g\geq 0}\frac{N^F}{N^{\frac{k}{2}}}\sharp\{\text{connected graphs with one star of type } q \text{ and } F \text{ faces}\}\end{aligned}$$

Recalling that $2-2g = F + \sharp$ vertices $-\sharp$ edges $= F+1-k/2$ completes the proof. □

Remark 4.13. *Above it is important to take μ_N to be the law of the GUE (and not GOE for instance) to insure that $E[X(ij)X(ji)] = 1/N$ but $E[X(ij)^2] = 0$. The GOE leads to the enumeration of other combinatorial objects (and in particular an expansion in N^{-1} rather than N^{-2}).*

Proof of Lemma 4.11. We let $q_i(X_1,\cdots,X_m) = X_{\ell_1^i}\cdots X_{\ell_{d_i}^i}$. As usual, we expand the traces;

$$\begin{aligned}&\int\prod_{i=1}^n(N\mathrm{Tr}(q_i(X_1,\cdots,X_m)))d\mu_N(X_1)\cdots d\mu_N(X_m)\\ &= N^n\sum_{\substack{i_1^k,\cdots,i_{d_k}^k\\1\leq k\leq n}}\mathbb{E}[\prod_{1\leq k\leq n}X_{\ell_1^k}(i_1^ki_2^k)\cdots X_{\ell_{d_k}^k}(i_{d_k}^ki_1^k)]\\ &= N^n\sum_{\substack{i_1^k,\cdots,i_{d_k}^k\\1\leq k\leq n}}\sum_{\pi\in PP(\sum d_i)}Z(\pi,\mathbf{i})\end{aligned}$$

where in the last line we used Wick formula, π is a pair partition of the edges $\{(i_j^k,i_{j+1}^k)_{1\leq j\leq d_k-1},(i_{d_k}^k,i_1^k), 1\leq k\leq n\}$ and $Z(\pi,\mathbf{i})$ is the product of the covariances over the corresponding blocks of the partition. A pictorial way to represent this sum over $PP(\sum d_i)$ is to represent $X_{\ell_1^k}(i_1^ki_2^k)\cdots X_{\ell_{d_k}^k}(i_{d_k}^ki_1^k)$ by its associated star of type q_k, for $1\leq k\leq n$. Note that in the counting this star will be labeled (here by the number k). A partition π is represented by a pairwise gluing of the half-edges of the stars. $Z(\pi)$, as the product of the covariance, is null unless each pairwise gluing is done in such a way that the indices written at the end of the glued half-edges coincides and the number of the variable (or color of the half-edges) coincide. Otherwise, each covariance being equal to N^{-1}, $Z(\pi,\mathbf{i}) = N^{-\sum_{i=1}^n k_i/2}$. Note also that once the gluing is done, by construction the indices are fixed on the boundary of each face of the graph (this is due to the fact that $E[X_{ij}X_{kl}]$ is null unless $kl = ji$). Hence, there are exactly N^F possible choices of indices for a given graph, if F is the number of faces of this graph (note here that if the graph is disconnected, we count the number of faces of each connected parts, including their external faces and sum the resulting numbers over all connected components). Thus,

$$\sum_{\substack{i_1^k,\cdots,i_{d_k}^k\\1\leq k\leq n}}\sum_{\pi\in PP(\sum d_i)}Z(\pi,\mathbf{i}) = \sum_{F\geq 0}\sum_{G\in G_F((q_i,1),1\leq i\leq n)}N^{-\sum_{i=1}^n d_i/2}N^F$$

where G_F denotes the union of connected maps with a total number of faces equal to F. Note that for a connected graph, $2-2g = F-\sharp edges - \sharp vertices$. Because the total number of edges of the graphs is $\sharp edges = \sum_{i=1}^n d_i/2$ and the total number of vertices is $\sharp vertices = n$. We see that if $g_i, 1\leq i\leq c$ are the genus of each connected component of our graph, we must have

$$2c-2\sum_{i=1}^c g_i = F-\sum_{i=1}^n d_i/2-n.$$

This completes the proof. □

We then claim that we find (42), namely

Lemma 4.14. *Let $q_1, \cdots, q_n$ be monomials. Then,*

$$\log\left(\int e^{\sum_{i=1}^n t_i N\mathrm{Tr}(q_i(X_1,\cdots,X_m))} d\mu_N(X_1)\cdots d\mu_N(X_m)\right)$$

$$= \sum_{g\geq 0} \frac{1}{N^{2g-2}} \sum_{k_1,\cdots,k_n\in\mathbb{N}} \prod_{i=1}^n \frac{(t_i)^{k_i}}{k_i!} \mathcal{M}_g((q_i,k_i), 1\leq i\leq n) \tag{49}$$

where the equality means that derivatives of all orders at $t_i = 0, 1 \leq i \leq n$, match.

Note here that the sum in the right hand side is not absolutely converging (in fact the left hand side is in general infinite if the t_i's do not have the appropriate signs). However, we shall see in the next sections that if we stop the expansion at $g \leq G < \infty$ (but keep the summation over all k_i's) the expansion is absolutely converging for sufficiently small t_i's.

Note as well that the right hand side of (49) is a generating function of numbers of connected objects; this is a rather well known trick in combinatorics that the logarithm of generating function of disconnected labeled objects gives rise to generating function of connected labeled objects. We do not wish to precise this statement here but the reader will check that the proof below is robust and only based on the relation (51) giving the number of disconnected graphs in terms of their connected subsets.

Proof of Lemma 4.14. The idea is to develop the exponential. Again, this has no meaning in terms of convergent series (and so we do not try to justify uses of Fubini's theorem etc) but can be made rigorous by the fact that we only wish to identify the derivatives at $t = 0$ (and so the formal expansion is only a way to compute these derivatives). So, we find that

$$\begin{aligned} L &:= \int e^{\sum_{i=1}^n t_i N\mathrm{Tr}(q_i(X_1,\cdots,X_m))} d\mu_N(X_1)\cdots d\mu_N(X_m) \\ &= \sum_{k_1,\cdots,k_n\in\mathbb{N}} \frac{(t_1)^{k_1}\cdots(t_n)^{k_n}}{k_1!\cdots k_n!} \int \prod_{i=1}^n (N\mathrm{Tr}(q_i(X_1,\cdots,X_m)))^{k_i} d\mu_N(X_1)\cdots d\mu_N(X_m) \\ &= \sum_{k_1,\cdots,k_n\in\mathbb{N}} \frac{(t_1)^{k_1}\cdots(t_n)^{k_n}}{k_1!\cdots k_n!} \sum_{g\geq 0}\sum_{c\geq 0} \frac{1}{N^{2g-2c}} \sharp\{G_{g,c}((q_i,k_i), 1\leq i\leq n)\} \qquad (50) \end{aligned}$$

where we finally used Lemma 4.11. Note that the case $c = 0$ is non empty only when all the k_i's are null, and the resulting contribution is one. Now, we relate $\sharp\{G_{g,c}((q_i,k_i), 1 \leq i \leq n)\}$ with the number of maps. Since graphs in $G_{g,c}((q_i,k_i), 1 \leq i \leq n)$ can be decomposed into a union of disconnected maps, $\sharp\{G_{g,c}((q_i,k_i), 1 \leq i \leq n)\}$ is related with the ways to distribute the stars and the genus among the c maps, and the number of each of these maps. More precisely, we have (since all stars are labeled)

$$\begin{aligned} &\sharp\{G_{g,c}((q_i,k_i), 1\leq i\leq n)\} \qquad (51) \\ &= \sum_{\substack{\sum_{i=1}^c g_i = g \\ g_i\geq 0}} \frac{g!}{g_1!\cdots g_c!} \sum_{\substack{\sum_{j=1}^c l_i^j = k_i \\ 1\leq j\leq n}} \prod_{i=1}^n \frac{k_i!}{l_i^1!\cdots l_i^c!} \prod_{j=1}^c \mathcal{M}_g((q_i,l_i^j), 1\leq i\leq n). \end{aligned}$$

Plugging this expression into (50) we get

$$\begin{aligned} L &= \sum_{c\geq 0} \frac{1}{c!}\left(\sum_{g\geq 0} \frac{1}{N^{2g-2}} \sum_{l_1,\cdots l_n\geq 0} \prod_{i=1}^n \frac{(t_i)^{l_i}}{l_i!} \mathcal{M}_g((q_i,l_i), 1\leq i\leq n)\right)^c \\ &= \exp\left(\sum_{g\geq 0} \frac{1}{N^{2g-2}} \sum_{l_1,\cdots l_n\geq 0} \prod_{i=1}^n \frac{(t_i)^{l_i}}{l_i!} \mathcal{M}_g((q_i,l_i), 1\leq i\leq n)\right) \end{aligned}$$

which completes the proof. □

The goal of the next sections is to justify that this equality does not only hold formally but as a large N expansion. Instead of using Wick formula, we shall base our analysis on differential calculus and its relations with Gaussian calculus (note here that Wick formula might also have been proved by use of differential calculus). The point here will be that we can design a natural asymptotic framework for differential calculus, that will then encode the combinatorics of the first order term term in 't Hooft expansion, that is planar maps. To make this statement clear, we shall see that a nice set up is when the potential $V = \sum t_i q_i$ possesses some convexity property.

Exercise 4.15.

(1) *We let G_N be a $N \times N$ matrix with independent centered complex Gaussian entries with covariance N^{-1}. We denote G_N^* its adjoint; $G_N^*(ij) = \bar{G}_N(ji)$. So $\mathbb{E}[G_N(ij)G_N(kl)] = 1_{ij=lk}/N$. Show that for any $m, p \in \mathbb{N}$,*

$$\mathbb{E}[\frac{1}{N}\mathrm{Tr}\,(((G^m(G^m)^*)^p)] = \sum_{g\geq 0} \frac{1}{N^{2g}} C(m,p,g)$$

with $C(m,p,g)$ the number of connected graphs that can be embedded into a surface of genus g (or higher) with

- *one vertex with $2mp$ half-edges with two colors (later referred as blue and red) , one distinguished half-edge and one orientation. We label the half-edges by one for the marked edge and proceed following the orientation by labeling by 2 the second half-edge etc The half-edges corresponding to the labels $\{k+2ml, 1 \leq k \leq m, 0 \leq l \leq p-1\}$ will be blue, the other half-edges red.*
- *The half-edges can be glued (or matched) two by two iff they have different colors.*

$\{C(m,p,0), m \geq 0, p \geq 0\}$ are called the Fuss Catalan numbers.

(2) *Take $G_{M,N}$ to be an $N \times M$ matrix with independent equidistributed centered complex Gaussian entries with variance $1/N$. Show that for all integer number p,*

$$\mathbb{E}[\frac{1}{N}\mathrm{Tr}((G_{M,N}G_{M,N}^*)^p)] = \sum_{g\geq 0}\frac{1}{N^{2g}} \sum_{0\leq F_1\leq p+3-2g} (\frac{M}{N})^{F_1} W(F_1,p,g)$$

with $W(f,p,g)$ the number of connected graphs that can be embedded into a surface of genus g (or higher) with

- *one vertex with $2p$ half-edges with two colors (later blue and red) , one distinguished half-edge and one orientation. We label the half edges by one for the marked edge and proceed following the orientation by labeling by 2 the second half-edge etc The odd half-edges will be blue, the other half-edges red.*
- *The half-edges can be glued (or matched) two by two iff they have different colors.*
- *If we draw the half-edge number k as a segment at position $\pi k/p$, we say that a face is gray if it contains the interior of the angle $2\pi(1+2k)/2p, 2\pi(2+2k)/2p$ in the vicinity of the origin, for some $k \leq p-1$. The number of Gray faces of the graph is f.*

Conclude that if M/N converges to some $\alpha \in [0,1]$, the spectral measure of $G_N G_N^*$ converges in moments; i.e

$$m_p(\alpha) = \lim_{N\to\infty} \mathbb{E}[\frac{1}{N}\mathrm{Tr}((G_N G_N^*)^p)]$$

and give a formula for $m_p(\alpha)$

$G_N G_N^*$ is called a Wishart matrix and the limit law with moments $m_p(\alpha)$ is known as the Pastur-Marchenko law.

Show that when $\alpha = 1$, it corresponds to the law of x^2 under σ the semi-circular law,i.e

$$m_p(1) = \int x^{2p} d\sigma(x)$$

3. First order expansion for the free energy

At the end of this lecture (see Theorem 4.24) we will have proved that Lemma 4.14 holds as a first order limit, i.e.

$$\lim_{N\to\infty} \frac{1}{N^2} \log \int e^{\sum_{i=1}^n t_i N \mathrm{Tr}(q_i(X_1,\cdots,X_m))} d\mu_N(X_1)\cdots d\mu_N(X_m)$$

$$= \sum_{k_1,\cdots,k_n \in \mathbb{N}^n \backslash \{0,\cdots,0\}} \prod_{i=1}^{n} \frac{(t_i)^{k_i}}{k_i!} \mathcal{M}_0((q_i,k_i), 1 \le i \le n)$$

provided the parameters t_i's are such that the polynomial $V = \sum t_i q_i$ is 'strictly convex' and 'sufficiently small'. To prove this result we first show that, under the same assumptions, $\bar{\mathbf{L}}_t^N(q) = \mu^N_{\sum t_i q_i}(N^{-1}\mathrm{Tr}(q))$ converges as N goes to infinity towards a limit that is as well related with map enumeration (see Theorem 4.19).

The central tool in our asymptotic analysis will be the so-called Schwinger-Dyson's (or Master loops) equations. They are simple emanation of the integration by parts formula (or, somewhat equivalently, of the symmetry of the Laplacian in $L^2(dx)$). These equations will be shown to pass to the large N limit and be then given as some asymptotic differential equation for the limit points of $\bar{\mathbf{L}}_t^N$. These equations will in turn uniquely determine these limit points in some small range of the parameters and show that the limit points have to be given as some generating function of maps.

3.20. Finite dimensional Schwinger-Dyson's equations

Property 4.16. *For all $P \in \mathbb{C}\langle X_1, \cdots, X_m\rangle$, all $i \in \{1, \cdots, m\}$,*

$$\mu^N_{V_{\mathbf{t}}}\left(\hat{\mathbf{L}}^N \otimes \hat{\mathbf{L}}^N(\partial_i P)\right) = \mu^N_{V_{\mathbf{t}}}\left(\hat{\mathbf{L}}^N((X_i + D_i V_{\mathbf{t}})P)\right) \quad \mathrm{FSD}[\mathrm{V}_{\mathbf{t}}]$$

Proof. A simple integration by part shows that for any differentiable function f on $\mathbb{R}$ such that $f e^{-N\frac{x^2}{2}}$ goes to zero at infinity,

$$N \int f(x) x e^{-N\frac{x^2}{2}} dx = \int f'(x) e^{-N\frac{x^2}{2}} dx.$$

Such a result generalizes to complex Gaussian by the remark that

$$\begin{aligned} N(x+iy)e^{-N\frac{|x|^2}{2} - N\frac{|y|^2}{2}} &= -(\partial_x + i\partial_y)e^{-N\frac{|x|^2}{2} - N\frac{|y|^2}{2}} \\ &= -\partial_{x-iy} e^{-N\frac{|x|^2}{2} - N\frac{|y|^2}{2}}. \end{aligned}$$

As a consequence, applying such a remark to the entries of a Gaussian random matrix, we obtain for any differentiable function f of the entries, all $r,s \in \{1, \cdots, N\}^2$, all $r \in \{1, \cdots, m\}$,

$$N \int A_l(rs) f(A_k(ij), 1 \le i,j \le N, 1 \le k \le m) d\mu_N(A_1) \cdots d\mu_N(A_m) =$$

$$\int \partial_{A_l(sr)} f(A_k(ij), 1 \le i,j \le N, 1 \le k \le m) d\mu_N(A_1) \cdots d\mu_N(A_m).$$

Using repeatedly this equality and section 1.18.2, we arrive at at

$$\begin{aligned} \int \frac{1}{N}\mathrm{Tr}(A_k P) d\mu_V^N(\mathbf{A}) &= \frac{1}{2N^2} \sum_{i,j=1}^{N} \int \partial_{A_k(ji)}(P e^{-N\mathrm{Tr}(V)})_{ji} \prod d\mu_N(A_i) \\ &= \int \left(\frac{1}{N^2}(\mathrm{Tr} \otimes \mathrm{Tr})(\partial_k P) - \frac{1}{N}\mathrm{Tr}(D_k V P)\right) d\mu_V^N(\mathbf{A}) \end{aligned}$$

which yields

$$\int \left(\hat{\mathbf{L}}^N((X_k + D_k V)P) - \hat{\mathbf{L}}^N \otimes \hat{\mathbf{L}}^N(\partial_k P)\right) d\mu_V^N(\mathbf{A}) = 0. \tag{52}$$

□

3.21. Tightness and limiting Schwinger-Dyson's equations

We say that $\tau \in \mathbb{C}\langle X_1, \cdots, X_m\rangle^*$ satisfies the Schwinger-Dyson equation with potential V, denoted in short $\mathbf{SD[V]}$, if and only if for all $i \in \{1, \cdots, m\}$ and $P \in \mathbb{C}\langle X_1, \cdots, X_m\rangle$,

$$\tau(I) = 1, \quad \tau \otimes \tau(\partial_i P) = \tau((D_i V + X_i)P) \qquad \mathbf{SD[V]}.$$

We shall now prove that $\mathbf{SD[V]}$ describes the asymptotic $\mathbf{FSD[V]}$ equations.

Property 4.17. *Assume that $V_{\mathbf{t}}$ is c-convex. Then, $(\bar{\mathbf{L}}_t^N, N \in \mathbb{N})$ is tight. Its limit points satisfy* $\mathbf{SD}[V_{\mathbf{t}}]$ *and*

$$|\tau(X_{\ell_1} \cdots X_{\ell_r})| \le M_0^r \tag{53}$$

for all $\ell_1, \cdots, \ell_r \in \mathbb{N}$, all $r \in \mathbb{N}$, with an M_0 that only depends on c.

Proof. By Lemma 3.21, we find that for all $\ell_1, \cdots, \ell_r$,

$$\begin{aligned} |\bar{\mathbf{L}}_t^N(X_{\ell_1} \cdots X_{\ell_r})| &\le \mu_{V_{\mathbf{t}}}^N(|\lambda_{max}(\mathbf{A})|^r) \\ &= \int_0^\infty r x^{r-1} \mu_{V_{\mathbf{t}}}^N(|\lambda_{max}(\mathbf{A})| \ge x) dx \\ &\le M_0^r + \int_{M_0}^\infty r x^{r-1} e^{-\alpha N x} dx \\ &= M_0^r + (\alpha N)^{-r} \int_0^\infty r x^{r-1} e^{-x} dx \end{aligned} \tag{54}$$

Hence, with the notations of Lemma 4.6, $\bar{\mathbf{L}}_t^N \in K(C)$ with $C(\ell_1, \cdots, \ell_r) = M_0^r + r\alpha^{-r} \int_0^\infty x^{r-1} e^{-x} dx$. $(\bar{\mathbf{L}}_t^N, N \in \mathbb{N})$ is therefore tight. Let us consider now its limit points; let τ be such a limit point. By (54),

$$|\tau(X_{\ell_1} \cdots X_{\ell_r})| \le M_0^r. \tag{55}$$

Moreover, by concentration inequalities (see Lemma 3.24), we find that

$$\lim_{N\to\infty} \left| \int \hat{\mathbf{L}}_{\mathbf{A}}^N \otimes \hat{\mathbf{L}}_{\mathbf{A}}^N(\partial_k P) d\mu_V^N(\mathbf{A}) - \int \hat{\mathbf{L}}_{\mathbf{A}}^N d\mu_V^N(\mathbf{A}) \otimes \int \hat{\mathbf{L}}_{\mathbf{A}}^N d\mu_V^N(\mathbf{A})(\partial_k P) \right| = 0$$

so that Property 4.16 results with

$$\limsup_{N\to\infty} \left| \bar{\mathbf{L}}_t^N\left((X_i + D_i V_{\mathbf{t}})P)\right) - \bar{\mathbf{L}}_t^N \otimes \bar{\mathbf{L}}_t^N(\partial_k P) \right| = 0. \tag{56}$$

Hence, (52) shows that

$$\tau((X_k + D_k V)P) = \tau \otimes \tau(\partial_k P). \tag{57}$$

□

3.21.1. *Uniqueness of the solutions to Schwinger-Dyson's equations for small parameters.* Let $R \in \mathbb{R}^+$ (we will always assume $R \ge 1$ in the sequel). We set

(CS(R)) *An element $\tau \in \mathbb{C}\langle X_1, \cdots, X_m\rangle^*$ satisfies* **(CS(R))** *if and only if for all $k \in \mathbb{N}$,*

$$\max_{1 \le i_1, \cdots, i_k \le m} |\tau(X_{i_1} \cdots X_{i_k})| \le R^k.$$

As we have seen in Property 4.17, the limit points of $\bar{\mathbf{L}}_{\mathbf{t}}^N$ satisfy $(\mathbf{CS}(M_0))$. In the sequel, we denote D the degree of V, that is the maximal degree of the $q_i's$; $q_i(X) = X_{j_1^i} \cdots X_{j_{d_i}^i}$ with, for $1 \le i \le n$, $\deg(q_i) =: d_i \le D$ and equality holds for some i.

The main result of this paragraph is

Theorem 4.18. *For all $R \ge 1$, there exists $\epsilon > 0$ so that for $|\mathbf{t}| = \max_{1 \le i \le n} |t_i| < \epsilon$, there exists at most one solution $\tau_{\mathbf{t}}$ to* $\mathbf{SD}[V_{\mathbf{t}}]$ *that satisfies* $(\mathbf{CS(R)})$.

Proof. Let us assume we have two solutions τ and τ'. Then, by the equation **SD[V]**, for any monomial function P of degree $l-1$, for $i \in \{1, \cdots, m\}$,

$$(\tau - \tau')(X_i P) = ((\tau - \tau') \otimes \tau)(\partial_i P) + (\tau' \otimes (\tau - \tau'))(\partial_i P) - (\tau - \tau')(D_i V P)$$

Hence, if we let for $l \in \mathbb{N}$

$$\Delta_l(\tau, \tau') := \sup_{\text{monomial } P \text{ of degree l}} |\tau(P) - \tau'(P)|$$

we get, since if P is of degree $l-1$, $\partial_i P = \sum_{k=0}^{l-2} p_k^1 \otimes p_{l-2-k}^2$ where p_k^i, $i = 1, 2$ are monomial of degree k or the null monomial, and $D_i V$ is a finite sum of monomials of degree smaller than $D-1$,

$$\begin{aligned}\Delta_l(\tau, \tau') &= \max_{P \text{ of degree } l-1} \max_{1 \le i \le m} \{|\tau(X_i P) - \tau'(X_i P)|\} \\ &\le 2 \sum_{k=0}^{l-2} \Delta_k(\tau, \tau') R^{l-2-k} + C|t| \sum_{p=0}^{D-1} \Delta_{l+p-1}(\tau, \tau')\end{aligned}$$

with a finite constant C (that depends on n only). For $\gamma > 0$, we set $d_\gamma(\tau, \tau') = \sum_{l \ge 0} \gamma^l \Delta_l(\tau, \tau')$. Note that under **(CS(R))**, this sum is finite for $\gamma < (R)^{-1}$. Summing the two sides of the above inequality times γ^l we arrive at

$$d_\gamma(\tau, \tau') \le 2\gamma^2 (1 - \gamma R)^{-1} d_\gamma(\tau, \tau') + C|t| \sum_{p=0}^{D-1} \gamma^{-p+1} d_\gamma(\tau, \tau').$$

We finally conclude that if $(R, |t|)$ are small enough so that we can choose $\gamma \in (0, R^{-1})$ so that $2\gamma^2(1-\gamma R)^{-1} + C|t| \sum_{p=0}^{D-1} \gamma^{-p+1} < 1$ then $d_\gamma(\tau, \tau') = 0$ and so $\tau = \tau'$ and we have at most one solution. Taking $\gamma = (2R)^{-1}$ shows that this is possible provided $\frac{1}{R^2} + C|t| \sum_{p=0}^{D-1} (2R)^{p-1} < 1$ so that when $|t|$ goes to zero, we see that we need R to be at most of order $|t|^{-\frac{1}{D-2}}$. □

3.22. Convergence of the empirical distribution

We are now in position to state the main result of this part;

Theorem 4.19. *For all $c > 0$, there exists $\eta > 0$ and $M_0 \in \mathbb{R}^+$ (given in Lemma 3.21) so that for all $\mathbf{t} \in U_c \cap B_\eta$, $\hat{\mathbf{L}}^N$ (resp. $\bar{\mathbf{L}}_t^N$) converges almost surely (resp. everywhere) towards the unique solution of $\mathbf{SD}[V_\mathbf{t}]$ such that*

$$|\tau(X_{\ell_1} \cdots X_{\ell_r})| \le M_0^r$$

for all choices of $\ell_1, \cdots, \ell_r$.

Proof. By Property 4.17, the limit points of $\bar{\mathbf{L}}_t^N$ satisfy $\mathbf{CS}(M_0)$ and $\mathbf{SD}[V_\mathbf{t}]$. Since M_0 does not depend on $\mathbf{t}$, we can apply Theorem 4.18 to see that if $\mathbf{t}$ is small enough, there is only one such limit point. Thus, by Corollary 4.6 we can conclude that $(\bar{\mathbf{L}}_t^N, N \in \mathbb{N})$ converges towards this limit point. From Lemma 3.24, we have that

$$\mu_V^N(|(\hat{\mathbf{L}}^N - \bar{\mathbf{L}}_t^N)(P)|^2) \le BC(P, M)N^{-2} + C^{2d} N^2 e^{-\alpha M N/2}$$

insuring by Borel-Cantelli's lemma that $\hat{\mathbf{L}}^N$ also converges almost surely to τ. □

Exercise 4.20. *The exercise generalizes the previous approach to unitary matrices following the Haar measure on the unitary group (without interaction to simplify) to recover asymptotic freeness (proved first in* [**101**] *by moments techniques). Let $(A_1^N, \cdots A_m^N)_{N \ge 0}$ be a family of $N \times N$ (eventually random) matrices. Assume that the algebra generated by $(A_1^N, \cdots A_m^N)_{N \ge 0}$ is closed under the involution $*$ and that the operator norm of the A_N^i is bounded independently of N. Finally, suppose that*

$$\lim_{N \to \infty} \hat{\mathbf{L}}_{A_1^N, \cdots, A_m^N}^N = \mu.$$

Let $U_1^N, \cdots, U_m^N$ be m independent unitary matrices, independent of the A_N^i's, following the Haar measure on $\mathcal{U}(N)$. Then $\hat{\mathbf{L}}_{A_i^N, U_N^i, (U_N^i)^{-1}}^N$ given by

$$\hat{\mathbf{L}}_{A_i^N, U_N^i, (U_N^i)^{-1}}^N(P) = \frac{1}{N} \mathrm{Tr}(P(A_i^N, U_N^i, (U_N^i)^{-1}))$$

converges as N goes towards infinity. Moreover the limit τ is described uniquely as

- *τ restricted to polynomials in the A_i's equal μ. For all polynomials and all i*

$$\tau\otimes\tau(\partial_i P)=0$$

with ∂_i the derivative that obeys the Leibniz rule

$$\partial_i(PQ)=\partial_i P\times 1\otimes Q+P\otimes 1\times\partial_i Q$$

and so that

$$\partial_i A_k=0\otimes 0\quad \partial_i U_j=1_{i=j}U_j\otimes 1,\quad \partial_i U_j^*=-1_{j=i}1\otimes U_j^*.$$

Hint : Use the invariance of the Haar measure by multiplication by unitaries (in particular by e^{tB_i} with B_i a matrix with null entries everywhere except at some kl and lk where it is equal to $+1$ and -1 respectively) to prove that $\mathbb{E}\mathrm{Tr}\otimes\mathrm{Tr}[(\partial_i P)]=0$. Then use concentration under the Haar measure to find that the limit points satisfy the above equations. Finally, prove the uniqueness of the solutions to such equations.

- *The law τ is as well described as the law which, once restricted to polynomials in the A_i's, is equal to μ, and once restricted to monomials in the U's, vanishes unless the monomial is a constant, and such that*

$$\tau(P_1(A)Q_1(U)P_2(A)\cdots Q_k(U))=0$$

for all polynomials $P_1,\cdots,P_k$ in the A_i's such that $\mu(P_i)=0$ and all polynomials $Q_1,\cdots,Q_k$ in the U's such that $\tau(Q_i(P))=0$.

One says that the U's and the A's are free under τ when they satisfy the above equality.

Hint: Show that such a law satisfy the previous equations.

3.23. Combinatorial interpretation of the limit

In this part, we are going to identify the unique solution $\tau_{\mathbf{t}}$ of Theorem 4.18 as a generating function for planar maps. Namely, we let for $\mathbf{k}=(k_1,\cdots,k_n)\in\mathbb{N}^n$ and P a monomial in $\mathbb{C}\langle X_1,\cdots,X_m\rangle$,

$$\mathcal{M}_{\mathbf{k}}(P)=\mathrm{card}\{\text{ planar maps with } k_i \text{ labeled stars of type } q_i \text{ for } 1\le i\le n$$
$$\text{and one of type } P\}.$$

This definition extends to $P\in\mathbb{C}\langle X_1,\cdots,X_m\rangle$ by linearity. Then, we shall prove that

Theorem 4.21.

(1) *The family $\{\mathcal{M}_{\mathbf{k}}(P),\mathbf{k}\in\mathbb{C}\langle X_1,\cdots,X_m\rangle,\ P\in\mathbb{C}\langle X_1,\cdots,X_m\rangle\}$ satisfies the induction relation: for all $i\in\{1,\cdots,m\}$, all $P\in\mathbb{C}\langle X_1,\cdots,X_m\rangle$, all $\mathbf{k}\in\mathbb{N}^n$,*

$$\mathcal{M}_{\mathbf{k}}(X_iP)=\sum_{\substack{0\le p_j\le k_j\\1\le j\le n}}\prod_{j=1}^n C_{k_j}^{p_j}\mathcal{M}_{\mathbf{p}}\otimes\mathcal{M}_{\mathbf{k}-\mathbf{p}}(\partial_i P)+\sum_{1\le j\le n}k_j\mathcal{M}_{\mathbf{k}-1_j}([D_iq_j]P)\tag{58}$$

where $1_j(i)=1_{i=j}$ and $\mathcal{M}_{\mathbf{k}}(1)=1_{\overline{k}=0}$. (58) defines uniquely the family $\{\mathcal{M}_{\mathbf{k}}(P),\mathbf{k}\in\mathbb{C}\langle X_1,\cdots,X_m\rangle,P\in\mathbb{C}\langle X_1,\cdots,X_m\rangle\}$.

(2) *There exists A,B finite constants so that for all $\mathbf{k}\in\mathbb{N}^n$, all monomial $P\in\mathbb{C}\langle X_1,\cdots,X_m\rangle$,*

$$|\mathcal{M}_{\mathbf{k}}(P)|\le\mathbf{k}!A^{\sum_{i=1}^n k_i}B^{deg(P)}\tag{59}$$

with $\mathbf{k}!:=\prod_{i=1}^n k_i!$.

(3) *For $\mathbf{t}$ in $B_{(A)^{-1}}$,*

$$\mathcal{M}_{\mathbf{t}}(P)=\sum_{\mathbf{k}\in\mathbb{N}^n}\prod_{i=1}^n\frac{(-t_i)^{k_i}}{k_i!}\mathcal{M}_{\mathbf{k}}(P)$$

is absolutely convergent. For $\mathbf{t}$ small enough, $\mathcal{M}_{\mathbf{t}}$ is the unique solution of $\mathbf{SD}[V_{\mathbf{t}}]$ that satisfies $\mathbf{CS(B)}$.

By Theorem 4.18 and Theorem 4.19, we therefore readily obtain that

Corollary 4.22. *For all $c > 0$, there exists $\eta > 0$ so that for $\mathbf{t} \in U_c \cap B_\eta$, $\hat{\mathbf{L}}^N$ converges almost surely and in expectation towards*

$$\tau_{\mathbf{t}}(P) = \mathcal{M}_{\mathbf{t}}(P) = \sum_{\mathbf{k}\in\mathbb{N}^n} \prod_{i=1}^n \frac{(-t_i)^{k_i}}{k_i!} \mathcal{M}_{\mathbf{k}}(P)$$

Let us remark that by definition of $\hat{\mathbf{L}}^N$, for all P, Q in $\mathbb{C}\langle X_1, \cdots, X_m\rangle$,

$$\hat{\mathbf{L}}^N(PP^*) \geq 0 \quad \hat{\mathbf{L}}^N(PQ) = \hat{\mathbf{L}}^N(QP).$$

These conditions are closed for the weak topology and hence we find that

Corollary 4.23. *There exists $\eta > 0$ ($\eta \geq (4A)^{-1}$) so that for $\mathbf{t} \in B_\eta$, $\mathcal{M}_{\mathbf{t}}$ is a linear form on $\mathbb{C}\langle X_1, \cdots, X_m\rangle$ such that for all P, Q*

$$\mathcal{M}_{\mathbf{t}}(PP^*) \geq 0 \quad \mathcal{M}_{\mathbf{t}}(PQ) = \mathcal{M}_{\mathbf{t}}(QP) \quad \mathcal{M}_{\mathbf{t}}(1) = 1.$$

Remark. This means that $\mathcal{M}_{\mathbf{t}}$ is a tracial state. The traciality property can easily be derived by symmetry properties of the maps. However, I do not know any other way (and in particular any combinatorial way) to prove the positivity property $\mathcal{M}_{\mathbf{t}}(PP^*) \geq 0$, except by using matrix models. This property will be seen to be useful to actually solve the combinatorial problem (i.e. find an explicit formula for $\mathcal{M}_{\mathbf{t}}$).

Proof of Theorem 4.21.

(1) *Proof of the induction relation* (58).

- We first check them for $\mathbf{k} = \mathbf{0} = (0, \cdots, 0)$. By convention, there is one planar map with a single vertex, so $\mathcal{M}_{\mathbf{0}}(1) = 1$. We now check that

$$\mathcal{M}_{\mathbf{0}}(X_iP) = \mathcal{M}_{\mathbf{0}} \otimes \mathcal{M}_{\mathbf{0}}(\partial_i P) = \sum_{P=p_1X_ip_2} \mathcal{M}_{\mathbf{0}}(p_1)\mathcal{M}_{\mathbf{0}}(p_2)$$

But in any planar map with only one star of type X_iP, the half-edge corresponding to X_i has to be glued with another half-edge of P. If X_i is glued with the half-edge X_i coming from the decomposition $P = p_1X_ip_2$, the map is then split into two (independent) planar maps with stars p_1 and p_2 respectively (note here that p_1 and p_2 inherits the structure of stars since they inherit the orientation from P as well as a marked half-edge corresponding to the first neighbour of the glued X_i.)

- We now proceed by induction over $\mathbf{k}$ and the degree of P; we assume that (58) is true for $\sum k_i \leq M$ and all monomials, and for $\sum k_i = M + 1$ when $\deg(P) \leq L$. Note that $\mathcal{M}_{\mathbf{k}}(1) = 0$ for $|\mathbf{k}| \geq 1$ since we can not glue a vertex with no half-edges with any star. Hence, this induction can be started with $L = 0$. Now, consider $R = X_iP$ with P of degree less than L and the set of planar maps with a star of type X_iP and k_j stars of type q_j, $1 \leq j \leq n$, with $|\mathbf{k}| = \sum k_i = M + 1$. Then,
$\diamond$ either the half-edge corresponding to X_i is glued with an half-edge of P, say to the half-edge corresponding to the decomposition $P = p_1X_ip_2$; we then can use the argument as above ; the map M is cut into two disjoint planar maps M_1 (containing the star p_1) and M_2 (resp. p_2), the stars of type q_i being distributed either in one or the other of these two planar maps; there will be $r_i \leq k_i$ stars of type q_i in M_1, the rest in M_2. Since all stars all labeled, there will be $\prod C_{k_i}^{r_i}$ ways to assign these stars in M_1 and M_2.
Hence, the total number of planar maps with a star of type X_iP and k_i stars of type q_i, such that the marked half-edge of X_iP is glued with an half-edge of P is

$$\sum_{P=p_1X_ip_2} \sum_{\substack{0\leq r_i\leq k_i\\ 1\leq i\leq n}} \prod_{i=1}^n C_{k_i}^{r_i} \mathcal{M}_{\mathbf{r}}(p_1)\mathcal{M}_{\mathbf{k}-\mathbf{r}}(p_2) \tag{60}$$

$\diamond$ Or the half-edge corresponding to X_i is glued with an half-edge of another star, say q_j; let's say with the edge coming from the decomposition of q_j into $q_j = q_j^1X_iq_j^2$. Then, once we are giving this gluing of the two edges, we can replace the two stars X_iP and $q_j^1X_iq_j^2$ glued by their X_i by the star $q_j^2q_j^1P$.

We have k_j ways to choose the star of type q_j and the total number of such maps is

$$\sum_{q_j=q_j^1 X_i q_j^2} k_j \mathcal{M}_{\mathbf{k}-1_j}(q_j^2 q_j^1 P)$$

Summing over j, we obtain by linearity of $\mathcal{M}_{\mathbf{k}}$

$$\sum_{j=1}^{n} k_j \mathcal{M}_{\mathbf{k}-1_j}([D_i q_j]P) \tag{61}$$

(60) and (61) give (58). Moreover, it is clear that (58) defines uniquely $\mathcal{M}_{\mathbf{k}}(P)$ by induction.

(2) The proof of (59) now follows by induction from the previous induction relations, see e.g. [**51, 49**]. □

3.24. Convergence of the free energy

Theorem 4.24. *Let $c > 0$. Then, for η small enough, for all $\mathbf{t} \in B_\eta \cap U_c$,*

$$\lim_{N\to\infty} \frac{1}{N^2} \log \frac{Z_N^{V_{\mathbf{t}}}}{Z_N^0} = \sum_{\mathbf{k}\in\mathbb{N}^n\setminus(0,..,0)} \prod_{1\le i\le n} \frac{(-t_i)^{k_i}}{k_i!} \mathcal{M}_{\mathbf{k}}.$$

Moreover, the limit depends analytically on $\mathbf{t}$ *in a neighborhood of the origin.*

Proof. We may assume without loss of generality that $c \in (0,1]$. We then let, for $\alpha \in [0,1]$, $V_{\alpha\mathbf{t}} = \alpha V_{\mathbf{t}}$. $V_{\alpha\mathbf{t}}$ is c-convex. Set

$$F_N(\alpha) = \frac{1}{N^2} \log Z_N^{V_{\alpha\mathbf{t}}}.$$

Then,

$$\frac{1}{N^2} \log \frac{Z_N^{V_{\mathbf{t}}}}{Z_N^0} = F_N(1) - F_N(0) = \int_0^1 \partial_\alpha F_N(\alpha) d\alpha = -\int_0^1 \mu^N_{V_{\alpha\mathbf{t}}}\left(\hat{\mathbf{L}}_N(V_{\mathbf{t}}\right) d\alpha.$$

By Theorem 4.19, we know that for all α (since $V_{\alpha\mathbf{t}}$ is c-convex for all $\alpha \in [0,1]$),

$$\lim_{N\to\infty} \mu^N_{V_{\mathbf{t}}^{\alpha}}(V_{\mathbf{t}}) = \tau_{\alpha\mathbf{t}}(V_{\mathbf{t}})$$

whereas by (54), we know that $\mu^N_{V_{\alpha\mathbf{t}}}\left(\hat{\mathbf{L}}_N(V_{\mathbf{t}}\right)$ stays uniformly bounded. Therefore, a simple use of dominated convergence theorem shows that

$$\lim_{N\to\infty} \frac{1}{N^2} \log \frac{Z_N^{V_{\mathbf{t}}}}{Z_N^0} = -\int_0^1 \tau_{\alpha\mathbf{t}}(V_{\mathbf{t}}) d\alpha = -\sum_{i=1}^{n} t_i \int_0^1 \tau_{\alpha\mathbf{t}}(q_i) d\alpha. \tag{62}$$

Now, observe that by Corollary 4.22,

$$\begin{aligned}
\tau_{\mathbf{t}}(q_i) &= \sum_{\mathbf{k}\in\mathbb{N}^n} \prod_{1\le j\le n} \frac{(-t_j)^{k_j}}{k_j!} \mathcal{M}_{\mathbf{k}+1_i} \\
&= -\partial_{t_i} \sum_{\mathbf{k}\in\mathbb{N}^n\setminus\{0,\cdots,0\}} \prod_{1\le j\le n} \frac{(-t_j)^{k_j}}{k_j!} \mathcal{M}_{\mathbf{k}}
\end{aligned}$$

so that (62) results with

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N^2} \log \frac{Z_N^{V_{\mathbf{t}}}}{Z_N^0} &= -\int_0^1 \partial_\alpha [\sum_{\mathbf{k}\in\mathbb{N}^n\setminus\{0,\cdots,0\}} \prod_{1\le j\le n} \frac{(-\alpha t_j)^{k_j}}{k_j!} \mathcal{M}_{\mathbf{k}}] d\alpha \\
&= -\sum_{\mathbf{k}\in\mathbb{N}^n\setminus\{0,\cdots,0\}} \prod_{1\le j\le n} \frac{(-t_j)^{k_j}}{k_j!} \mathcal{M}_{\mathbf{k}}.
\end{aligned}$$

□

4. Discussion

In the first part, we discuss, following [**26**], how to compute exactly the generating function of quadrangulation from the matrix models approach. In the second part, we underline some open problems.

4.25. Exact computation of the generating function of planar maps

Let us consider the case where $m = 1$ and

$$d\mu_{\mathbf{t}}^N(X) = (Z_{\mathbf{t}}^N)^{-1} e^{-N\mathrm{Tr}(V_{\mathbf{t}}(X))} d\mu^N(X)$$

Assume now that there exists $c > 0$ such that $V_{\mathbf{t}}$ is c-convex. Then, Theorem 4.19 and Corollary 4.22 assert that the limit $\mu_{\mathbf{t}}$ is also a generating function for planar maps;

$$\mu_{\mathbf{t}}(x^p) = \sum_{k\in\mathbb{N}^n} \prod_{i=1}^n \frac{(-t_i)^{k_i}}{k_i!} \mathcal{M}_{\mathbf{k}}(x^p)$$

with $\mathcal{M}_{\mathbf{k}}(x^p)$ the number of planar maps with k_i stars of type x^i and one star of type x^p.

Let us show how to deduce formulae for $\mathcal{M}_{\mathbf{k}}(x^p)$ when $V_{\mathbf{t}}(x) = tx^4$ from the above large deviation result, i.e count quadrangulations and recover the result of Tutte [**100**] from matrix models approach. The analysis below is inspired from [**12**]. As can be guessed, formulae become more complicated as $V_{\mathbf{t}}$ becomes more complex (see [**34**] for a more general treatment)

To find an explicit formula for $\mu_{\mathbf{t}}$ from the Schwinger-Dyson's equation SD$[V_{\mathbf{t}}]$, take $P(x) = (z-x)^{-1}$ to obtain

$$G\mu_{\mathbf{t}}(z)^2 = -4t(\alpha_t + z^2) - 1 + 4tz^3 G\mu_{\mathbf{t}}(z) + zG\mu_{\mathbf{t}}(z)$$

with $G\mu_{\mathbf{t}}(z) = \int (z-x)^{-1} d\mu_{\mathbf{t}}(x)$, $z \in \mathbb{C}\backslash\mathbb{R}$ and $\alpha_t = \int x^2 d\mu_t(x)$. Solving this equation yields

$$G\mu_{\mathbf{t}}(z) = \frac{1}{2}\left(4tz^3 + z - \sqrt{(4tz^3+z)^2 - 4(4t(\alpha_t + z^2)+1)}\right)$$

where we have chosen the solution so that $G\mu_{\mathbf{t}}(z) \approx z^{-1}$ as $|z| \to \infty$. The square root is chosen as the analytic continuation in $\mathbb{C}\backslash\mathbb{R}^-$ of the square root on $\mathbb{R}^+$. Recall that if p_ϵ is the Cauchy law with parameter $\epsilon > 0$, for $x \in \mathbb{R}$,

$$\Im(G\mu_{\mathbf{t}}(x+i\epsilon)) = \int \frac{\epsilon}{(x-y)^2+\epsilon^2} d\mu_{\mathbf{t}}(y) = \pi p_\epsilon * \mu_{\mathbf{t}}(x).$$

Hence, if $\Im(G\mu_{\mathbf{t}}(x+i\epsilon))$ converges as ϵ decreases towards zero, its limit is the density of $\mu_{\mathbf{t}}$. Thus, we in fact have

$$\frac{d\mu_{\mathbf{t}}}{dx} = -\frac{1}{\pi}\lim_{\epsilon\downarrow 0} \Im\left(\sqrt{(4t(x+i\epsilon)^3 + (x+i\epsilon))^2 - 4(4t(\alpha_t + (x+i\epsilon)^2)+1)}\right).$$

To analyse this limit, we write,

$$(4tz^3+z)^2 - 4(4t(\alpha_t+z^2)+1) = (4t)^2(z^2-a_1)(z^2-a_2)(z^2-a_3).$$

for some $a_1, a_2, a_3 \in \mathbb{C}$. Note that since $G\mu_{\mathbf{t}}$ is analytic on $\mathbb{C}\backslash\mathbb{R}$, either we have a double root and a real non negative root, or three real non negative roots. We now argue that when $V_{\mathbf{t}}$ is convex, $a_1 = a_2$ and $a_3 \in \mathbb{R}^+$. In fact, the function

$$f(x) := -2\int \log|x-y| d\mu_{\mathbf{t}}(y) + V_{\mathbf{t}}(x) + \frac{1}{2}x^2$$

is strictly convex on $\mathbb{R}\backslash\mathrm{support}(\mu_{\mathbf{t}})$ and it is continuous at the boundaries of the support of $\mu_{\mathbf{t}}$ as a bounded density. By large deviation analysis and the study of the critical points of the resulting rate function(see e.g. [**10**]), it is well known that $\mu_{\mathbf{t}}$ is such that

$$f(x) = \ell \text{ on the support of } \mu_{\mathbf{t}},\ f(x) \geq \ell \text{ outside}.$$

Thus, we deduce that if there is a hole in the support of $\mu_{\mathbf{t}}$, f must also be constant equal to ℓ on this hole. This contradicts the strict convexity of f outside the support. Hence, we must have $b = a_1 = a_2 \in \mathbb{R}$ and $a = a_3 \in \mathbb{R}^+$. Plugging back this equality give a and b.

Remark Note that the connectivity argument for the support of the optimizing measure is valid for any c-convex potential, $c > 0$. It was shown in [**34**] that the optimal measure has always the form, in the small parameters region,

$$d\mu_{\mathbf{t}}(x) = ch(x)\sqrt{(x-a_1)(a_2-x)}dx$$

with h a polynomial. However, as the degree of $V_{\mathbf{t}}$ grows, the equations for the parameters of h become more and more complex.

4.26. Discussion and open problems

• In this lecture we have shown that

$$\lim_{N\to\infty} \mu_V^N[\hat{\mathbf{L}}_N(P)] = \sigma_0^V(P)$$

Lemma 5.7 when V is a sufficiently small convex potential. This in particular entails that

$$F_V = \lim_{N\to\infty} \frac{1}{N^2} \log \int e^{-N\mathrm{Tr}(V(X_1,\cdots,X_m))} d\mu^N(X_1)\cdots d\mu^N(X_m)$$

exists.

It is a natural question to wonder how much this type of results generalize to non-perturbative situations (that is general potential V).

In the next lecture, we shall rapidly describe the dynamical approach that allows to generalize the convergence and many results such as the analyticity of the limit F_V as a function of the parameters of the potential to the setting of a convex potential V.

In the case of one matrix $m = 1$, the knowledge of the joint law of the eigenvalues given by Theorem 2.1 allows to prove by large deviation techniques (see [**10**]) that F_V exists and is given by

$$F_V = -\inf\{\int V(x)d\mu(x) - \int\int \log|x-y|d\mu(x)d\mu(y)\} + c$$

where the infimum is taken over all probability measure μ on $\mathbb{R}$ and c is some universal constant.

However, it is not known in a more general context whether F_V exists. In [**15**], a related result concerning Voiculescu's free entropy allows only to give lower and upper bounds on the number F_V.

• In a similar spirit, it is not known whether any non-commutative law can be approximated with laws of the form $\mathbf{L}_{A_1,\cdots,A_m}$. Namely, being given $\tau \in \mathbb{C}\langle X_1,\cdots,X_m\rangle^*$, can we find a sequence $A_1^N,\cdots,A_m^N$ of (possibly random) matrices such that for any polynomial P,

$$\lim_{N\to\infty} \mathbf{L}_{A_1^N,\cdots,A_m^N}(P) = \tau(P).$$

In the case $m = 1$, this is the famous Birkhoff theorem that asserts that for any probability measure μ on $\mathbb{R}$, there exists a sequence of real numbers $(\lambda_1^N,\cdots,\lambda_N^N)$ such that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i^N} = \mu.$$

The question is still open when $m \geq 2$ and was posed in his PhD thesis by A. Connes.

• There is no large deviation results for the spectral measure of a Wigner matrix with entries that are not Gaussian.

• In [**56**], large deviations techniques where used to obtain the asymptotics of

$$\lim_{N\to\infty} \frac{1}{N^2} \log \int e^{N\mathrm{Tr}(AUBU^*)} dU$$

when U denotes the Haar measure and A, B two diagonal matrices with converging spectral distribution. Such an asymptotic allows to show that F_V exists for potentials V of two indeterminates when there is at most one term that depends on both indeterminates and it is quadratic (or of the form $X_1^p X_2^q$ for some $p, q \in \mathbb{N}$).

The central tool was to use the Hermitian Brownian motion (see next part) $(A + H^N(t), t \in [0,1])$ and study the process of its spectral distribution. Because the eigenvalues of this process are solution to (67), they can be viewed as a N dimensional Brownian motion with trajectories that are conditioned not to collide. This makes a connection with the results of [**72**] since tilling with rhombi can be seen to be in bijection with random walks that do not intersect.

One can wonder if there would be a general approach to large deviations for the empirical measures of particles under such a conditioning.

• Among the matrix models with several matrices, only very few could be solved even on a physical basis.

Among the models that could be solved, let us cite the Ising model that corresponds to

$$V(X_1, X_2) = \beta X_1 X_2 + g X_1^4 + g X_2^4.$$

We refer here to the work of Mehta [**78**], as well as Boulatov and Kazakov [**22**] and more recently Eynard [**41, 42**](see as well a mathematical review of these results in [**50**] and the large deviation approach in [**48**]). Similar results were found by Bousquet-Mélou and Scheaffer [**23**] by a combinatorial approach. This approach is completely different and based on bijections with trees; it allows to get more precise geometrical information such as the typical diameter of the maps.

However, matrix models approach allowed to study more diverse models so far (such as the q-Potts model for instance). The advantage of this approach is that it allows the use of matrix tricks that may have no counterpart in combinatorics. In particular, it gives the interpretation of the generating function for amps in terms of moments of a probability measure (or a tracial state in the colored setting). The limit so far is that only models where the Cauchy-Stieljes transform of some limiting spectral distribution satisfies some algebraic equation could be solved explicitly.

The next lecture takes a completely different route by trying to get less exact and more qualitative results by constructing the limiting objects by dynamics.

5
Random matrices and dynamics

Processes can be used to obtain non-perturbative results. In [**27, 28, 15, 56**], processes were the key to obtain large deviation estimates for Gaussian matrices. In this lecture, we want to show how they can help to extend the results of the previous lecture, and in particular to weaken the hypothesis on the potential V. This lecture summarizes a recent article with D. Shlyakhtenko [**53**].

In the previous lecture, we have seen that as N goes to infinity, if $V = V_{\mathbf{t}}$ is 'convex and small'.

$$\lim_{N\to\infty} \mu_V^N[\hat{\mathbf{L}}_N(P)] = \tau_V(P) \tag{63}$$

where $\tau_V(:= \tau_{\mathbf{t}})$ is a non commutative law, that is a linear functional on the space of polynomials in m-non commutative variables (see section 1.18.3). τ_V was as well characterized (see section 3.22) as the unique solution to the so-called Schwinger-Dyson equation

$$\tau_V \otimes \tau_V(\partial_i P) = \tau_V((X_i + D_i V)P) \tag{64}$$

satisfying a bound such as $|\tau_V(X_{i_1} \cdots X_{i_p})| \le R^p$ for all $i_j \in \{1, \cdots, m\}$. Then, in section 3.23, $\tau_V(P)$ was identified with the generating function $\mathcal{M}_V(:= \mathcal{M}_{\mathbf{t}})$ for the enumeration of the associated maps because it satisfies the same equation under the same type of constraints.

We provide yet here another type of characterization of the law τ_V as the invariant measure of a stochastic process (and then as a long time limit of this process). This idea is quite reminiscent to Monte Carlo approximation; we shall see in the second part of this lecture that such an approximation has a few nice consequences.

The idea is that for fixed N, μ_V^N is a Gibbs measure that is the invariant measure of dynamics such as Langevin dynamics. It turns out that the process, say $(X_N(t), t \ge 0)$, given by Langevin dynamics converges (in the sense of moments of time marginals) towards a limiting process $(X(t), t \ge 0)$ as N goes to infinity. Note here that $(X_N(t), t \ge 0)$ was denoting a matrix-valued process (in fact each time marginal is a m tuple of $N \times N$ Hermitian matrices) and so the limit $(X(t), t \ge 0)$ is some operator-valued process. Such a process is naturally defined in the free probability setting. In these lecture notes, I will not try to introduce precisely the setting of free probability (the definition of non-commutative laws given in section 1.18.3 being sufficient to our purpose) nor to give detailed proofs but rather invoke the analogy with the classical probability setting. Indeed, arguments are similar, even though the objects are now non-commutative. We hope however this will motivate the reader to learn more about free probability, see [**103, 14, 16, 3, 49**].

Roughly speaking one can hope that, at least for good potentials V, the process $X(t)$ will converge in law as t goes to infinity towards τ_V, so that the following diagram is commutative;

$$\begin{array}{ccc}
\mathcal{L}(X_N(t)) & \rightarrow t \rightarrow \infty \rightarrow & \mu_V^N \\
\downarrow & & \downarrow \\
N \to \infty & & N \to \infty \\
\downarrow & & \downarrow \\
\mathcal{L}(X(t)) & \rightarrow t \rightarrow \infty \rightarrow & \tau_V
\end{array}$$

We shall show that this diagram holds true when $V + \frac{1}{2}\sum X_i^2$ satisfies some convexity property (in fact local convexity is enough, but we shall restrict ourselves to strict convexity in these notes), and then deduce some properties of τ_V from this construction (in particular the fact that any polynomial in $(X_1, \cdots, X_m)$ will have a connected support, a result that generalizes the one dimensional situation that we examined in section 4.25). Thus, processes allow to obtain (63) without the assumption that the potentials are small perturbation of quadratic potentials, provided they stay strictly convex (and even to locally convex potential up to add an appropriate cutoff to μ_V^N). We shall see that they can be used to generalize other results, such as the convergence in operator norms. However, the techniques used here are heavily dependent upon the convexity of the interaction.

1. Free Brownian motions and related stochastic differential calculus

We introduce the free Brownian motion and related stochastic differential calculus as a limit of the Hermitian Brownian motion that is simply a Hermitian matrix with Brownian motions entries.

1.27. Hermitian Brownian motion

We let $\{(B_{ij}(t))_{t\geq 0}, (\tilde{B}_{ij}(t))_{t\geq 0}, i \leq j\}$ be independent Brownian motions, and set $H^N : \mathbb{R}^+ \to \mathcal{H}_N^{(2)}$ to be the Hermitian-matrix valued process given by

$$H^N_{k\ell}(t) = \frac{1}{\sqrt{2N}}\left(B_{k\ell}(t) + i\tilde{B}_{k\ell}(t)\right)$$

for $k < \ell$, $H^N_{k\ell}(t) = \bar{H}^N_{\ell k}(t)$ when $\ell < k$ and $H^N_{kk}(t) = \frac{1}{\sqrt{N}}B_{kk}(t)$ when $k = \ell$.

1.28. Classical Itô's calculus

In this section, we recall the basic results on classical Itô's calculus that we shall need.

Theorem 5.1 (Itô, Kunita-Watanabe)**.** [**68**, pg. 149] *Let $B = \{B_t, \mathcal{F}_t; 0 \leq t < \infty\}$ be a n-dimensional Brownian motion. Let $k : \mathbb{R}^n \to \mathbb{R}^n$ be a bounded Lipschitz function. Then,*

(1) *There exists a unique strong solution X_t to*

$$X_t = X_0 + B_t - \int_0^t k(X_s)ds$$

starting from a given initial condition $X_0 = Z$.

(2) *For any function $f : \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^2$, almost-surely,*

$$f(X_t) = f(X_0) + \int_0^t \nabla f'(X_s).(dB_s - k(X_s)ds) + \frac{1}{2}\int_0^t \Delta f(X_s)ds, \quad 0 \leq t < \infty$$

with $u.v = \sum_{i=1}^n u_i v_i$ and $\Delta f = \sum_{i=1}^n \partial_i^2 f$.

Remark 5.2. *Note that one way to prove the first point is to see $X_t, t \in [0,T]$, as the unique fixed point of $\Phi : C([0,T],\mathbb{R}^n) \to C([0,T],\mathbb{R}^n)$ given by*

$$\Phi_{B,X_0}(X)(t) = X_0 + B_t - \int_0^t k(X_t)dt.$$

This shows in particular that X_t is a continuous function of $\{X_0 + B_s, s \leq t\}$ for any $t \geq 0$.

In relation with the second point we find that

Theorem 5.3. *Let $W : \mathbb{R}^n \to \mathbb{R}$ be a bounded twice continuously differentiable function and consider the unique strong solution to*

$$X_t = X_0 + B_t - \int_0^t \frac{1}{2}\nabla W(X_s)ds.$$

Then, the law

$$d\mu^n(x_1, \cdots, x_n) = \frac{1}{Z_n}e^{-W(x_1,\cdots,x_n)}dx_1 \cdots dx_n$$

is an invariant measure of the process X_t (i.e the law of X_t is μ^n for all $t > 0$ if the law of X_0 is μ^n).

Note that Theorem 5.1 easily shows that $\partial_t E(f(X_t)) = 0$ at $t = 0$ when X_0 has law μ^n, but showing that in fact $\mu^n(f(X_t)) = \mu^n(f(X_0))$ for all $t \geq 0$ requires more thoughts around the domain of the generator (see the Hille-Yoshida Theorem).

1.29. Itô's calculus and Random matrices

As we said earlier, we would like to construct dynamics whose invariant measure is given by the matrix model of the previous part;

$$d\mu_V^N(X_1, \cdots, X_m) = \frac{1}{Z_V^N}e^{-N\mathrm{Tr}(V(X_1,\cdots,X_m))}d\mu_N(X_1)\cdots d\mu_N(X_m)$$

for a self-adjoint potential V. Recall that $d\mu_N(X_m)$ is just the Gaussian law on the entries and so is absolutely continuous with respect to Lebesgue measure and so we can construct, if V is sufficiently smooth, invariant dynamics as in Theorem 5.3, i.e dynamics on each entries of the random matrices $X_1, \cdots, X_m \in \mathcal{H}^{(N)}$ with invariant measure μ_V^N.

Recall (see paragraph 1.18.2) that for each $i \in \{1, \cdots, m\}$ and each $k, \ell \in \{1, \cdots, N\}$,

$$\partial_{X^i_{k\ell}} \mathrm{Tr}(P(X_1, \cdots, X_m)) = (D_i P)_{k\ell}.$$

Hence, we can write the dynamics of Theorem 5.3 in a matricial way as

$$dX_t^{N,i} = dH_t^{N,i} - \frac{1}{2}(D_i V(X_t^{N,1}, \cdots, X_t^{N,m}) + X_t^{N,i})dt \tag{65}$$

with $H_t^{N,i}, 1 \leq i \leq m$ m independent Hermitian Brownian motions. We denote in short $W = V + \frac{1}{2}\sum_{i=1}^m X_i^2$.

Lemma 5.4. *Take V to be a polynomial, denote $W = V + \frac{1}{2}\sum X_i^2$ and assume that*

$$\phi : X_1, \cdots, X_m \in (\mathcal{H}_N^{(2)})^m \to \mathrm{Tr}(W)$$

is strictly convex (i.e the Hessian of ϕ is bounded below by cI for some $c > 0$). Then, for any integer number N,

1. *There exists a unique strong solution $X^N_.$ to (65) for all times. Moreover, $X_t^{N,i}$ is Hermitian for all $i \in \{1, \cdots, m\}$ and $t \geq 0$.*
2. *μ_V^N is an invariant measure for this process.*
3. *The law of $X_t^N = (X_t^{N,i}, 1 \leq i \leq m)$ converges towards μ_W^N as t goes to infinity, independently of the initial data.*

Proof. The first part of the lemma is a direct consequence of the previous section except for the fact that V, as a polynomial, is a priori unbounded (as well as its derivatives). Convexity however insures the existence and uniqueness of a solution, together with the fact that the solution will remain finite almost surely. Moreover, since we assume $\mathrm{Tr}V$ convex, it is in particular real valued and therefore $\mathrm{Tr}V = \mathrm{Tr}V^*$. Differentiating both sides of this equality shows that (see (46)) $(D_i V)^* = D_i V^* = D_i V$for all $i \in \{1, \cdots, m\}$. Therefore, taking the adjoint of both sides of (65) we find that $(X^N_.)^*$ is as well solution of (46) and thus $X^N = (X^N_.)^*$ by uniqueness.

Moreover, μ_V^N is an invariant measure by Theorem 5.3. To prove the convergence in law of X_t^N we simply take two solutions X_t^N and $\tilde{X}_t^N$ starting from two initial data X_0^N and $\tilde{X}_0^N$ and constructed with the same Hermitian Brownian motion. Then, we get that

$$d(X_t^{N,i} - \tilde{X}_t^{N,i}) = \frac{1}{2}\left(D_i W(X_t^N) - D_i W(\tilde{X}_t^N)\right) dt.$$

Therefore, because by convexity

$$\sum_{i=1}^m (X_t^{N,i} - \tilde{X}_t^{N,i}).\left(D_i W(X_t^N) - D_i W(\tilde{X}_t^N)\right) \geq c \sum_{i=1}^m (X_t^{N,i} - \tilde{X}_t^{N,i})^2$$

in the sense of self-adjoint operators (here $X.Y = XY^* + YX^*$), we find that

$$d\sum_{i=1}^m (X_t^{N,i} - \tilde{X}_t^{N,i})^2 \leq -c\sum_{i=1}^m (X_t^{N,i} - \tilde{X}_t^{N,i})^2 dt$$

and therefore

$$\sum_{i=1}^m (X_t^{N,i} - \tilde{X}_t^{N,i})^2 \leq e^{-ct}\sum_{i=1}^m (X_0^{N,i} - \tilde{X}_0^{N,i})^2.$$

Thus, X_t^N will not depend much of its initial law provided it has a finite operator norm, that is finite entries. We can now take $\tilde{X}_0^N$ with law μ_V^N (recall that by Brascamp Lieb inequalities Lemma 3.20 the operator norm of $\tilde{X}_0^N$ is then well controlled) to deduce that the law of the entries of $(X_t^{N,i}, 1 \leq i \leq m)$ will converge to μ_V^N, whatever is the initial condition with finite operator norm. □

We have as well the following application of Itô's calculus.

Lemma 5.5. *Let P be a polynomial in m non-commutative variables. Let X^N be the process considered in Lemma 5.4 and denote $W = V + \frac{1}{2}\sum_{i=1}^m X_i^2$. Then,*

$$P(X_t^N) = P(X_0^N) + \int_0^t \sum_{i=1}^m \partial_i P(X_t^N)\sharp(dH_t^{N,i} - \frac{1}{2}D_iW(X_t^N)dt) + \frac{1}{2}\int_0^t \mathbb{L}^N P(X_t^N)dt$$

with $\mathbb{L}^N = \sum_{i=1}^m I \otimes (\frac{1}{N}\mathrm{Tr})\Delta^{N,i}$ *if*

$$\Delta^{N,i} = M \circ \partial_i \circ \partial_i$$

where $M(A \otimes B \otimes C) = AC \otimes B$*. As a consequence,*

$$\frac{1}{N}\mathrm{Tr}\left(P(X_t^N)\right) = \frac{1}{N}\mathrm{Tr}\left(P(X_0^N)\right) + \int_0^t \sum_{i=1}^m \frac{1}{N}\mathrm{Tr}(D_iP(X_t^N)(dH_t^{N,i} - \frac{1}{2}D_iW(X_t^N)dt)$$

$$+\frac{1}{2}\int_0^t \mathcal{L}^N P(X_t^N)dt \tag{66}$$

with $\mathcal{L}^N = \sum_{i=1}^m (\frac{1}{N}\mathrm{Tr}) \otimes (\frac{1}{N}\mathrm{Tr})\partial_i \circ D_i$*.*

The proof of the first point is a direct consequence of Theorem 5.1 with $f(X_t^N) = P(X_t^N)_{k\ell}$ for all $k, \ell \in \{1, \cdots, N\}$ and is left to the reader. The second equality is obtained by taking the trace of the first.

Note that when $m = 1$, the previous Lemma can be simplified by diagonalizing the matrices $X_t^N, t \geq 0$. In that case, if we put $\mathbf{L}_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i^N(t)}$ to be the spectral measure of X_t^N, we have $\frac{1}{N}\mathrm{Tr}\left(P(X_t^N)\right) = \int P(x)d\mathbf{L}_t^N(x)$. Another way to derive Lemma 5.5 is then to use the representation of the dynamics of the eigenvalues $(\lambda_i^N(t), 1 \leq i \leq N)$ as solution of a stochastic differential system. In fact, Dyson [**38**] showed that

$$d\lambda_N^i(t) = \frac{1}{\sqrt{2N}}dW_t^i + \frac{1}{N}\sum_{j\neq i}\frac{1}{\lambda_N^i(t) - \lambda_N^j(t)} - \frac{1}{2}W'(\lambda_N^i(t))dt \tag{67}$$

with initial condition $\lambda_N(0)$, the eigenvalues of Z and $(W^1, \cdots, W^N)$ a N-dimensional Brownian motion.

Exercise 5.6. *We leave the reader to check that Itô's calculus implies that* $M_t^P := \int P(x)d\mathbf{L}_t^N - \int P(x)d\mathbf{L}_0^N - \int_0^t \int_{x\neq y}(\frac{P'(x)-P'(y)}{2(x-y)})d\mathbf{L}_s^N(x)d\mathbf{L}_s^N(y)ds$ *is a local martingale (as in* (66)*).*

1.30. The free Brownian motion

1.30.1. *Law of the free Brownian motion.* The free Brownian motion can be thought as the large N limit of the Hermitian Brownian motion, in the sense of weak convergence. In fact, by Theorem 1.17, any moment of the increments of an Hermitian Brownian motion converges when N goes to infinity; for any $t_1 < \cdots < t_m$, any polynomial function P of m non-commutative variables,

$$\lim_{N\to\infty}\mathbb{E}[\frac{1}{N}\mathrm{Tr}(P(H_{t_1}^N, H_{t_2}^N - H_{t_1}^N, \cdots, H_{t_m}^N - H_{t_{m-1}}^N))]$$
$$= \sigma_m[P(\sqrt{t_1}X_1, \sqrt{t_2 - t_1}X_2, \cdots, \sqrt{t_m - t_{m-1}}X_m))].$$

As in the classical setting, we can also gives sense of σ_m with $m = \infty$ (i.e construct the joint law of an infinite number of free semi-circular) and then construct the law ϕ of a continuous process $(S_t, t \geq 0)$, such that for all $m \in \mathbb{N}$, all polynomials P, all times $t_1 < t_2 < \cdots < t_m$,

$$\phi(P(S_{t_1}, S_{t_2} - S_{t_1}, \cdots, S_{t_m} - S_{t_{m-1}})) = \sigma_m[P(\sqrt{t_1}X_1, \sqrt{t_2 - t_1}X_2, \cdots, \sqrt{t_m - t_{m-1}}X_m)].$$

ϕ is here seen as a linear form on the set $\mathcal{C}$ of cylinder functions of the form

$$F(S) = P(S_{t_1}, S_{t_2}, \cdots, S_{t_m})$$

for some choice of $t_1 < \cdots < t_m$, $m \in \mathbb{N}$ and polynomials P.

To give a sense to the continuity of the process, we need to have some notion of positivity and norms. To this end, note that as the σ_m's, ϕ is a tracial state in the sense that

$$\phi(FF^*) \geq 0,\ \phi(FG) = \phi(GF) \text{ and } \phi(1) = 1$$

for any $F, G \in \mathcal{C}$. Here $*$ is the involution as defined in (45) ($(zS_{t_1}\cdots S_{t_m})^* = \bar{z}S_{t_m}\cdots S_{t_1}$ for all choices of $(t_1, \cdots, t_m)$ and all $m \in \mathbb{N}$).

The positivity property $\phi(FF^*) \geq 0$ allows to think of ϕ as an expectation, and to endow $\mathcal{C}$ with useful notions such as a semi-norm or a partial order. Indeed,

$$\|F\|_\infty := \sup_{n\geq 0} \phi((FF^*)^n)^{\frac{1}{2n}}$$

defines a semi-norm on the vector space $\mathcal{C}$ and ϕ is continuous with respect to this semi-norm ($|\phi(P)| \leq \|P\|_\infty$ for all $P \in \mathcal{C}$) Moreover, we can define an order on $\mathcal{C}$ by $P \geq 0$ iff P is self-adjoint, $P = P^*$, and $\phi(QPQ^*) \geq 0$ for any $Q \in \mathcal{C}$.

The continuity of $(S_t, t \geq 0)$ is insured by the fact (deduced from the definition of σ_m) that for all $s < t$, all $m \in \mathbb{N}$,

$$\phi[(S_t - S_s)^{2m}] \leq (4(t-s))^m$$

which insures that $\|S_t - S_s\|_\infty \leq 2\sqrt{|t-s|}$.

1.30.2. *Realization of the free Brownian motion.* In the previous subsection, we described the law of the free Brownian motion. However, as probabilists, we like to think about random variables. Usually, random variables are thought as measurable functions X from a probability space (Ω, P) into the space of values of X, say $\mathbb{R}$. If X is bounded, X is just some element of $L^\infty(\Omega, P)$.

A similar concern exists in free probability and the generalization goes as follows. The idea is to think of $L^\infty(\Omega, P)$ as a space of functions that acts by left multiplication on $L^2(\Omega, P)$; i.e we identify the function $f \in L^\infty(\Omega, P)$ with the operator $\pi_f : L^2(\Omega, P) \to L^2(\Omega, P)$ that associates to $g \in L^2(\Omega, P)$, $\pi_f(g) = fg$. The interests in $L^2(\Omega, P)$ is that it can be equipped with a scalar product $\langle g, h\rangle := \int ghdP$ and we can construct a Hilbert space H by separating $L^2(\Omega, P)$ (i.e. by taking the quotient of $L^2(\Omega, P)$ by the ideal $\{g :< g, g >= 0\}$) so that $\|g\|_2 := \langle g, g\rangle^{\frac{1}{2}}$ is a norm on H. Then, π_f can be seen as an element of the space $B(H)$ of bounded operators on H for the norm

$$\|\pi_f(g)\|_\infty = \sup_{g\in H} \frac{\|\pi_f(g)\|_2}{\|g\|_2}.$$

The same construction can be generalized to the non-commutative setting; P is then replaced by the tracial state ϕ, and the scalar product by $\langle P, Q\rangle := \phi(PQ^*)$ (note here that the positivity of ϕ is crucial). A Hilbert space H is obtained by completing and separating $L^2(\phi)$ (the closure of polynomials by the norm induced by $\langle *, *\rangle$) and random variables are interpreted as elements of the space $B(H)$ of bounded operators on H (in fact as left multiplication operators as above). ϕ is then a linear form on $B(H)$ and the operator norm on $B(H)$ is nothing but

$$\|P\|_\infty := \sup_{n\geq 0} \phi((PP^*)^n)^{\frac{1}{2n}}.$$

This is the so-called Gelfand-Neimark-Segal construction (see [**83**] for details)

Thus, $(S_t, t \geq 0)$ can be thought as a continuous (for $\|\cdot\|_\infty$) process with values in $B(H)$, the space of bounded operators on a Hilbert space H.

1.30.3. *Free Brownian motion and freeness.* $(S_t, t \geq 0)$ is called the free Brownian motion. The word freeness is used to say that increments are not independent as for the classical Brownian motion but free in the following sense. Freeness means (see also exercise 4.20) that if $X = S_t - S_s$, for any polynomials $P_1, \cdots, P_k$ and any elements $A_1, \cdots, A_k$ in the algebra generated by $(S_u, u < s)$ (or equivalently the cylinder functions that only depends on $(S_u, u < s)$),

$$\phi((P_1(X) - \phi(P_1(X)))(A_1 - \phi(A_1))(P_2(X) - \phi(P_2(X))) \cdots (A_k - \phi(A_k))) = 0.$$

Note here that this relation determines uniquely the joint law of $(X, (S_u, u < s))$ from the moments of X and $(S_u, u < s)$ respectively; in other words, as in the classical case, the law of free (resp. independent) variables is uniquely determined by its marginals.

Therefore, we shall thereafter consider ϕ as a linear form on some bigger space than $\mathcal{C}$. In tune with the previous section, we shall denote this space $B(H)$.

As a remark, note that the word 'freeness' emphasizes the relation with the usual notion of freeness in groups; indeed, if ϕ, evaluated at a monomial, is one if the monomial is the neutral element and zero otherwise, we see that the above relation exactly means that non trivial words in X and the A_i's can not be the neutral element; i.e they are free in the usual sense.

1.31. Free stochastic calculus

Free stochastic integrals with respect to the Free Brownian motion can be build exactly as in the classical case; we consider for a continuous adapted (i.e Y_t depends only on $S_s, s < t$ for all t) process $Y : \mathbb{R}^+ \to B(H)$ the sums $I(t) = \sum_{i=1}^n Y_{t_{i-1}}(S_{t_i} - S_{t_{i-1}})$ and prove that they converge in $L^2(\phi)$, when the sequence $t = (0 = t_0 < \cdots < t_n = t)$'s is such that $\sup |t_i - t_{i-1}|$ goes to zero. Moreover, the limit does not depend on the choice of the t_i's. The limit is then denoted $\int_0^t Y_s dS_s$. It shares many properties with its classical analogues. In particular it is a martingale with respect to the filtration of the free Brownian motion.

A property, that is in fact specific to the non-commutative setting, is that a Burkhölder-Davis inequality for integrals with respect to free Brownian motion holds for the L^p norm even with $p = \infty$(see Theorem 3.2.1 of [**89**]). More precisely, the following estimate holds for any adapted process Y_t

$$\left\| \sum_{i=1}^m \int_0^s Y_t^i dS_t^i \right\|_\infty \le 2\sqrt{2} \left(\int_0^s \| \sum_{i=1}^m (Y_t^i)^2 \|_\infty dt \right)^{\frac{1}{2}}. \tag{68}$$

This result is related with the fact that S_t is uniformly bounded (on the contrary to the standard Brownian motion).

1.32. Stochastic differential calculus

The same construction holds to construct m free Brownian motions $S_t^i, t \ge 0, 1 \le i \le m$ as the weak limit of $H_t^{N,i}, t \ge 0, 1 \le i \le m$, where $H^{N,i}, 1 \le i \le m$ are m independent Hermitian Brownian motions.

Since $(S_t, t \ge 0)$ can be thought as the limit of the Hermitian Brownian motion and, by construction, solutions to stochastic differential equations are continuous functions of the Hermitian Brownian motion, it is no surprise that the theory of stochastic differential equations passes to the large N limit. Let us state the results;

Lemma 5.7. *Let $W = V + \frac{1}{2}\sum X_i^2$ be a strictly convex self-adjoint polynomial, i.e there exists $c > 0$ such that for any m-tuples $X = (X_1, \cdots, X_m)$ and $Y = (Y_1, \cdots, Y_m)$ of self-adjoint operators in $\mathcal{C}$*

$$\sum_{i=1}^m (D_i W(X) - D_i W(Y)).(X_i - Y_i) \ge c \sum_{i=1}^m (X_i - Y_i)^2.$$

Then,

(1) *Let X_0 be a m-tuple of self-adjoint variables in $B(H)$. There exists a unique solution to*

$$X_t^i = X_0^i + S_t^i - \frac{1}{2}\int_0^t D_i W(X_s) ds, 1 \le i \le m$$

Moreover for each $t \ge 0$, X_t is a m-tuple of self-adjoint operators.

(2) *There exists $B(c) \le M(c) < \infty$ such that, for any X_0,*

$$\limsup_{t\to\infty} \|X_t^i\|_\infty \le B(c), \quad \sup_{t \ge 0} \|X_t\|_\infty \le \max\{M(c), \|X_0\|_\infty\}$$

where $\|\cdot\|_\infty$ is the operator norm given by

$$\|X\|_\infty = \sup_{n \ge 0} \phi((X)^{2n})^{\frac{1}{2n}}.$$

(3) *The following analogue of Itô's calculus holds for any polynomial of m non-commutative variables*

$$\phi(P(X_t)) = \phi(P(X_0)) + \frac{1}{2}\int_0^t \sum_{i=1}^m (\phi \otimes \phi(\partial_i \circ D_i P(X_s)) - \phi(D_i P(X_s) D_i W(X_s))) ds.$$

(4) *The distribution of X_t converges weakly as $t \ge 0$, independently of $X_0 \in B(H)$. The limit τ is invariant and satisfies for any polynomial P*

$$\tau \otimes \tau(\sum_{i=1}^m \partial_i \circ D_i P) = \tau(\sum_i D_i P.D_i W). \tag{69}$$

Reciprocally, any non-commutative measure (i.e. linear form on $\mathbb{C}\langle X_1, \cdots, X_m\rangle$ such that $|\tau(X_{i_1}\cdots X_{i_\ell})| \le R^\ell$ for some R and all ℓ) satisfying this equation for all polynomial P is invariant. Hence, there exists a unique such measure. There exists a unique solution τ_V to the Schwinger-Dyson equation (64); $\tau_V = \tau$.

The proof of the first point copies the classical proof; during a short time before the X_s eventually explode, $D_i W$ is uniformly Lipschitz (for the operator norm) and so the usual Picard type argument applies to construct the solution (i.e we let $\phi(S, X)_t = X_0 + S_t - \frac{1}{2}\int_0^t D_i W(X_s)ds$ and show that this application is a contraction (at least on processes bounded by some threshold)). We then show that its unique fixed point is bounded uniformly because of (68) in the following lines; we let $Y.X = \sum(X^i Y^i + X^i Y^i)$ and find

$$\begin{aligned} dX_t.X_t &= -\frac{1}{2}(DW(X_t) - DW(0)).(X_t)dt - \frac{1}{2}DW(0).X_t dt + X_t.dS_t \\ &\le -\frac{c}{4}X_t.X_t dt + \frac{1}{2c}DW(0).DW(0)dt + X_t.dS_t \end{aligned}$$

Thus, we find that $A_t = \|X_t.X_t\|_\infty$ is bounded by

$$\begin{aligned} A_t &\le e^{-\frac{t}{2c}}A_0 + \frac{1}{2c}DW(0).DW(0) + e^{-\frac{t}{2c}}\|\int_0^t e^{\frac{u}{2c}}X_u.dS_u\|_\infty \\ &\le e^{-\frac{t}{2c}}A_0 + \frac{1}{2c}DW(0).DW(0) + e^{-\frac{t}{2c}}2\sqrt{2}\left(\int_0^t e^{\frac{u}{c}}A_u du\right)^{\frac{1}{2}} \end{aligned}$$

from which it is easy to deduce that A_t stays uniformly bounded by some constant that only depends on c and $DW(0).DW(0)$. This allows to show the second point of the theorem and thus complete the proof of the first. Itô's calculus formula can be derived as in the classical case or, at least intuitively, as a large N limit of (66). By the condition that a stationary distribution satisfies $\partial_t\phi(P(X_t))|_{t=0} = 0$ we deduce the fourth point. We can finally copy the proof of Lemma 5.4 to obtain the uniform-time convergence of the process, and obtain

$$\|X_t^Z - X_t^{Z'}\|_\infty \le e^{-ct}\|Z - Z'\|_\infty \tag{70}$$

when X_t^Z starts from the initial condition Z. This entails the uniqueness of the laws satisfying (69) and therefore uniqueness of the solution to Schwinger-Dyson equation with potential W (that imposes stronger conditions).

Remark 5.8. *Observe that because of the Burkholder Davis Gundy inequality in L^∞, the process stays uniformly bounded by $M(c)$; hence, as long as we do not start our process with initial data's says larger than $2B_0(c)$, we need only to assume that V is strictly convex when applied to operators bounded by $M(c)$. This allows to weaken our hypothesis to a 'locally convex hypothesis' and thus include any small perturbation of a quadratic potential.*

2. Consequences

In this section we summarize a few applications of Lemma 5.7 valid in the case of a convex potential; first we study the resulting properties of the invariant measure $\tau = \tau_V$ of Lemma 5.7, then the large N limit of $\mu_V^N(\hat{\mathbf{L}}_N)$ and finally the analyticity of τ_V as a function of the parameters of the potential V, i.e we discuss the absence of phase transition in the domain of convexity of the potential.

2.33. Approximation by a continuous function of the free Brownian motion

Taking Z with law τ the invariant measure and $Z' = 0$, we deduce from (70) that since X_t^Z has law τ as well, we can approximate the law τ by X_t^0, with t large. We now claim that

Lemma 5.9. *For any $t > 0$, there exists a sequence $\psi_{n,t}$ of continuous functions (for the uniform operator norm $\|Y.\|_\infty^t = \sup_{s\le t}\|Y_s\|_\infty$) such that*

$$\|\psi_{n,t}(S) - X_t^0\|_\infty \le \frac{1}{n}.$$

Therefore, for any $\epsilon > 0$, there exists a continuous (for $\|\cdot\|_\infty^\infty$) function ψ_ϵ of the free Brownian motion, there exists Z with law τ, such that

$$\|Z - \psi_\epsilon(S)\| \le \epsilon$$

Proof. Indeed, as in Remark 5.2, we can see $(X_s, s \le T)$, the solution of the free SDE as in Lemma 5.7, as the unique fixed point of

$$\Phi_S(X)(t) = S_t - \int_0^t \frac{1}{2} DW(X_t) dt$$

with $k = \frac{1}{2}DW$ on variables bounded by some M larger than the uniform bound $M(c)$ on X, but extended outside of this set so that k is uniformly Lipschitz. Therefore, we can approximate $X(t)$ by $\Phi_S^n(0)(t)$ uniformly on $t \le T$. $\Phi_S^n(0)(t)$, as integrals of polynomials in S, is a continuous function of $S_s, s \le T$ (because the later are uniformly bounded; $\|S_.\|_\infty^u \le 2\sqrt{u}$). □

This result has at least two nice consequences. The first is really due to a non-commutative phenomenon; there is no operator that is a continuous function of the free Brownian motion that can have a disconnected spectrum. Thus we have

Lemma 5.10. *Assume that $W = V + \frac{1}{2}\sum X_i^2$ is strictly convex and let $(X_1, \cdots, X_m)$ be non-commutative variables with law τ_V. Then, for any polynomial function P, the operator $P(X_1, \cdots, X_m)$ in $B(H)$ has a connected spectrum.*

Proof. Let us summarize the arguments for the motivated reader, even though they are less classical in probability. In fact, the algebra of operators constructed as polynomials of the free Brownian motion, or as approximations of such operators for the operator norm, is projectionless, i.e we cannot build a non trivial projection with such elements (that is P such that $P^2 = P$ and $\phi(P) \in (0,1)$) (see [**85**] (see also [**58**] for a random matrix proof)). This implies in particular that the spectrum of any self-adjoint element A of this family (defined as the support of the probability measure given by $\mu_A(P) = \phi(P(A))$ for any polynomial P) has to be connected since otherwise we can build a non trivial projection (basically the projection on the eigenspace of one connected component). Thus, for any $\epsilon > 0$, ψ_ϵ of Lemma 5.9 must have a continuous spectrum. This makes impossible that Z has a disconnected support since if its spectrum had a hole of width δ, ψ_ϵ should also have a hole in its spectrum for $\epsilon < 1/4\delta$. □

Remark 5.11. *Note that the proof above is very different from that given in section 4.25 that was based on the statement of a large deviation principle and the analysis of the resulting rate function. Such an approach is not yet available in the multi-matrix case, as stressed in section 4.26.*

The second application concerns the extension of a result of Haagerup and Thorbjornsen [**59**] where it was proved that if $X_1^N, \cdots, X_m^N$ are m independent matrices following the **GUE**,

$$\lim_{N\to\infty} \|P(X_1^N, \ldots, X_m^N)\|_\infty = \|P(X_1, \ldots, X_m)\|_\infty \text{ a.s.}$$

with $(X_1, \cdots, X_m)$ m free semi-circular variables. We can extend this result to the case where the m-tuple $(X_1^N, \cdots, X_m^N)$ has law μ_V^N (and $(X_1, \cdots, X_m)$ has law τ) by writing that $(X_1^N, \cdots, X_m^N)$ and $(X_1, \cdots, X_m)$ are approximately some nice polynomials in the increments of the Hermitian Brownian motions (respectively the free Brownian motion) and use at this level Haagerup and Thorbjornsen result [**59**]. We thus prove that

Lemma 5.12 ([**53**], Lemma 6.1)**.** *Let $(X_1^N, \ldots, X_m^N)$ be a m-tuple of matrices with law μ_V^N and assume that $V + \frac{1}{2}\sum X_i^2$ is strictly convex. Let τ_V be the unique invariant measure τ_V of Lemma 5.7 and $(X_1, \ldots, X_m)$ with law τ_V. Then, for any polynomial function P,*

$$\lim_{N\to\infty} \|P(X_1^N, \ldots, X_m^N)\|_\infty = \|P(X_1, \ldots, X_m)\|_\infty \ a.s.$$

2.34. Convergence of the matrix model

We can generalize Theorem 4.19 concerning the convergence of the empirical distribution of matrices following μ_V^N as well as Theorem 4.24 to any convex potential V. Indeed, by Lemma 4.17, all limit points of $\mu_V^N(\hat{\mathbf{L}}_N)$ satisfy the Schwinger-Dyson equation and are bounded laws. Therefore, Lemma 5.7 insuring the uniqueness of such solutions, shows that $\mu_V^N(\hat{\mathbf{L}}_N)$ has a unique limit point and thus converges toward τ. The result for the free energy is then a trivial consequence. Thus, we have for any convex potential V

$$\lim_{N\to\infty} \frac{1}{N^2} \log \frac{Z_N^V}{Z_N^0} = F_0(V).$$

2.35. Analyticity of the limit as a function of the parameters of the potential
A last consequence of Lemma 5.7 is to see that, for any polynomial P, $\tau_V(P)$ depends analytically upon the parameters of V within the set where V stays convex. More precisely

Lemma 5.13. *Let $V = V_\beta = \sum_{i=1}^n \beta_i q_i$ be a polynomial, where $\beta = (\beta_i)_{1\le i\le n}$ are (complex) parameters and $(q_i)_{1\le i\le n}$ are monomials. For $c > 0$ let $T(c) \subset \mathbb{C}^n$ be the interior of the subset of parameters $\beta = (\beta_i)_{1\le i\le n}$ for which V is c-convex. Let $\tau_\beta = \tau_{V_\beta}$ be the unique stationary measure of Lemma 5.7. Then for any polynomial $P \in \mathbb{C}\langle X_1, \cdots, X_m\rangle$, the map $\beta \in T(c) \to \tau_\beta(P)$ is analytic. As a consequence, $\beta \in T(c) \to F_0(V_\beta)$ is analytic.*

By Lemma 4.24, for sufficiently small β's,

$$F_0(V_\beta) = \sum_{\mathbf{k}\in\mathbb{N}^n\setminus(0,..,0)} \prod_{1\le i\le n} \frac{(-\beta_i)^{k_i}}{k_i!} \mathcal{M}_{\mathbf{k}}.$$

Hence, the above Lemma shows that this generating function extends analytically on $T(c)$. The proof of the Lemma goes as follows. We denote X_t^β the process of Lemma 5.7 build with the potential V_β. We then prove that there exists a family $X_t^{(k_1,\ldots,k_n)}, k_i \in \mathbb{N}, 1 \le i \le n$ of operator-valued processes such that for $\eta \in \mathbb{C}^n$, $|\beta - \eta| := \max_{1\le i\le n} |\beta_i - \eta_i|$ small enough,

$$X_t^\eta = X_t^\beta + \sum_{\substack{k_1,\ldots,k_n\in\mathbb{N}^n \\ \sum k_i \ge 1}} \prod_{i=1}^n (\eta_i - \beta_i)^{k_i} X_t^{(k_1,\ldots,k_n)} \tag{71}$$

The variables $X_t^{(k_1,\ldots,k_n)}, k_i \in \mathbb{N}, 1 \le i \le n$ are just obtained as solution of some stochastic differential system obtained by differentiating (5.7) with respect to the β_i's. Moreover, $X_t^{(k_1,\ldots,k_n)}$ are processes such that there exists a constant C that only depends on c and the degree of V so that

$$\sup_{t\in\mathbb{R}^+} \|X_t^{(k_1,\ldots,k_n)}\|_\infty \le C^{\sum k_i} \tag{72}$$

and hence the right hand side of (71) converges in norm for any η so that $|\eta - \beta| < 1/C$. Finally the distribution of $\left(X_t^{(k_1,\ldots,k_n)}\right)_{k_1,\ldots,k_n\in\mathbb{N}^n}$ converges (in the sense of finite marginals, i.e., on polynomials involving only a finite number of the $\left(X_t^{(k_1,\ldots,k_n)}\right)_{k_1,\ldots,k_n\in\mathbb{N}^n}$) towards the law of $\left(X_\infty^{(k_1,\ldots,k_n)}\right)_{k_1,\ldots,k_n\in\mathbb{N}^n}$ as t goes to infinity. We do not detail the proof of the facts above but just underline that it is based on the uniform boundedness of X_t^β. They allow us, by letting t going to infinity, to prove that for $\eta \subset B(\beta, \frac{1}{C}) \cap T(c)$,

$$X^\eta = X^\beta + \sum_{\substack{k_1,\ldots,k_n\in\mathbb{N}^n \\ \sum k_i \ge 1}} \prod_{i=1}^n (\eta_i - \beta_i)^{k_i} X^{(k_1,\ldots,k_n)}$$

with X^β with law τ_{V_β} and operators $(X^{(k_1,\ldots,k_n)}, k_1, \cdots, k_n \in \mathbb{N}^n)$ uniformly bounded by C. We thus get the analyticity. Moreover, when $\beta = 0$, the $X^{(k_1,\ldots,k_n)}$ can be explicitly constructed via the free Brownian motion, thus giving a 'free probabilist' representation of the laws $\tau_{V_\beta} = \mathcal{M}_\beta$.

3. Discussion

- Uniqueness of the solution to the Schwinger-Dyson equation is not true in general as well as convergence of the dynamics to a solution is not true in general (see [**17**], section 7.1). The situation is even more dramatic than in the classical case since, because of the boundedness of S, the process may not be able to quit a well of the potential W inside which it started (with probability one).

- The study of a low temperature phase should be doable. Dembo, Maurel Segala and myself started the analysis of the static case but not of the associated dynamics.
- In this paragraph, we have constructed via a free Brownian motion the solution τ_V, that, as we have seen, is related with the enumeration of maps. Can this be useful to understand better the enumeration question? Note (see section 4.25) that in the case $m = 1$, the fact that τ_V was a probability measure with connected support allowed to get a priori an explicit formula for $\mathcal{M}_V$.

  In the same vain, one can wonder how critical exponents (that govern the polynomial decay of the numbers $M_0((k_i, q_i), 1 \le i \le n))$ could be detected by our dynamics. Note here that they are related with the analyticity (or absence of analyticity) of $\tau_V(P)$.
- The fact that τ_V is a state (i.e that $\tau_V(PP^*) \ge 0$ for all polynomials) is a triviality from the random matrices points of view. However, when thought as the same property for the associated enumeration of graphs, it becomes much less transparent. Could there be a combinatorial explanation why τ_V is a state? This question also applies for one matrix when we wonder why τ_V is a probability measure.
- In classical statistical mechanics, it is known, for instance in spin glasses, that the dynamics may have a phase transition before the statics do. Could this happen here?

Bibliography

[1] Albeverio S., P. L., and M., S. On the $1/n$ expansion for some unitary invariant ensembles of random matrices. *Comm. Math. Phys. 224*, 1 (2001), 271–305. Dedicated to Joel L. Lebowitz.

[2] Ambjorn, J., Chekhov, Kristjansen, L., C.F., and Makeenko, Y. Matrix model calculations beyond the spherical limit. *Nuclear Physics B 404* (1993), 127–172.

[3] Anderson, G., Guionnet, A., and Zeitouni, O. *Lectures notes on random matrices*.

[4] Anderson, G., and Zetouni, O. A clt for a band matrix model. *Probab. Theory Related Fields To appear* (2005).

[5] Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C., and Scheffer, G. *Sur les inégalités de Sobolev logarithmiques*, vol. 10 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.

[6] Auffinger, A., Ben Arous, G., and Péché, S. Poisson convergence for the largest eigenvalues of heavy tailed random matrices. *http://front.math.ucdavis.edu/0710.3132* (2007).

[7] Bai, Z. D. Convergence rate of expected spectral distributions of large random matrices. II. Sample covariance matrices. *Ann. Probab. 21*, 2 (1993), 649–672.

[8] Bai, Z. D. Circular law. *Ann. Probab. 25*, 1 (1997), 494–529.

[9] Baik, J., Deift, P., and Johansson, K. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc. 12*, 4 (1999), 1119–1178.

[10] Ben Arous, G., and Guionnet, A. Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Related Fields 108*, 4 (1997), 517–542.

[11] Ben Arous, G., and Guionnet, A. The spectrum of heavy tailed random matrices. *To appear in Comm. Math. Phys.* (2007).

[12] Bessis, D., Itzykson, C., and Zuber, J. B. Quantum field theory techniques in graphical enumeration. *Adv. in Appl. Math. 1*, 2 (1980), 109–157.

[13] Bhatia, R. *Perturbation bounds for matrix eigenvalues*, vol. 53 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2007. Reprint of the 1987 original.

[14] Biane, P. Free probability for probabilists. In *Quantum probability communications, Vol. XI (Grenoble, 1998)*, QP-PQ, XI. World Sci. Publishing, River Edge, NJ, 2003, pp. 55–71.

[15] Biane, P., Capitaine, M., and Guionnet, A. Large deviation bounds for matrix Brownian motion. *Invent. Math. 152*, 2 (2003), 433–459.

[16] Biane, P., and Speicher, R. Stochastic calculus with respect to free Brownian motion and analysis on Wigner space. *Probab. Theory Related Fields 112*, 3 (1998), 373–409.

[17] Biane, P., and Speicher, R. Free diffusions, free entropy and free Fisher information. *Ann. Inst. H. Poincaré Probab. Statist. 37*, 5 (2001), 581–606.

[18] Bobkov, S. G., and Götze, F. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal. 163*, 1 (1999), 1–28.

[19] Borodin, A., Ferrari, P. L., and Prähofer, M. Fluctuations in the discrete TASEP with periodic initial configurations and the Airy_1 process. *Int. Math. Res. Pap. IMRP*, 1 (2007), Art. ID rpm002, 47.

[20] Borodin, A., Ferrari, P. L., Prähofer, M., and Sasamoto, T. Fluctuation properties of the TASEP with periodic initial configuration. *J. Stat. Phys. 129*, 5-6 (2007), 1055–1080.

[21] Borodin, A., and Olshanski, G. Random partitions and the gamma kernel. *Adv. Math. 194*, 1 (2005), 141–202.

[22] Boulatov, D. V., and Kazakov, V. A. The Ising model on a random planar lattice: the structure of the phase transition and the exact critical exponents. *Phys. Lett. B 186*, 3-4 (1987), 379–384.

[23] Bousquet-Melou, M., and Schaeffer, G. The degree distribution in bipartite planar maps: applications to the Ising model. *arXiv:math.CO/0211070* (2002).

[24] Boutet de Monvel, A., Khorunzhy, A., and Vasilchuk, V. Limiting eigenvalue distribution of random matrices with correlated entries. *Markov Process. Related Fields 2*, 4 (1996), 607–636.

[25] Brascamp, H. J., and Lieb, E. H. On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Functional Analysis 22*, 4 (1976), 366–389.

[26] Brézin, E., Itzykson, C., Parisi, G., and Zuber, J. B. Planar diagrams. *Comm. Math. Phys. 59*, 1 (1978), 35–51.

[27] Cabanal Duvillard, T., and Guionnet, A. Large deviations upper bounds for the laws of matrix-valued processes and non-communicative entropies. *Ann. Probab. 29*, 3 (2001), 1205–1261.

[28] Cabanal-Duvillard, T., and Guionnet, A. Discussions around Voiculescu's free entropies. *Adv. Math. 174*, 2 (2003), 167–226.

[29] Caffarelli, L. A. Monotonicity properties of optimal transportation and the FKG and related inequalities. *Comm. Math. Phys. 214*, 3 (2000), 547–563.

[30] Chadha, S., Mahoux, G., and Mehta, M. L. A method of integration over matrix variables. II. *J. Phys. A 14*, 3 (1981), 579–586.

[31] Cizeau, P., and Bouchaud, J.-P. Theory of lévy matrices. *Physical Review E 50*, 3 (1994), 1810–1822.

[32] Conrey, B., and Gamburd, A. Pseudomoments of the Riemann zeta-function and pseudomagic squares. *J. Number Theory 117*, 2 (2006), 263–278.

[33] Conrey, B., and Gamburd, A. Pseudomoments of the Riemann zeta-function and pseudomagic squares. *J. Number Theory 117*, 2 (2006), 263–278.

[34] Deift, P., Kriecherbauer, T., and McLaughlin, K. T.-R. New results on the equilibrium measure for logarithmic potentials in the presence of an external field. *J. Approx. Theory 95*, 3 (1998), 388–475.

[35] Di Francesco P. D., G. P., and J., Z.-J. 2d gravity and random matrices. *Phys. Rep.*, 254 (1995).

[36] Diaconis, P., and Gamburd, A. Random matrices, magic squares and matching polynomials. *Electron. J. Combin. 11*, 2 (2004/06), Research Paper 2, 26 pp. (electronic).

[37] Dobrushin, R., Groeneboom, P., and Ledoux, M. *Lectures on probability theory and statistics*, vol. 1648 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1996. Lectures from the 24th Saint-Flour Summer School held July 7–23, 1994, Edited by P. Bernard.

[38] Dyson, F. J. A Brownian-motion model for the eigenvalues of a random matrix. *J. Mathematical Phys. 3* (1962), 1191–1198.

[39] El Karoui, N. A rate of convergence result for the largest eigenvalue of complex white Wishart matrices. *Ann. Probab. 34*, 6 (2006), 2077–2117.

[40] Ercolani, N. M., and McLaughlin, K. D. T.-R. Asymptotics of the partition function for random matrices via Riemann-Hilbert techniques and applications to graphical enumeration. *Int. Math. Res. Not.*, 14 (2003), 755–820.

[41] Eynard, B. An introduction to random matrices. CEA/SPhT, Saclay.

[42] Eynard, B. Master loop equations, free energy and correlations for the chain of matrices. *J. High Energy Phys.*, 11 (2003), 018, 45 pp. (electronic).

[43] Füredi, Z., and Komlós, J. The eigenvalues of random symmetric matrices. *Combinatorica 1*, 3 (1981), 233–241.

[44] Ge, L. Applications of free entropy to finite von Neumann algebras. *Amer. J. Math. 119*, 2 (1997), 467–485.

[45] Ge, L. Applications of free entropy to finite von Neumann algebras. II. *Ann. of Math. (2) 147*, 1 (1998), 143–157.

[46] Gessel, I., and Viennot, G. Binomial determinants, paths, and hook length formulae. *Adv. in Math. 58*, 3 (1985), 300–321.

[47] Guhr, T., M.-G.-A., and Weidenmuller, H. *random matrix theory in quantum Physics : Common concepts*. arXiv:cond-mat/9707301. 1997.

[48] Guionnet, A. First order asymptotics of matrix integrals; a rigorous approach towards the understanding of matrix models. *Comm. Math. Phys. 244*, 3 (2004), 527–569.

[49] Guionnet, A. *Lectures on Random Matrices : Macroscopic asymptotics*. Lecture Notes for the 2006 Saint-Flour Probability Summer School. www.umpa.ens-lyon.fr/ aguionne, 2007.

[50] Guionnet, A., and Maurel-Segala, E. Combinatorial aspects of matrix models. *arviv http://front.math.ucdavis.edu/math.PR/0503064* (2005).

[51] Guionnet, A., and Maurel-Segala, E. Combinatorial aspects of matrix models. *ALEA Lat. Am. J. Probab. Math. Stat. 1* (2006), 241–279 (electronic).

[52] Guionnet, A., and Maurel-Segala, E. Second order asymptotics for matrix models. *Ann. Probab. 35*, 6 (2007), 2160–2212.

[53] Guionnet, A., and Shlyakhtenko, D. Free diffusions and matrix models with strictly convex interaction. *http://arxiv.org/abs/math/0701787* (2007).

[54] Guionnet, A., and Zegarlinski, B. Lectures on logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXVI*, vol. 1801 of *Lecture Notes in Math.* Springer, Berlin, 2003, pp. 1–134.

[55] Guionnet, A., and Zeitouni, O. Concentration of the spectral measure for large matrices. *Electron. Comm. Probab. 5* (2000), 119–136 (electronic).

[56] Guionnet, A., and Zeitouni, O. Large deviations asymptotics for spherical integrals. *J. Funct. Anal. 188*, 2 (2002), 461–515.

[57] Guionnet, A., and Zeitouni, O. Addendum to large deviations asymptotics for spherical integrals. *J. Funct. Anal. To appear* (2004).

[58] Haagerup, U., Schultz, H., and Thorbjørnsen, S. A random matrix approach to the lack of projections in $C^*_{\text{red}}(\mathbb{F}_2)$. *Adv. Math. 204*, 1 (2006), 1–83.

[59] Haagerup, U., and Thorbjørnsen, S. A new application of random matrices: $\text{Ext}(C^*_{\text{red}}(F_2))$ is not a group. *Ann. of Math. (2) 162*, 2 (2005), 711–775.

[60] Harer, J., and Zagier, D. The Euler characteristic of the moduli space of curves. *Invent. Math. 85*, 3 (1986), 457–485.

[61] Hargé, G. A convex/log-concave correlation inequality for Gaussian measure and an application to abstract Wiener spaces. *Probab. Theory Related Fields 130*, 3 (2004), 415–440.

[62] Johansson, K. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J. 91*, 1 (1998), 151–204.

[63] Johansson, K. Shape fluctuations and random matrices. *Comm. Math. Phys. 209*, 2 (2000), 437–476.

[64] Johansson, K. Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Comm. Math. Phys. 215*, 3 (2001), 683–705.

[65] Johansson, K. Toeplitz determinants, random growth and determinantal processes. In *Proceedings of the International Congress of Mathematicians, Vol. III (Beijing, 2002)* (Beijing, 2002), Higher Ed. Press, pp. 53–62.

[66] Johansson, K. The arctic circle boundary and the Airy process. *Ann. Probab. 33*, 1 (2005), 1–30.

[67] Johnstone, I. M. On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist. 29*, 2 (2001), 295–327.

[68] Karatzas, I., and Shreve, S. E. *Brownian motion and stochastic calculus*, second ed., vol. 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991.

[69] Karlin, S., and McGregor, J. Coincidence probabilities. *Pacific J. Math. 9* (1959), 1141–1164.

[70] Keating, J. P. Random matrices and the Riemann zeta-function. In *Highlights of mathematical physics (London, 2000)*. Amer. Math. Soc., Providence, RI, 2002, pp. 153–163.

[71] Keating, J. P. L-functions and the characteristic polynomials of random matrices. In *Recent perspectives in random matrix theory and number theory*, vol. 322 of *London Math. Soc. Lecture Note Ser.* Cambridge Univ. Press, Cambridge, 2005, pp. 251–277.

[72] Kenyon, R., Okounkov, A., and Sheffield, S. Dimers and amoebae. *Ann. of Math. (2) 163*, 3 (2006), 1019–1056.

[73] Kontsevich, M. Intersection theory on the moduli space of curves and the matrix Airy function. *Comm. Math. Phys. 147*, 1 (1992), 1–23.

[74] Ledoux, M. *The concentration of measure phenomenon*, vol. 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[75] Ledoux, M. A remark on hypercontractivity and tail inequalities for the largest eigenvalues of random matrices. In *Séminaire de Probabilités XXXVII*, vol. 1832 of *Lecture Notes in Math.* Springer, Berlin, 2003, pp. 360–369.

[76] Marčenko, V. A., and Pastur, L. A. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.) 72 (114)* (1967), 507–536.

[77] Maurel-Segala, E. High order expansion for matrix models. *http://front.math.ucdavis.edu/0608.5192* (2006).

[78] Mehta, M. L. A method of integration over matrix variables. *Comm. Math. Phys. 79*, 3 (1981), 327–340.

[79] Mehta, M. L. *Random matrices*, third ed., vol. 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, 2004.

[80] Okounkov, A. Infinite wedge and random partitions. *Selecta Math. (N.S.) 7*, 1 (2001), 57–81.

[81] Pastur, L., and Shcherbina, M. Universality of the local eigenvalue statistics for a class of unitary invariant random matrix ensembles. *J. Statist. Phys. 86*, 1-2 (1997), 109–147.

[82] Péché, S., and Soshnikov, A. Wigner random matrices with non-symmetrically distributed entries. *arxiv :math/0702035* (2007).

[83] Pedersen, G. K. *C^*-algebras and their automorphism groups*, vol. 14 of *London Mathematical Society Monographs*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1979.

[84] Peres, Y., and Virág, B. Zeros of the i.i.d. Gaussian power series: a conformally invariant determinantal process. *Acta Math. 194*, 1 (2005), 1–35.

[85] Pimsner, M., and Voiculescu, D. K-groups of reduced crossed products by free groups. *J. Operator Theory 8*, 1 (1982), 131–156.

[86] Potters, M., Bouchaud, J.-P., and Laloux, L. Financial applications of random matrix theory: old laces and new pieces. *Acta Phys. Polon. B 36*, 9 (2005), 2767–2784.

[87] Ramirez, J., Rider, B., and Virag, B. Beta ensembles, stochastic airy spectrum, and a diffusion. *http://arxiv.org/abs/math/0607331*.

[88] Ruzmaikina, A. Universality of the edge distribution of eigenvalues of Wigner random matrices with polynomially decaying distributions of entries. *Comm. Math. Phys. 261*, 2 (2006), 277–296.

[89] Silverstein, J. W., and Bai, Z. D. On the empirical distribution of eigenvalues of a class of large-dimensional random matrices. *J. Multivariate Anal. 54*, 2 (1995), 175–192.

[90] Sinai, Y., and Soshnikov, A. Central limit theorem for traces of large random symmetric matrices with independent matrix elements. *Bol. Soc. Brasil. Mat. (N.S.) 29*, 1 (1998), 1–24.

[91] Soshnikov, A. Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys. 207*, 3 (1999), 697–733.

[92] Soshnikov, A. Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys. 207*, 3 (1999), 697–733.

[93] Soshnikov, A. Poisson statistics for the largest eigenvalues in random matrix ensembles. In *Mathematical physics of quantum mechanics*, vol. 690 of *Lecture Notes in Phys.* Springer, Berlin, 2006, pp. 351–364.

[94] 't Hooft, G. A planar diagram theory for strong interactions. *Nuclear Physics B 72*, 3 (1974), 461–473.

[95] Tao, T., and Vu, V. Random matrices:the circular law. *http://front.math.ucdavis.edu/0708.2895* (2007).

[96] Tracy, C. A., and Widom, H. Level-spacing distributions and the Airy kernel. *Comm. Math. Phys. 159*, 1 (1994), 151–174.

[97] Tracy, C. A., and Widom, H. The distribution of the largest eigenvalue in the Gaussian ensembles: $\beta = 1, 2, 4$. In *Calogero-Moser-Sutherland models (Montréal, QC, 1997)*, CRM Ser. Math. Phys. Springer, New York, 2000, pp. 461–472.

[98] Tracy, C. A., and Widom, H. Universality of the distribution functions of random matrix theory. In *Integrable systems: from classical to quantum (Montréal, QC, 1999)*, vol. 26 of *CRM Proc. Lecture Notes*. Amer. Math. Soc., Providence, RI, 2000, pp. 251–264.

[99] Tse, D., and Zeitouni, O. Linear multiuser receivers in random environments. *IEEE trans. IT 46* (2000), 171–188.

[100] Tutte, W. T. On the enumeration of planar maps. *Bull. Amer. Math. Soc. 74* (1968), 64–74.

[101] Voiculescu, D. Limit laws for random matrices and free products. *Invent. Math. 104*, 1 (1991), 201–220.

[102] Voiculescu, D. The analogues of entropy and of Fisher's information measure in free probability theory. III. The absence of Cartan subalgebras. *Geom. Funct. Anal. 6*, 1 (1996), 172–199.

[103] Voiculescu, D. Lectures on free probability theory. In *Lectures on probability theory and statistics (Saint-Flour, 1998)*, vol. 1738 of *Lecture Notes in Math.* Springer, Berlin, 2000, pp. 279–349.

[104] Wigner, E. P. On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2) 67* (1958), 325–327.

[105] WISHART, J. The generalized product moment distribution in samples from a normal multivariate population. *Biometrika 20* (1928), 35–52.

[106] Zakharevich, I. A generalization of wigner's law.

[107] Zvonkin, A. Matrix integrals and map enumeration: an accessible introduction. *Math. Comput. Modelling 26*, 8-10 (1997), 281–304. Combinatorics and physics (Marseilles, 1995).