Lecture Notes, Minneapolis

Alice Guionnet[1] ,

July 2, 2012

[1]UMPA, CNRS UMR 5669, ENS Lyon, 46 allée d'Italie, 69007 Lyon, France. aguionne@ens-lyon.fr. This work was partially supported by the ANR project ANR-08-BLAN-0311-01.

Chapter 1

Introduction

In these notes we will consider the asymptotics of the empirical measure of random matrices, or more generally the trace of functions in random matrices. For instance, we will consider a Wigner matrix X_N , that is a $N\times N$ matrix so that $X_N = X_N^*$ with i.i.d entries $(X_N(ij))_{1\le i\le j\le N}$ with law μ (on the real line or the complex plan), independent from the i.i.d entries $(X_N(ii))_{1\le i\le N}$ with law ν on the real line. As a self-adjoint matrix, its eigenvalues $(\lambda_1,\dots,\lambda_N)$ are real. We shall wonder about the asymptotics of the spectral measure

$$L_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i/\sqrt{N}}$$

and more precisely

- Derive the asymptotics of $\mathbb{E}[\int f(x)dL_N(x)]$,
- Study the central limit theorem for the random variable $\int f(x)dL_N(x)$
- Study the large deviations from this limit in the case of Gaussian entries (GOE/GUE)

We will then start going on even finer analysis which is specific to random matrices, that is large dimension expansion, which was shown to be a "topological expansion " by 'thooft and Brézin-Parisi-Itzykson and Zuber. More precisely, if X_N follows the GUE law, then we have

$$\mathbb{E}[\frac{1}{N}\mathrm{tr}X_N^p] = \sum_{p\ge 0}\frac{1}{N^{2g}}M(g,p)$$

where $M(g,p)$ are integer numbers that counts the number of maps with genus g with one vertex of degree p. Maps are graphs which can be properly embedded (that is so that edges do not cross) into a surface with given genus, and its genus is the minimal genus of such a surface. This expansion was shown to hold at list formally on a much greater generality if one adds a potential. Namely, let $V(x) = \sum t_i x^i$ and consider the law

$$d\mu_N^V(X) = \frac{1}{Z_N^V}\exp\{N\mathrm{Tr}(V(X))\}d\mu_N(X)$$

with μ_N the law of the GUE. Then, at least as formal expansion we have the equality

$$\int \frac{1}{N}\mathrm{tr}X^p d\mu_N^V(X) = \sum_{g\geq 0} \frac{1}{N^{2g}} \sum_{n_i\geq 0} \prod \frac{t_i^{n_i}}{n_i!} M(g,p,n_1,\ldots,n_p)$$

where $M(g,p,n_1,\ldots,n_p)$ is the number of maps with genus g build on on vertex of degree p, and n_i vertices of degree i. We shall prove that this expansion can be turned into an asymptotic expansion following a joint work with Gaetan Borot [28].

This combinatorial link was used in [32] to indeed enumerate planar maps (that is connected graphs which can be properly embedded into the sphere) build on vertices with degree 3, that is triangulations of the sphere. It turns out in fact that considering the same type of enumeration but with several matrices is even more fruitful and allows to enumerate many topological objects. Indeed, several matrix models are related with the enumeration of colored maps which are extremly rich objects. We shall also tackle the topological expansion of several matrix models.

In a last part of the course we shall consider diverse generalizations. They can take different directions. First one can try to build matrix models to enumerate a given combinatorial object. We can think about loop models for instance, but more intricated models can be related to discrete objects such as plane partitions [54] or to topological strings [13]. Another question is whether non Gaussian matrix models can also exhibit such a topological expansion. A natural candidate is the Haar measure on the unitary group. Surprisingly it turns out that at list at first order we also have a kind of topological expansion [42]. For instance, if A_i are matrices so that

$$\frac{1}{N}\mathrm{Tr}(P(A_i, 1 \leq i \leq n))$$

converges for all polynomial P, if V is a polynomial function so that $V(U,U^*,A_i,1 \leq i \leq N)$ is self-adjoint, for t real small enough

$$\frac{1}{N^2}\log \int e^{tN\mathrm{Tr}V(U,U^*,A_i,1\leq i\leq N)} dU$$

converges as N goes to infinity. Furthermore, the limit is a converging series in t whose coefficients are integer numbers that can be determined as the enumeration of some planar maps. This applies in particular for the so-called Harich-Chandra–Itzykson-Zuber integral

$$\frac{1}{N^2}\log \int e^{tN\mathrm{Tr}(UAU^*B)} dU\,.$$

Chapter 2

Wigner matrices: generalities

2.1 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; $\mathbf{X}^N_{ij} = \overline{\mathbf{X}}^N_{ji}$. 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

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

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 (1) from moment convergence is done in section 2.2.

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

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

Then,

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

where the convergence holds in expectation and almost surely. $(C_k)_{k\geq 0}$ *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 $i \neq j$ and $\cup V_i = S$. The $V_i, 1 \leq i \leq r$ 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 π of $\{1, \cdots, n\}$ is said to be *crossing* if there exist $1 \leq p_1 < q_1 < p_2 < q_2 \leq n$ 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

$$\mathbb{E}\left[\frac{1}{N}\mathrm{tr}\left((\mathbf{X}^N)^k\right)\right] = \sum_{i_1, \cdots, i_k = 1}^{N} N^{-\frac{k}{2}-1} \mathbb{E}[Y_{i_1 i_2} Y_{i_2 i_3} \cdots Y_{i_k i_1}] \tag{3}$$

   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

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

   By (2) and Hölder's inequality, $P(\mathbf{i})$ is bounded uniformly 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 (3), 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 $G(\mathbf{i})$.

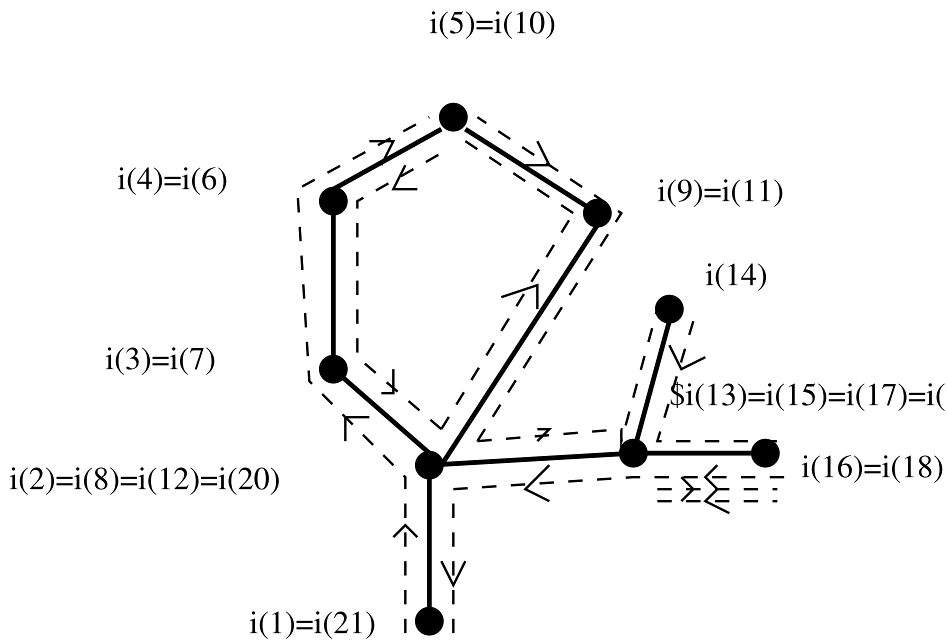


Figure 2.1: Figure of $G(\mathbf{i})$ (in dash) versus $\tilde{G}(\mathbf{i})$ (in bold), $|\tilde{E}(i)| = 9, |\tilde{V}(\mathbf{i})| = 9$

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

$$|V| \leq |E| + 1. \tag{4}$$

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 proven 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 which 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 $(G_j)_{1 \leq j \leq r}$ gives

$$\begin{aligned} |V| - 1 &\leq \sum_{i=1}^{r} (|E_j| + 1) \\ &= |E| + r - l \leq |E| \end{aligned} \tag{5}$$

which proves (4). 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 which 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 (5) (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

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

   and so by (4) applied to the skeleton $\tilde{G}(\mathbf{i})$ we find

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

   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 which contribute to the sum (3) and so we have

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

   where we used (2). 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 which 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, $P(\mathbf{i})=E[|\sqrt{N}X^N_{ij}|^2]^{\frac{k}{2}}=1$ does not depend on $\mathbf{i}$.

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

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

with

$$P(\mathbf{i},\mathbf{i}'):=\mathbb{E}[Y_{i_1i_2}Y_{i_2i_3}\cdots Y_{i_ki_1}Y_{i'_1i'_2}\cdots Y_{i'_ki'_1}].$$

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

$$|\tilde V(\mathbf{i},\mathbf{i}')|\le|\tilde E(\mathbf{i},\mathbf{i}')|+1\le k+1. \tag{6}$$

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 (6). 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,

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

which implies

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

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 $\delta > 0$

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

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. Take for $L \in \mathbb{N}$, $\mathbf{X}^{N,L}$ the $N \times N$ self-adjoint matrix such that $\mathbf{X}^{N,L}_{ij} = (2L)^{-\frac{1}{2}} 1_{|i-j|\leq L} X_{ij}$ with $(X_{ij}, 1 \leq i \leq j \leq N)$ 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 $k \geq 2$

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

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 $i_2,\cdots,i_k$.

2.2 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 2. *Let $(\lambda_i)_{1\leq i\leq N}$ 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 (2) 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

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

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

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

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 $[-2,2]$) which goes to zero as p goes to infinity. Consequently,

$$\begin{aligned}\left|\int f(x)d(L_{\mathbf{X}^N}(x)-\sigma(x))\right| &\leq \left|\int P_\delta(x)d(L_{\mathbf{X}^N}(x)-\sigma(x))\right| \\ &\quad +\delta+\left|\int_{|x|\geq B}(f-P_\delta)(x)dL_{\mathbf{X}^N}(x)\right| \end{aligned} \tag{7}$$

goes to zero as N goes to infinity. □

2.2.1 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 $(X^N_{ij})_{1\le i\le j\le N}$ are independent. The generalization of Wigner's theorem to possibly mildly dependent entries can be found for instance in [29]. 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 2. 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 [9]).

Corollary 3. *Assume that $\{\sqrt{N}\mathbf{X}^N_{ij}, i\le j\}$ 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 2.

Remark. When the entries are not equidistributed, the convergence in probability can be proved when $\{\sqrt{N}\mathbf{X}^N_{ij}, i\le j\}$ 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 [40] and rigorous studies in [105, 17]). Eventhough 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.

2.2.2 Relaxation of the hypothesis on the centering of the entries

A last generalization concerns the hypothesis on the mean of the variables $\sqrt{N}X^N_{ij}$ which, as we shall see, is irrelevant in the statement of Corollary 3. More precisely, we shall prove that (proof originated from [67])

Lemma 4. 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,

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

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,

$$\left| \int f(x) dL_{\mathbf{Z}^N}(x) - \int f(x) dL_{\mathbf{Z}^N+\lambda ee^*}(x) \right| \leq \frac{2}{N} \|f\|_\infty. \tag{8}$$

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

Theorem 3. *[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*

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

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

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

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

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

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

These two bounds prove (8). We leave the reader extend this result from $\mathbf{Y}^N = \lambda ee^*$ with rank 1 to $\mathbf{Y}^N$ with rank $r(N)$.

□

By Corollary 3 and Lemma 4, we find that

Corollary 5. *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 3 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 2 when y^N goes to zero.

2.3 Words in several independent Wigner matrices

In this chapter, we consider m independent Wigner $N\times N$ matrices $\{\mathbf{X}^{N,\ell}, 1\le \ell\le m\}$ with real or complex entries. In other words, the $\mathbf{X}^{N,\ell}$ are self-adjoint random matrices with independent entries $\left(\mathbf{X}^{N,\ell}_{ij}, 1\le i\le j\le N\right)$ above the diagonal which are centered and with variance one. Moreover, the $\left(\mathbf{X}^{N,\ell}_{ij}, 1\le i\le j\le N\right)_{1\le\ell\le m}$ are independent. We shall generalize Theorem 4 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 to the context of several matrices as a first step towards part 4. Let us first describe the combinatorial objects that we shall need.

2.3.1 Partitions of colored elements

Because we now have m different matrices, the partitions which 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 6. 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 which 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).

2.3.2 Voiculescu's theorem

The aim of this chapter 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 4. *[Voiculescu [102]] Assume that for all $k\in\mathbb{N}$,*

$$B_k := \sup_{1\le\ell\le m}\sup_{N\in\mathbb{N}}\sup_{ij\in\{1,\cdots,N\}^2}\mathbb{E}[|\sqrt{N}\mathbf{X}^{N,\ell}_{ij}|^k] < \infty. \tag{9}$$

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

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

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 7. 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 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 which vanishes when edges of different colors have been paired.

Exercise 8. The next exercise concerns a special case of what is called 'Asymptotic freeness' and was proved in greater generality by D. Voiculescu.

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

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

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

$$B_k = \sup_{N \in \mathbb{N}} \sup_{ij \in \{1, \cdots, N\}^2} \mathbb{E}[|\sqrt{N}\mathbf{X}^N_{ij}|^k] < \infty \tag{10}$$

Let D^N be a deterministic diagonal matrix such that

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

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. Conditionning 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).

2.4 The case of heavy tails random matrices

In the case where the entries have no finite second moment, we can not a priori use moment strategy except if we add a cutoff. This idea was developped by Zakharevich [105] and we shall briefly sketch her result and in particular stress why we are not anymore in the domain of universality of the semicircle law. So let $\mathbf{X}^N_{ij}, 1 \le i \le j \le N$ be independent with a symmetric heavy tail law, that is

$$\mathbb{P}\left(|\mathbf{X}^N_{ij}| \ge t\right) \simeq L(t)t^{-\alpha}$$

with a slowly varying function L and $\alpha < 2$. Then, set

$$\mathbf{X}^{N,B}_{ij} = 1_{|\mathbf{X}^{N,B}_{ij}| \le BN^{1/\alpha}} \mathbf{X}^{N,B}_{ij}$$

and let us consider the symmetric matrix $\mathbf{X}^{N,B}$ with entries $\mathbf{X}^{N,B}_{ij}$. Observe that for $p \ge 1$

$$\mathbf{E}[(\mathbf{X}^{N,B}_{ij})^{2p}] \sim C(p)B^{2p}N^{2p/\alpha}N^{-1}.$$

Hence, we find that if G' denotes a skeleton (that is a rooted connected graph with simple non-oriented edges) with less than p edges and P the set of rooted loops on G' with length $2p$ so that if $m(P,e)$ denotes the number of times that the path P goes through the edge e of G', $m(P,e) \ge 2$ then

$$\begin{aligned}
\mathbb{E}[\frac{1}{N}\mathrm{tr}((N^{-1/\alpha}\mathbf{X}^{N,B})^{2p}] &= \sum_{G'=(V',E')} N^{|V'|-|E'|-1} \sum_P \prod_{e\in E'} (N\mathbb{E}[(\frac{\mathbf{X}_{11}}{N^{1/\alpha}})^{m(e)}]) \\
&\sim \sum_{G'=(V',E') \text{ rooted tree}} \sum_P \prod_{e\in E'} N\mathbb{E}C(m(e))B^{2m(e)}
\end{aligned}$$

where we have used that the main contribution comes from connected graphs so that

$$|V'| = |E'| + 1$$

that is trees.

2.5 Central limit theorem

In the previous section, we proved Wigner's theorem by evaluating $\int x^p dL_{\mathbf{A}^N}(x)$ for $p \in \mathbb{N}$. We shall push this computation one step further here and prove a central limit theorem. Namely, setting

$$\int x^k d\overline{L}_{\mathbf{A}^N}(x) := \mathbb{E}[\int x^k dL_{\mathbf{A}^N}(x)],$$

we shall prove that

$$M_k^N := N\left(\int x^k dL_{\mathbf{A}^N}(x) - \int x^k d\overline{L}_{\mathbf{A}^N}(x)\right) = \sum_{i=1}^N \left(\lambda_i^k - \mathbb{E}[\lambda_i^k]\right)$$

converges in law to a centered Gaussian variable. Since in chapter 4 we shall give a complete and detailed proof of the central limit theorem in the case of Gaussian entries with a weak interaction, we will be rather sketchy here. We refer to [10] for a complete and clear treatment and [9] for a simplified exposition of the full proof of the theorem we state below. To simplify, we assume here that $\mathbf{A}^N$ is a Wigner matrix with

$$A_{ij}^N = \frac{B_{ij}}{\sqrt{N}},$$

where $(B_{ij}, 1 \le i \le j \le N)$ are independent real equidistributed random variables. Their marginal distribution μ has all moments finite (in particular (2) is satisfied) and satisfies

$$\int x d\mu(x) = 0 \text{ and } \int x^2 d\mu(x) = 1.$$

We shall show why the following statement holds.

Theorem 5. *Let*

$$\sigma_k^2 = k^2 [C_{\frac{k-1}{2}}]^2 + \frac{k^2}{2}[C_{\frac{k}{2}}]^2 [\int x^4 d\mu(x) - 1] + \sum_{r=3}^{\infty} \frac{2k^2}{r}\left(\sum_{\substack{k_i \ge 0 \\ 2\sum_{i=1}^r k_i = k-r}} \prod_{i=1}^r C_{k_i}\right)^2,$$

In this formula, C_x equals zero if x is not an integer and otherwise is equal to the Catalan number.

Then, M_k^N converges in moments to the centered Gaussian variable with variance σ_k^2, i.e., for all $l \in \mathbb{N}$,

$$\lim_{N\to\infty} \mathbb{E}\left[(M_k^N)^l\right] = \frac{1}{\sqrt{2\pi}\sigma_k} \int x^l e^{-\frac{x^2}{2\sigma_k^2}} dx.$$

Remark. Unlike the standard central limit theorem for independent variables, the variance here depends on $\mu(x^4)$.

Outline of the proof.

- *We first prove that the statement is true when $l = 2$.* (It is clearly true for $k = 1$ since A_k^N is centered.) We thus want to show

$$\sigma_k^2 = \lim_{N\to\infty} \mathbb{E}\left[(M_k^N)^2\right]. \tag{11}$$

Below (6), we proved that $\mathbb{E}\left[(A_k^N)^2\right]$ is bounded, uniformly in N. Furthermore, we can write

$$\mathbb{E}\left[(M_k^N)^2\right] = \frac{1}{N^k}\sum_{\mathbf{i},\mathbf{i}'}[P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')]$$

where the sum over $\mathbf{i},\mathbf{i}'$ will hold on graphs $\tilde{G}(\mathbf{i},\mathbf{i}') = (\tilde{V}(\mathbf{i},\mathbf{i}'),\tilde{E}(\mathbf{i},\mathbf{i}'))$ so that

$$|\tilde{V}(\mathbf{i},\mathbf{i}')| \leq k, \quad |\tilde{E}(\mathbf{i},\mathbf{i}')| \leq k.$$

Since $[P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}')]$ is uniformly bounded, the only contributing graphs to the leading order will be those for which $|\tilde{V}(\mathbf{i},\mathbf{i}')| = k$. Then, since we always have $|\tilde{V}(\mathbf{i},\mathbf{i}')| \leq |\tilde{E}(\mathbf{i},\mathbf{i}')| + 1$, we have two cases :

• $|\tilde{E}(\mathbf{i},\mathbf{i}')| = k-1$ in which case the skeleton $\tilde{G}(\mathbf{i},\mathbf{i}')$ will again be a tree but with one edge less than the total number possible; this means that one edge appears with multiplicity four and belongs to $\tilde{E}(\mathbf{i}) \cap \tilde{E}(\mathbf{i}')$, the other edges appearing with multiplicity 2. Hence, the graphs of $\tilde{E}(\mathbf{i})$ and $\tilde{E}(\mathbf{i}')$ are both trees (which implies that k is even); there are $C_{\frac{k}{2}}^2$ such trees, and they are glued by a common edge, to choose among $\frac{k}{2}$ edges in each of the tree. Finally, there are two possible choice to glue the two trees according to the orientation. Thus, there are

$$2(\frac{k}{2})^2 C_{\frac{k}{2}}^2 = (\frac{k^2}{2}) C_{\frac{k}{2}}^2$$

such graphs and then

$$P(\mathbf{i},\mathbf{i}') - P(\mathbf{i})P(\mathbf{i}') = \int x^4 d\mu(x) - 1.$$

We hence obtain the contribution $(\frac{k^2}{2})C_{\frac{k}{2}}^2(\int x^4 d\mu(x) - 1)$ to the variance.

• $|\tilde{E}(\mathbf{i},\mathbf{i}')| = k$. In this case, the graph is not a tree anymore and because $|\tilde{E}(\mathbf{i},\mathbf{i}')| - |\tilde{V}(\mathbf{i},\mathbf{i}')| = 1$, it contains exactly one cycle. This can be seen either by closer inspection of the arguments given after (4) or by using the formula which relates the genus of a graph and its number of vertices, faces and edges;

$$\sharp\text{vertices} + \sharp\text{faces} - \sharp\text{edges} = 2 - 2g \leq 2$$

The faces are defined by following the boundary of the graph; each of these boundaries are exactly one cycle of the graph except one (since a graph has always one boundary) and therefore

$$\sharp\text{faces} = 1 + \sharp\text{cycles}.$$

So we get, for a connected graph with skeleton $(\tilde{V},\tilde{E})$,

$$|\tilde{V}| \leq |\tilde{E}| + 1 - \sharp\text{cycles} \tag{12}$$

In our case, $\sharp\text{vertices} = \sharp\text{edges} = k$ and $\sharp\text{cycles} \geq 1$ (since the graph is not a tree), which implies that $\sharp\text{cycles} = 1$. Counting the number of such graphs completes the proof of the convergence of $\mathbb{E}\left[(M_k^N)^2\right]$ to σ_k^2 (see [10] for more details).

- *Convergence to the Gaussian law.*

  We next show that M_k^N is asymptotically Gaussian. This amounts to prove that $\lim_{N\to\infty}\mathbb{E}[(M_k^N)^{2l+1}]=0$ whereas

$$\lim_{N\to\infty}\mathbb{E}[(M_k^N)^{2l}]=\sharp\{\text{number of pair partitions of } 2l \text{ elements}\}\times\sigma_k^{2l}.$$

Again, we shall expand the expectation in terms of graphs and write for $l\in\mathbb{N}$,

$$\mathbb{E}[(M_k^N)^l]=\frac{1}{N^{\frac{kl}{2}}}\sum_{\mathbf{i}_1,\cdots,\mathbf{i}_l}P(\mathbf{i}^1,\cdots,\mathbf{i}^l)$$

with $P(\mathbf{i}^1,\cdots,\mathbf{i}^l)$ given by

$$=\mathbb{E}[\Big(B_{i_1^1i_2^1}\cdots B_{i_k^1i_1^1}-\mathbb{E}[B_{i_1^1i_2^1}\cdots B_{i_k^1i_1^1}]\Big)\cdots\Big(B_{i_1^li_2^l}\cdots B_{i_k^li_1^l}-\mathbb{E}[B_{i_1^li_2^l}\cdots B_{i_k^li_1^l}]\Big)].$$

We denote by $G(\mathbf{i}^1,\cdots,\mathbf{i}^l)=(V(\mathbf{i}^1,\cdots,\mathbf{i}^l),E(\mathbf{i}^1,\cdots,\mathbf{i}^l))$ the corresponding graph; $V(\mathbf{i}^1,\cdots,\mathbf{i}^l)=\{i_n^j,1\le j\le l,1\le n\le k\}$ and $E(\mathbf{i}^1,\cdots,\mathbf{i}^l)=\{(i_n^j,i_{n+1}^j),1\le j\le l,1\le n\le k\}$ with the convention $i_{l+1}^j=i_1^j$. As before, $P(\mathbf{i}^1,\cdots,\mathbf{i}^l)$ equals zero unless each edge appears with multiplicity 2 at least. Also, because of the centering, it vanishes if there exists a $j\in\{1,\cdots,l\}$ so that $E(\mathbf{i}^1,\cdots,\mathbf{i}^l)\cap E(\mathbf{i}^j)$ does not intersect $E(\mathbf{i}^1,\cdots,\mathbf{i}^{j-1},\mathbf{i}^{j+1},\cdots,\mathbf{i}^l)$. Let us decompose $G(\mathbf{i}^1,\cdots,\mathbf{i}^l)$ into its connected components $(G_1,\cdots,G_c)$. We claim that

$$|V(\mathbf{i}^1,\cdots,\mathbf{i}^l)|\le c-l+[\frac{l(k+1)}{2}]. \tag{13}$$

This type of bound is rather intuitive; if a connected component G_i contains $G(\mathbf{i}^{j_1}),\cdots,G(\mathbf{i}^{j_p})$, each gluing of the $G(\mathbf{i}^{j_l})$ we should create either a cycle or an edge with multiplicity 4, the total number of vertices decreasing at least by one in each gluing. Hence, $|V(\mathbf{i}^1,\cdots,\mathbf{i}^l)|$ should grow linearly with the number of connected components. The proof is given in the Appendix 7.3 for completeness (see [9] or [10]). With (13), we conclude that the only indices which will contribute are such that

$$c-l+[\frac{l(k+1)}{2}]\ge\frac{kl}{2}$$

with $c\le[\frac{l}{2}]$. This implies that

$$\frac{kl}{2}\le[\frac{l}{2}]-l+[\frac{l(k+1)}{2}]\le\frac{l}{2}-l+\frac{l(k+1)}{2}=\frac{kl}{2}$$

resulting with all inequalities being equalities. Thus, to get a first order contribution we must have l even and $c=\frac{l}{2}$. In that case, we write $(s_j,r_j)_{1\le j\le l}$ the pairing so that $(G(\mathbf{i}_{s_j}),G(\mathbf{i}_{r_j}))_{1\le j\le l}$ are connected for all $1\le j\le l$ (with the convention $s_j<r_j$). By independence of the entries, we have

$$P(\mathbf{i}_1,\cdots,\mathbf{i}_{2l})=\prod_{j=1}^{l}P(\mathbf{i}_{s_j},\mathbf{i}_{r_j})$$

and so we have proved that

$$N^{-kl}\sum_{\mathbf{i}_1,\cdots,\mathbf{i}_{2l}}P(\mathbf{i}_1,\cdots,\mathbf{i}_{2l})=\sum_{\substack{s_1<\cdots<s_l\\ r_j>s_j}}\left(N^{-k}\sum_{\mathbf{i}_1,\mathbf{i}_2}P(\mathbf{i}_1,\mathbf{i}_2)\right)^l+o(1)=\sigma_k^{2l}\sum_{\substack{s_1<\cdots<s_l\\ r_j>s_j}}1+o(1)$$

which proves the claim since

$$\frac{1}{\sqrt{2\pi}}\int x^{2l}e^{-\frac{x^2}{2}}dx=\sum_{\substack{s_1<\cdots<s_l\\ r_j>s_j}}1=(2l-1)(2l-3)(2l-5)\cdots 1.$$

This completes the proof of the moments convergence. □

Exercise 9. Show that Theorem 5 implies that M_k^N converges weakly to the centered Gaussian variable with variance σ_k^2. *Hint: control tails to approximate bounded continuous functions by polynomials.*

Exercise 10. Show that in the case of heavy tails entries, see section 2.4, one needs to renormalize by $\sqrt{N}$ to get the convergence to the Gaussian law by computing the asymptotics of the covariance.

2.6 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 are known to hold. We first consider the case where the underlying measure satisfies a log-Sobolev inequality; we show how to prove this inequality in a simple situation 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 $\int f(x)dL_{\mathbf{X}^N}(x)$ with f Lipschitz, are smooth functions of the entries of the matrix $\mathbf{X}^N$ 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 chapter. 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 chapter.

To motivate the reader, let us state the type of result we want to obtain in this chapter.

To this end, we introduce some extra notations. Let us recall that if X is a symmetric (Hermitian) matrix and f is a bounded measurable function, $f(X)$ is defined as the matrix with the same eigenvectors than X but with eigenvalues which are the image by f of those of X; namely, if e is an eigenvector of X with eigenvalue λ, $Xe=\lambda e$, $f(X)e:=f(\lambda)e$. In terms of the spectral decomposition $X=UDU^*$ with U orthogonal (unitary) and D diagonal real, one has $f(X)=Uf(D)U^*$ with $f(D)_{ii}=f(D_{ii})$. For $M\in$

$\mathbb{N}$, 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 $||x||_2^2 := \langle x,x\rangle$.

Throughout this section, we denote the Lipschitz constant of a function $G:\mathbb{R}^M \to \mathbb{R}$ by

$$|G|_{\mathcal{L}} := \sup_{x\neq y\in\mathbb{R}^M} \frac{|G(x)-G(y)|}{\|x-y\|_2},$$

and call G a *Lipschitz function* if $|G|_{\mathcal{L}} < \infty$.

Lemma 11. Let $g:\mathbb{R}\to\mathbb{R}$ be Lipschitz with Lipschitz constant $|g|_{\mathcal{L}}$. Then, with X^N denoting the Hermitian (or symmetric) matrix with entries $(X_{ij}^N)_{1\le i,j\le N}$, the map $\{X_{ij}^N\}_{1\le i\le j\le N} \mapsto \mathrm{tr}(g(X^N))$ is a Lipschitz function with constant $\sqrt{N}|g|_{\mathcal{L}}$. Therefore, if the joint law of $(X_{ij}^N)_{1\le i\le j\le N}$ is 'good', there exists $\alpha > 0$, constants $c > 0$ and $C < \infty$ so that for all $N\in\mathbb{N}$

$$\mathbb{P}\left(\left|\mathrm{tr}(g(X^N)) - \mathbb{E}[\mathrm{tr}(g(X^N))]\right| > \delta|g|_{\mathcal{L}}\right) \le Ce^{-c|\delta|^\alpha}.$$

'Good' here means for instance that the law satisfies a log-Sobolev inequality; an example is when the $\{X_{ij}^N\}_{1\le i\le j\le N}$ are independent Gaussian variables with uniformly bounded covariance (see Theorem 6).

The interest of results such as Lemma 11 is that they provide bounds on deviations which do not depend on the dimension. They can be used to show law 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 a central limit theorem (when $\alpha = 2$, Lemma 11 indeed shows that $\mathrm{tr}(g(X^N)) - \mathbb{E}[\mathrm{tr}(g(X^N))]$ has a sub-Gaussian tail, providing tightness arguments for free).

2.6.1 Concentration for the spectral measure; a universal bound

We first give a concentration result due to C. Bordenave, P. Caputo and D. Chafai [26] which is based on Azuma's-Hoeffding inequality.

Lemma 12. *Let $\|f\|_{TV}$ be the total variation norm,*

$$\|f\|_{TV} = \sup_{x_1<\cdots<x_n} \sum_{i=2}^n |f(x_i) - f(x_{i-1})|$$

Then, for either the Wigner or the Wishart matrices, for any $\delta > 0$ and any function f with finite total variation norm so that $E[|\frac{1}{N}\sum_{i=1}^N f(\lambda_i)|] < \infty$,

$$P\left(\left|\frac{1}{N}\sum_{i=1}^N f(\lambda_i) - \mathbb{E}[\frac{1}{N}\sum_{i=1}^N f(\lambda_i)]\right| \ge \delta\|f\|_{TV}\right) \le 2e^{-\frac{N\delta^2}{8c_X}}$$

where $c_X = 1$ for Wigner's matrices and M/N for Wishart matrices.

Remark 13. Note that the above speed is not optimal for laws μ,ν which have sufficiently fast decaying tails as we will see below, in which case $\sum_{i=1}^N f(\lambda_i) - \mathbb{E}[\sum_{i=1}^N f(\lambda_i)]$ is of order one. However it is the optimal rate for instance for heavy tails matrices where the central limit theorem holds for $N^{-1/2}(\sum_{i=1}^N f(\lambda_i) - \mathbb{E}[\frac{1}{N}\sum_{i=1}^N f(\lambda_i)])$.

Remark 14. Note that we only required independence of the vectors, rather than the entries.

Proof. Let us first recall Azuma-Hoeffding's inequality

Lemma 15. *(Azuma-Hoeffding's inequality) Suppose $M_k, k \geq 0$ is a martingale for the filtration $\mathcal{F}_k$ and $|M_k - M_{k-1}| \leq c_k$. Then for all $t \geq 0$*

$$P(M_n - M_0 \geq t) \leq \exp\{-\frac{t^2}{2\sum_{k=1}^n c_k^2}\}.$$

Proof. By Tchebychev's inequality for all $\lambda \geq 0$

$$P(M_n - M_0 \geq t) \leq e^{-\lambda t}\mathbb{E}[e^{\lambda \sum_{k=1}^n (M_k - M_{k-1})}]. \tag{14}$$

We first integrate conditionnaly to $\mathcal{F}_{n-1}$, that is control uniformly

$$f(\lambda) = \log \mathbb{E}[e^{\lambda(M_n - M_{n-1})}|\mathcal{F}_{n-1}].$$

Clearly, $f(0) = f'(0) = 0$ whereas

$$f''(\lambda) = \frac{\mathbb{E}[(M_n - M_{n-1})^2 e^{\lambda(M_n - M_{n-1})}|\mathcal{F}_{n-1}]}{\mathbb{E}[e^{\lambda(M_n - M_{n-1})}|\mathcal{F}_{n-1}]} \leq c_n^2.$$

Therefore, we have the uniform bound

$$\mathbb{E}[e^{\lambda(M_n - M_{n-1})}|\mathcal{F}_{n-1}] \leq e^{\frac{1}{2}\lambda^2 c_n^2}$$

and proceeding by induction we deduce

$$\mathbb{E}[e^{\lambda \sum_{k=1}^n (M_k - M_{k-1})}] \leq e^{\frac{1}{2}\lambda^2 \sum_{k=1}^n c_k^2}.$$

Plugging back this control into (14) and taking $\lambda = t/\sum c_k^2$ yields the lemma. □

We finally prove Lemma 12 for a continuously differentiable function f, the generalization to all functions with finite variation norm then holds by density. We then have $\|f\|_{TV} = \int |f'(x)|dx$. We apply Azuma-Hoeffding 's inequality to

$$M_k = \mathbb{E}[\frac{1}{N}\sum_{i=1}^N f(\lambda_i)|\mathcal{F}_k]$$

where $\mathcal{F}_k$ is the filtration generated by the kth first column vectors of $Y_{N,M}$ for Wishart matrices and by $\{\mathbf{X}_N(i,j), 1 \leq i \leq j \leq k\}$ for Wigner matrices. M_k is a martingale obviously and

$$M_N - M_0 = \frac{1}{N}\sum_{i=1}^N f(\lambda_i) - \mathbb{E}[\frac{1}{N}\sum_{i=1}^N f(\lambda_i)].$$

Therefore we need to bound for each $k \in \{1, \cdots, N\}$

$$\begin{aligned} M_k - M_{k-1} &= \mathbb{E}[\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i) - \mathbb{E}[\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i)|\mathcal{F}_{k-1}]] \\ &= \mathbb{E}[\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i) - \frac{1}{N}\sum_{i=1}^{N} f(\tilde{\lambda}_i)|\mathcal{F}_k]\,. \end{aligned}$$

where in the above expectation λ_i (resp $\tilde{\lambda}_i$) are the iegenvalues of $Z_N = \mathbf{X}_N$ or W_N with the same entries except for the kth vector of $Y_{N,M}$ for Wishart matrices or over $\{\mathbf{X}_{i,k}, i \leq k\}$ for Wigner matrices where we take independent copies. Hence the eigenvalues λ and $\tilde{\lambda}$ are the eigenvalues of two operators which differ at most by a rank one perturbation. Therefore, by Theorem 22, the eigenvalues λ_i and $\tilde{\lambda}_i$ are interlaced, that is if they are ordered

$$\tilde{\lambda}_{i-1} \leq \lambda_i \leq \lambda_{i+1}\,.$$

If g is increasing we deduce that

$$\sum_{i=1}^{N-2} g(\tilde{\lambda}_i) \leq \sum_{i=2}^{N-1} g(\lambda_i) \leq \sum_{i=3}^{N} g(\tilde{\lambda}_i)$$

which implies

$$|\sum_{i=1}^{N} g(\lambda_i) - \sum_{i=1}^{N} g(\tilde{\lambda}_i)| \leq 2\|g\|_\infty$$

Decomposing $f(x) - f(0)$ as the sum of the two increasing functions

$$f(x) - f(0) = \int_0^x f'(y) 1_{f'(y)\geq 0} dy - \int_0^x (-f')(y) 1_{f'(y)<0} dy$$

shows that

$$|M_k - M_{k-1}| \leq \frac{2}{N}\|f\|_{TV}$$

which allows to conclude that for all $\delta > 0$

$$P\left(\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i) - \mathbb{E}[\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i)] \geq \delta\right) \leq e^{-\frac{\delta^2 N}{8c_X}}\,.$$

The other bound is obtained by changing f into $-f$. $\square$

2.6.2 Concentration inequalities from the logarithmic Sobolev inequality

Throughout this section an integer number N will be fixed.

Definition 16. 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 $f:\mathbb{R}^N\to\mathbb{R}$,

$$\int f^2\log\frac{f^2}{\int f^2dP}dP\le 2c\int \|\nabla f\|_2^2 dP. \tag{15}$$

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 [64] or [76]. The Gaussian law, and any probability measure ν absolutely continuous with respect to the Lebesgue measure with a strictly log-concave density or satisfying the Bobkov and Götze [25] condition (including $\nu(dx)=Z^{-1}e^{-|x|^\alpha}dx$ for $\alpha\ge 2$, 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 [76], [64, Property 4.6].

Lemma 17 (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

$$\int e^{\lambda(G-E_P(G))}dP\le e^{c\lambda^2|G|_{\mathcal{L}}^2/2}, \tag{16}$$

and so for all $\delta>0$

$$P\left(|G-E_P(G)|\ge\delta\right)\le 2e^{-\delta^2/2c|G|_{\mathcal{L}}^2}. \tag{17}$$

Note that Lemma 17 also implies that E_PG is finite.
Proof of Lemma 17 We denote E_P the expectation $E_P[f]=\int f dP$. Note first that (17) follows from (16). Indeed, by Chebychev's inequality, for any $\lambda>0$,

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

Optimizing with respect to λ (by taking $\lambda=\delta/c|G|_{\mathcal{L}}^2$) yields the bound (17).

Turning to the proof of (16), let us first assume that G is a bounded differentiable function such that

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

Define

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

Then, taking $f=e^{\lambda(G-E_PG)}$ in (15), some algebra reveals that for $\lambda>0$,

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

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 (16) in case G is bounded and differentiable.

Let us now assume only that G is Lipschitz with $|G|_{\mathcal{L}} < \infty$. For $\varepsilon > 0$, define $\overline{G}_\varepsilon = G \wedge (-1/\varepsilon) \vee (1/\varepsilon)$, and note that $|\overline{G}_\varepsilon|_{\mathcal{L}} \leq |G|_{\mathcal{L}} < \infty$. Consider the regularization $G_\varepsilon(x) = p_\varepsilon * \overline{G}_\varepsilon(x) = \int \overline{G}_\varepsilon(y) p_\varepsilon(x-y) dy$ with the Gaussian density $p_\varepsilon(x) = e^{-|x|^2/2\varepsilon} dx / \sqrt{(2\pi\varepsilon)^N}$ such that $p_\varepsilon(x)dx$ converges weakly to the atomic measure δ_0 as ε converges to 0. Since for any $x \in \mathbb{R}^N$,

$$|G_\varepsilon(x) - \overline{G}_\varepsilon(x)| \leq |G|_{\mathcal{L}} \int ||y||_2 p_\varepsilon(y) dy = |G|_{\mathcal{L}} \sqrt{\varepsilon N},$$

G_ε converges pointwise to G. G_ε is also continuously differentiable and

$$\begin{aligned} |||| \nabla G_\varepsilon||_2^2||_\infty &= \sup_{x \in \mathbb{R}^M} \sup_{u \in \mathbb{R}^M} \{2\langle \nabla G_\varepsilon(x), u\rangle - ||u||_2^2\} \\ &\leq \sup_{u,x \in \mathbb{R}^M} \sup_{\delta > 0} \{2\delta^{-1}(G_\varepsilon(x+\delta u) - G_\varepsilon(x)) - ||u||_2^2\} \\ &\leq \sup_{u \in \mathbb{R}^M} \{2|G|_{\mathcal{L}}||u||_2 - ||u||_2^2\} = |G|_{\mathcal{L}}^2. \end{aligned} \tag{18}$$

Thus, we can apply the previous result to find that for any $\varepsilon > 0$ and all $\lambda \in \mathbb{R}$

$$E_P[e^{\lambda G_\varepsilon}] \leq e^{\lambda E_P G_\varepsilon} e^{c\lambda^2 |G|_{\mathcal{L}}^2/2} \tag{19}$$

Therefore, by Fatou's lemma,

$$E_P[e^{\lambda G}] \leq e^{\liminf_{\varepsilon \to 0} \lambda E_P G_\varepsilon} e^{c\lambda^2 |G|_{\mathcal{L}}^2/2}. \tag{20}$$

We next show that $\lim_{\varepsilon \to 0} E_P G_\varepsilon = E_P G$, which, in conjunction with (18), will conclude the proof. Indeed, (19) implies that

$$P(|G_\varepsilon - E_P G_\varepsilon| > \delta) \leq 2e^{-\delta^2/2c|G|_{\mathcal{L}}^2}. \tag{21}$$

Consequently,

$$\begin{aligned} E[(G_\varepsilon - E_P G_\varepsilon)^2] &= 2\int_0^\infty x P(|G_\varepsilon - E_P G_\varepsilon| > x) \, dx \leq 4 \int_0^\infty x e^{-\frac{x^2}{2c|G|_{\mathcal{L}}^2}} dx & (22) \\ &= 4c|G|_{\mathcal{L}}^2 & (23) \end{aligned}$$

so that the sequence $(G_\varepsilon - E_P G_\varepsilon)_{\varepsilon\geq 0}$ is uniformly integrable. Now, G_ε converges pointwise to G and therefore there exists a constant K, independent of ε, such that for $\varepsilon < \varepsilon_0$, $P(|G_\varepsilon| \leq K) \geq \frac{3}{4}$. On the other hand, (21) implies that $P(|G_\varepsilon - E_P G_\varepsilon| \leq r) \geq \frac{3}{4}$ for some r independent of ε. Thus,

$$\{|G_\varepsilon - E_P G_\varepsilon| \leq r\} \cap \{|G_\varepsilon| \leq K\} \subset \{|E_P G_\varepsilon| \leq K + r\}$$

is not empty, providing a uniform bound on $(E_P G_\varepsilon)_{\varepsilon<\varepsilon_0}$. We thus deduce from (23) that $\sup_{\varepsilon<\varepsilon_0} E_P G_\varepsilon^2$ is finite, and hence $(G_\varepsilon)_{\varepsilon<\varepsilon_0}$ is uniformly integrable. In particular,

$$\lim_{\varepsilon\to 0} E_P G_\varepsilon = E_P G < \infty,$$

which finishes the proof. □

Lemma 18. let $f : B \to \mathbb{R}$ such that

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

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

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

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

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

Note that, if $x \in B$, since $f(y) - f(x) - |f|_{\mathcal{L}}^B d(x,y) \leq 0$, the above supremum is taken at $y = x$ and $\tilde{f}(x) = f(x)$. Moreover, using the triangle inequality, we get that for any $x,z \in X$,

$$\begin{aligned} \tilde{f}(x) &\geq \sup_{y\in B}\{f(y) - |f|_{\mathcal{L}}^B(d(x,z) + d(z,y))\} \\ &= -|f|_{\mathcal{L}}^B d(x,z) + \tilde{f}(z) \end{aligned} \tag{24}$$

and hence $\tilde{f}$ is Lipschitz, with constant $|f|_{\mathcal{L}}^B$. Therefore, we find that

$$\mu(\{|f - \mu(f1_B)| \geq \delta\} \cap B) \leq \mu(|\tilde{f} - \mu(\tilde{f})| \geq \delta + \mu(|1_B f - \tilde{f}|))$$

Note that $\mu(|1_B f - \tilde{f}|) = \mu(1_{B^c}|\tilde{f}|)$. (24) with $z \in B$ shows that

$$|\tilde{f}(x)| \leq |f(z)| + |f|_{\mathcal{L}}^B d(z,x)$$

and so optimizing to minimize the distance gives

$$|\tilde{f}(x)| \leq \max_{z\in B}|f(z)| + |f|_{\mathcal{L}}^B d(B,x).$$

Hence,

$$\mu(|1_B f - \tilde{f}|) \leq \mu\left(1_{B^c}(\sup_{x\in B}|f(x)| + |f|_{\mathcal{L}}^B d(.,B))\right) =: \delta(f)$$

gives the desired estimate. □

2.6.3 Smoothness and convexity of the eigenvalues of a matrix and of traces of matrices

We shall not follow [65] 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 [9], rely on Weyl and Lidskii inequalities (see Theorems 20 and 23). We recall that we will denote, for $\mathbf{B} \in \mathcal{M}_N(\mathbb{C})$, $\|\mathbf{B}\|_2$ its Euclidean norm;

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

From Lidskii's Theorem 23, we will deduce that each eigenvalue of the matrix is a Lipschitz function of the entries of the matrix. We denote $\mathcal{E}_N^{(1)} = \mathbb{R}^{N(N+1)/2}$ (resp. $\mathcal{E}_N^{(2)} = \mathbb{C}^{N(N-1)/2} \times \mathbb{R}^N$) and $\mathbf{A}$ the symmetric (resp. Hermitian) $N \times N$ Wigner matrix such that $\mathbf{A} = \mathbf{A}^*; (\mathbf{A})_{ij} = A_{ij}, 1 \le i \le j \le N$ for $(A_{ij})_{1\le i\le j\le N} \in \mathcal{E}_N^{(\beta)}$, $\beta = 1$ (resp. $\beta = 2$).

Lemma 19. 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 $k \in \{1, \cdots, N\}$, all $\mathbf{A}, \mathbf{B} \in \mathcal{H}_N^{(2)}$,

$$|\lambda_k(\mathbf{A}+\mathbf{B}) - \lambda_k(\mathbf{A})| \le \|\mathbf{B}\|_2.$$

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

$$(A_{ij})_{1\le i\le j\le N} \in \mathcal{E}_N^{(2)} \to \lambda_k(\mathbf{A})$$

is Lipschitz with constant one.

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

$$(A_{ij})_{1\le i\le j\le N} \in \mathcal{E}_N^{(2)} \to \sum_{k=1}^{N} f(\lambda_k(\mathbf{A}))$$

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

$$\lim_{\varepsilon\to 0} \varepsilon^{-1} \left(\sum_{k=1}^{N} f(\lambda_k(\mathbf{A}+\varepsilon\mathbf{B})) - \sum_{k=1}^{N} f(\lambda_k(\mathbf{A})) \right) = \operatorname{tr}(f'(\mathbf{A})\mathbf{B}).$$

Let us point out that if $\mathbf{A}$ is only normal, the first result was generalized by Gersgorin.

Proof. The first inequality is a direct consequence of Lidskii's Theorem 23 and entails the same control on $\lambda_{\max}(\mathbf{A})$. For the second we only need to use Cauchy-Schwarz's inequality;

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

where we used Lidskii's Theorem 23 in the last line. For the last point, we check it for $f(x) = x^k$ where the result is clear since

$$\operatorname{tr}((\mathbf{A}+\varepsilon\mathbf{B})^k) = \operatorname{tr}(\mathbf{A}^k) + \varepsilon k \operatorname{tr}(\mathbf{A}^{k-1}\mathbf{B}) + O(\varepsilon^2) \tag{25}$$

and complete the argument by density of the polynomials. □

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

Lemma 20. Let P be a polynomial in m-non commutative indeterminates. For $1 \leq i \leq m$, 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 $(\mathbf{B}_1, \cdots, \mathbf{B}_m) \in \mathcal{H}_N^{(2)}$,

$$\lim_{\varepsilon \to 0} \varepsilon^{-1} \left(\operatorname{tr}(P(\mathbf{A}_1 + \varepsilon\mathbf{B}_1, \cdots, \mathbf{A}_m + \varepsilon\mathbf{B}_m)) - \operatorname{tr}(P(\mathbf{A}_1, \cdots, \mathbf{A}_m))\right)$$

$$= \sum_{i=1}^{m} \operatorname{tr}(D_i P(\mathbf{A}_1, \cdots, \mathbf{A}_m) \mathbf{B}_i).$$

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 $M < \infty$,

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

is Lipschitz with a Lipschitz norm bounded by $\sqrt{N} C(P, M)$ for a constant $C(P, M)$ which 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

$$\begin{aligned}&\mathrm{tr}(P(\mathbf{A}_1+\varepsilon\mathbf{B}_1,\cdots,\mathbf{A}_m+\varepsilon\mathbf{B}_m))-\mathrm{tr}(P(\mathbf{A}_1,\cdots,\mathbf{A}_m))\\&=\varepsilon\sum_{i=1}^m\sum_{P=P_1X_iP_2}\mathrm{tr}(P_1(\mathbf{A}_1,\cdots,\mathbf{A}_m)\mathbf{B}_iP_2(\mathbf{A}_1,\cdots,\mathbf{A}_m))+O(\varepsilon^2)\end{aligned}$$

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 X_i, $\sum_{i=1}^m d_i=d$ and D_iP is the sum of exactly d_i monomials of degree $d-1$. Hence, $D_iP(\mathbf{A}_1,\cdots,\mathbf{A}_m)$ has spectral radius bounded by d_iM^{d-1} when $(\mathbf{A}_1,\cdots,\mathbf{A}_m)$ are Hermitian matrices in $Ł_M^N$. Hence, by Cauchy-Schwarz's inequality, we obtain

$$\begin{aligned}|\sum_{i=1}^m\mathrm{tr}(D_iP(\mathbf{A}_1,\cdots,\mathbf{A}_m)\mathbf{B}_i)| &\le \left(\sum_{i=1}^m\mathrm{tr}(|D_iP(\mathbf{A}_1,\cdots,\mathbf{A}_m)|^2)\right)^{\frac12}\left(\sum_{i=1}^m\mathrm{tr}(\mathbf{B}_i^2)\right)^{\frac12}\\&\le\left(N\sum_{i=1}^m d_i^2M^{2(d-1)}\right)^{\frac12}\left(\sum_{i=1}^m\|\mathbf{B}_i\|_2^2\right)^{\frac12}\\&\le\sqrt{N}dM^{d-1}\left(\sum_{i=1}^m\|\mathbf{B}_i\|_2^2\right)^{\frac12}.\end{aligned}$$

□

Exercise 21. Prove that when $m=1$, $D_1P(x)=P'(x)$.

We now prove the following result originally due to Klein and which can be found for instance in [92].

Lemma 22 (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. Let $X,Y\in\mathcal{H}_N^{(2)}$. We shall show that if f is a convex continuously differentiable function

$$\mathrm{tr}\,(f(X)-f(Y))\ge\mathrm{tr}\left((X-Y)f'(Y)\right). \tag{26}$$

Taking $X=\mathbf{A}$ or $X=\mathbf{B}$ and $Y=2^{-1}(\mathbf{A}+\mathbf{B})$ and summing the two resulting inequalities shows that for any couple $\mathbf{A},\mathbf{B}$ of $N\times N$ Hermitian matrices,

$$\mathrm{tr}\left(f(\frac12\mathbf{A}+\frac12\mathbf{B})\right)\le\frac12\mathrm{tr}\,(f(\mathbf{A}))+\frac12\mathrm{tr}\,(f(\mathbf{B}))$$

which implies that $(A_{ij})_{1\le i\le j\le N}\to \mathrm{tr}(f(\mathbf{A}))$ is convex. The result follows for general convex functions f by approximations.

To prove (26), let us denote $\lambda_i(C)$ the eigenvalues of a Hermitian matrix C and $\xi_i(C)$ the associated eigenvector and write

$$\begin{aligned}\langle \xi_i(X),(f(X)-f(Y))\xi_i(X)\rangle &= f(\lambda_i(X))-\sum_{j=1}^{N}|\langle \xi_i(X),\xi_j(Y)\rangle|^2 f(\lambda_j(Y))\\ &= \sum_{j=1}^{N}|\langle \xi_i(X),\xi_j(Y)\rangle|^2 (f(\lambda_i(X))-f(\lambda_j(Y)))\\ &\ge \sum_{j=1}^{N}|\langle \xi_i(X),\xi_j(Y)\rangle|^2 (\lambda_i(X)-\lambda_j(Y))f'(\lambda_j(Y))\end{aligned}$$

where we have used the convexity of f to write $f(x)-f(y)\ge (x-y)f'(y)$. The right hand side of the last inequality just coincides with the right hand side of (26), which completes the first part of the proof of the lemma.

We give an other proof below, which 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 (25) one step further gives

$$\begin{aligned}\mathrm{tr}((\mathbf{A}+\varepsilon\mathbf{B})^k) &= \mathrm{tr}(\mathbf{A}^k)+\varepsilon\sum_{k=0}^{p-1}\mathrm{tr}(\mathbf{A}^k\mathbf{B}\mathbf{A}^{p-1-k})\\ &\quad +\varepsilon^2\sum_{0\le k+l\le p-2}\mathrm{tr}(\mathbf{A}^k\mathbf{B}A^l\mathbf{B}A^{p-2-k-l})+O(\varepsilon^3)\\ &= \mathrm{tr}(A^k)+\varepsilon p\,\mathrm{tr}(A^{p-1}\mathbf{B})+\frac{\varepsilon^2}{2}p\sum_{0\le l\le p-2}\mathrm{tr}(A^l\mathbf{B}A^{p-2-l}\mathbf{B})+O(\varepsilon^3)\end{aligned}\tag{27}$$

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 (27) reads

$$p\sum_{0\le l\le p-1}\mathrm{tr}(\mathbf{A}^l\mathbf{B}\mathbf{A}^{p-2-l}\mathbf{B})=\langle g_{x^p}(\mathbf{A},\mathbf{A}),\mathbf{B}\otimes\mathbf{B}\rangle \tag{28}$$

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 $\langle \mathbf{B},\mathbf{D}\rangle_2 = \sum_{i,j=1}^N B_{ij}\overline{D}_{ij}$. 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 (27) and (28), 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 $g_f\geq c$ when $f''\geq c$ we finally have proved

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

The proof is thus complete. □

Let us also notice that

Lemma 23. Assume $\lambda_1(\mathbf{A})\leq\lambda_2(\mathbf{A})\cdots\leq\lambda_N(\mathbf{A})$. 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 $\mathcal{M}_N^{(2)}$, $(A_{ij})_{1\leq i,j\leq N}\to\|\mathbf{A}\|$ is convex.

Proof. The first result is clear since we have already seen that $\lambda_N(\mathbf{A}+\mathbf{B})\leq\lambda_N(\mathbf{A})+\lambda_N(\mathbf{B})$. 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 $(A_{ij})_{1\leq i,j\leq N}\to\|\mathbf{A}\|$ is due to the definition of the norm. □

2.6.4 Concentration inequalities for random matrices

Concentration inequalities for the eigenvalues of random matrices

We consider a Hermitian random matrix $\mathbf{A}$ whose real or complex entries have joint law μ^N which satisfies one of the two hypotheses below.

Either the entries of $\mathbf{A}$ are independent and satisfy for some $c>0$ the following condition (H1).

• (H1)$\mathbf{A} = \mathbf{X}^N/\sqrt{N} = (\mathbf{A})^*$ *with* $(\mathbf{X}^N_{ij}, 1 \le i \le j \le N)$ *independent, with law* μ^N_{ij}, *which are probability measures on* $\mathbb{C}$ *or* $\mathbb{R}$*, all of them satisfying the log-Sobolev inequality with constant* $c < \infty$.
Or μ^N is a Gibbs measure with strictly convex potential, i.e satisfies (H2) below.

• (H2) *there exists a strictly convex twice continuously differentiable function* V : $\mathbb{R} \to \mathbb{R}$, $V''(x) \ge \frac{1}{c} > 0$, *so that*

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

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

Note that when $V = \frac{1}{2}x^2$, μ^N is the law of a Gaussian Wigner matrix but in any other case the entries of $\mathbf{A}$ with law μ^N are not independent.

We can now state the following theorem.

Theorem 6. *Suppose there exists* $c > 0$ *so that either (H1) or (H2) holds. Then:*

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

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

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

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

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

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

Proof of Theorem 6 For (H2), the assumption $V''(x) \ge \frac{1}{c}$ implies, by Lemma 22, that $(A_{ij})_{1\le i\le j\le N} \in \mathcal{E}_N^{(\beta)} \to N\mathrm{tr}(V(\mathbf{A}))$ is twice continuously differentiable with Hessian bounded below by $\frac{N}{c}$. The second case uses the product property of Lemma **??** 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 19, $G(A^N_{ij}, 1 \le i \le j \le N) = \mathrm{tr}(f(\mathbf{A}))$ 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 23. $\square$

Exercise 24. State the concentration result when the μ^N_{ij} only satisfy Poincaré inequality.

Exercise 25. 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 $\mathbf{A}$ satisfies a concentration of measure inequality.

When the laws satisfy rather Talagrand's type condition we state the induced concentration bounds

Theorem 7. *Let $\mu^N(f(\mathbf{A})) = \int f(\mathbf{X}/\sqrt{N}) \prod d\mu^N_{i,j}(X_{ij})$ with $(\mu^N_{i,j}, i \le j)$ compactly supported probability measures on a connected compact subset K of $\mathbb{C}$. Fix $\delta_1 = 8|K|\sqrt{\pi}$. Then, for any $\delta \ge \delta_1 N^{-1}$, for any convex function f,*

$$\mu^N\left(|tr(f(\mathbf{A})) - \mu^N[tr(f(\mathbf{A}))]| \ge N\delta|f|_{\mathcal{L}}\right) \tag{29}$$

$$\le \frac{32|K|}{\delta} \exp\left(-N^2 \frac{1}{16|K|^2 a^2} \frac{(\delta - \delta_1 N^{-1})^2}{16|K|}\right).$$

If $\lambda_{max}(\mathbf{A})$ is the largest (or smallest) eigenvalue of $\mathbf{A}$, or the spectral radius of $\mathbf{A}$, for $\delta \ge \delta_1(N)$,

$$\mu^N\left(|\lambda_{max}(\mathbf{A}) - E^N[\lambda_{max}(\mathbf{A})]| \ge \delta N^{\frac{1}{2}}\right) \le \frac{32|K|}{\delta} \exp\left(-\frac{1}{16|K|^2 a^2} \frac{(\delta - \delta_1 N^{-\frac{1}{2}})^2}{16|K|}\right).$$

Proof. Applying Corollary **??**, Lemmas 19 and 22 with $f : (\mathbf{A} \to \mathrm{tr}(f(\mathbf{A}))$ which is Lipschitz with Lipschitz constant $|f|_{\mathcal{L}}$ provides the first bound. □

Observe that the speed of the concentration we obtained is optimal for $\mathrm{tr}(f(\mathbf{X}^N))$ (since it agrees with the speed of the central limit theorem). It is also optimal in view of the large deviation principle we will prove in the next section. However, it does not capture the true scale of the fluctuations of $\lambda_{\max}(\mathbf{A})$ which are of order $N^{-\frac{1}{3}}$. Improvements of concentration inequalities in that direction were obtained by M. Ledoux [77].

We emphasize that Theorem 6 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 $|a|^2 c$).

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 **??**). 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 variables. Assume that for any $N \in \mathbb{N}$,

$$\varphi^N_V : ((A_k)_{ij})_{\substack{i \le j \\ 1 \le k \le m}} \in \mathcal{E}^{(2)}_N \to \mathrm{tr} V(\mathbf{A}_1, \cdots, A_m)$$

is convex. Let c be a positive real.

$$d\mu^{N,\beta}_V(\mathbf{A}_1, \cdots, \mathbf{A}_m) := \frac{1}{Z^N_V} e^{-N\mathrm{tr}(V(\mathbf{A}_1,\cdots,\mathbf{A}_m))} d\mu^{N,\beta}_c(\mathbf{A}_1) \cdots d\mu^{N,\beta}_c(\mathbf{A}_m)$$

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 variance $1/cN$, 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

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

We then have the following corollary.

Corollary 26. *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 $\Lambda_M^N = \{\mathbf{A}_i \in \mathcal{H}_N^{(2)}; max_{1 \le i \le m} \lambda_{\max}(A_i) \le M\}$ and P be a monomial of degree $d \in \mathbb{N}$. Then, for any $\delta > 0$*
$$\mu_V^{N,\beta}\left(\{|tr(P(\mathbf{A}_1, \cdots, \mathbf{A}_m)) - \mu_V^{N,\beta}(tr(P(\mathbf{A}_1, \cdots, \mathbf{A}_m) 1_{\Lambda_M^N}))| > \delta + \delta(M,N)\} \cap \Lambda_M^N\right)$$
$$\le 2e^{-\frac{c\delta^2}{d^2 M^{2(d-1)}}}$$
with
$$\delta(M,N) \le M^d \mu_V^{N,\beta}\left((1 + d\|\mathbf{A}\|_2) 1_{(\Lambda_M^N)^c}\right).$$

Proof. By assumption, the law $\mu_V^{N,\beta}$ of the entries of $(\mathbf{A}_1, \cdots, \mathbf{A}_m)$ is absolutly continuous with respect to Lebesgue measure. The Hessian of the logarithm of the density is bounded above by $-NcI$. Hence, by Corollary **??**, $\mu_V^{N,\beta}$ satisfies a log Sobolev inequality with constant $1/Nc$ and thus by Lemma 17 we find that $\mu_V^{N,\beta}$ satisfies the first statement of the Corollary. We finally conclude by using Lemma **??** and the fact that $X_1, \cdots, X_m \to \mathrm{tr}(P(X_1, \cdots, X_m))$ is locally Lipschitz by Lemma 20. □

2.7 Brascamp-Lieb inequalities; Applications to random matrices

We introduce first Brascamp-Lieb inequalities and show how they can be derived from results from optimal transport theory, following a proof of Hargé [71]. We then show how these inequalities 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 chapter.

2.7.1 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 8. (Brascamp-Lieb [30], Hargé [71, 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:*

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

where

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

This theorem was proved by Brascamp and Lieb [30, Theorem 7] (case $g(x) = |x_1|^{\alpha}$), by Caffarelli [37, Corollary 6] (case $g(x) = g(x_1)$) and then for a general convex function g by Hargé [71]. Hargé followed the idea introduced by Caffarelli to use optimal transport of measure. We can unfortunately not develop the theory of optimal transport here but shall still provide Hargé's proof (which is based, as for the proof of log Sobolev inequalities, on the use of a semi-group which interpolates between the two measures of interest) as well as the statement of the results in optimal transport theory that the proof requires. For more information on the later, we refer the reader to the two survey books of Villani [100, 101].

We shall denote $d\mu(x) = f(x)d\gamma(x)/\int f d\gamma$.

Brenier [31] (see also Mc Cann [80]) has shown that there exists a convex function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that

$$\int g(y)d\mu(y) = \int g(\nabla\varphi(x))d\gamma(x).$$

In other words, μ can be realized as the image (or push forward) of μ by the map $\nabla\varphi$.

Caffarelli [36, 35] then proved that if the density f is Hölder continuous with exponent $\alpha \in]0,1[$, φ is $\mathcal{C}^{2,\alpha}$ for any $\alpha \in]0,1[$ (i.e twice continuously differentiable with a second derivative Hölder continuous with exponent α). Moreover, by Caffarelli [37, Theorem 11], we know (and here we need to have γ, μ as specified above to get the upper bound) that for any vector $e \in \mathbb{R}^n$,

$$0 \leq \partial_{ee}\varphi = < \mathrm{Hess}(\varphi)e, e > \leq 1.$$

We now start the proof of Theorem 8. Observe first that we can assume without loss of generality that γ is the law of independent centered Gaussian variables with variance one (up to a linear transformation on the x's).

We let $\psi(x) = -\varphi(x) + \frac{1}{2}\|x\|_2^2$ so that $0 \leq \mathrm{Hess}(\psi) = \leq I$ (with I the identity matrix and where inequalities hold in the operator sense) and write

$$\int g(y)d\mu(y) = \int g(x - \nabla\psi(x))d\gamma(x).$$

The idea is then to consider the following interpolation

$$\theta(t) = \int g(x - P_t(\nabla\psi)(x))d\gamma(x)$$

with P_t the Ornstein-Uhlenbeck process given, for $h : \mathbb{R}^n \to \mathbb{R}$ by

$$P_t h(x) = \int h(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}y)d\gamma(y)$$

and $P_t(\nabla\psi) = (P_t(\nabla_1\psi), \cdots, P_t(\nabla_n\psi))$ with $\nabla_i\psi = \partial_{x_i}\psi$. Note that for a Lispshitz function h, for all $x \in \mathbb{R}^n$,

$$\begin{aligned}|P_t h(x) - h(x)| &\leq \int |h(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}y) - h(x)|d\gamma(y) \\ &\leq |h|_{\mathcal{L}}(\sqrt{1-e^{-t}} + (1-e^{-\frac{t}{2}}))\int (\|x\|_2 + \|y\|_2)d\gamma(y)\end{aligned}$$

goes to zero as t goes to zero (since $\int \|x\|_2 d\gamma(x) < \infty$). Similarly, for $t > 1$, there is a finite constant C such that

$$|P_t h(x) - \int h d\gamma| \leq C|h|_{\mathcal{L}} e^{-\frac{t}{2}}(\|x\|_2 + \int \|y\|_2 d\gamma(y))$$

which shows that $P_t h$ goes to $\int h d\gamma$ as t goes to infinity. Since ψ is twice continuously differentiable with Hessian bounded by one, each $\nabla_i\psi$, $1 \leq i \leq n$, has uniformly bounded derivatives (by one) and so is Lipschitz for the Euclidean norm (with norm bounded by $\sqrt{n}$). Hence, the above applies with $h = \nabla_i\psi$, $1 \leq i \leq n$.

Let us assume that g is smooth and ∇g is bounded. Then, we deduce from the above estimates that, again because $\int \|x\|_2 d\gamma(x)$ is finite,

$$\lim_{t\to 0}\theta(t) = \theta(0) = \int g(x - \nabla\psi(x))d\gamma(x) = \int g(x)d\mu(x), \lim_{t\to\infty}\theta(t) = \int g(x - \int \nabla\psi d\gamma)d\gamma(x).$$

Since

$$\int \nabla\psi d\gamma = \int (\nabla\psi - x)d\gamma + \int x d\gamma = \int x d\gamma - \int x d\mu$$

we see that Theorem 8 is equivalent to prove that $\theta(0) \leq \theta(\infty)$ and so it is enough to show that θ is non decreasing. But, $t \to \theta(t)$ is differentiable with derivative

$$\theta'(t) = -\int < \nabla g(x - P_t(\nabla\psi)(x)), \partial_t P_t(\nabla\psi)(x) > d\gamma(x) \tag{30}$$

with

$$\begin{aligned}\partial_t P_t(h)(x) &= \int < -\frac{1}{2}e^{-\frac{t}{2}}x + \frac{1}{2}e^{-t}(1-e^{-t})^{-\frac{1}{2}}y, \nabla h(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}y) > d\gamma(y) \\ &= -\frac{1}{2}e^{-\frac{t}{2}} < x, P_t(\nabla h)(x) > + \frac{1}{2}e^{-t}\int \Delta h(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}y)d\gamma(y) \\ &= -\frac{1}{2} < x, \nabla P_t h(x) > + \frac{1}{2}\Delta(P_t h)(x) := L(P_t h)(x)\end{aligned}$$

where in the second line we integrated by parts under the standard Gaussian law γ. Note also, again by integration by parts, that

$$\int h_1 L h_2 d\gamma(x) = -\frac{1}{2}\int < \nabla h_1, \nabla h_2 > d\gamma.$$

Hence, (30) implies

$$\begin{aligned}
\theta'(t) &= -\int \sum_{i=1}^{n} (\partial_i g)(x - P_t(\nabla\psi)) L P_t(\partial_i \psi) d\gamma \\
&= \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\int \partial_j((\partial_i g)(x - P_t(\nabla\psi)))\partial_j(P_t(\partial_i\psi)))d\gamma \\
&= \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\sum_{k=1}^{n}\int (1_{k=j} - \partial_j(P_t(\partial_k\psi)))(\partial_k\partial_i g)(x - P_t(\nabla\psi))\partial_j(P_t(\partial_i\psi))d\gamma.
\end{aligned}$$

Thus, if we let

$$M_{ij}(x) := \partial_j(P_t(\partial_i\psi))(x), \text{ and } C_{ij}(x) = (\partial_j\partial_i g)(x - P_t(\nabla\psi)),$$

we have written, with $I_{ij} = 1_{i=j}$ the identity matrix,

$$\begin{aligned}
\theta'(t) &= \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\sum_{k=1}^{n}\int (I - M(x))_{kj} C_{ik}(x) M_{ij}(x) d\gamma(x) \\
&= \frac{1}{2}\int \mathrm{tr}(C(x)(I - M(x))M^*(x))d\gamma(x) \geq 0
\end{aligned}$$

since by Caffarelli we know that $0 \leq M(x) \leq I$ for all x, whereas $C \geq 0$ by hypothesis.

This completes the proof for smooth g with bounded gradient. The generalization to all convex function g is easily done by approximation. The function can indeed be assumed as smooth as wished, since we can always restrict first the integral to a large ball $B(0,R)$, then on this large ball use Stone-Weierstrass theorem to approximate g by a smooth function, and extend again the integral. We can assume the gradient of g bounded by approximating g by

$$g_R(x) = \sup_{y\in B(0,R)} \{g(y) + < \nabla g(y), x - y >\}.$$

g_R is convex and with bounded gradient. Moreover, since $g(x) \geq g(y) + < \nabla g(y), x - y >$ by convexity of g, $g_R = g$ on $B(0,R)$, while $g(0) + < \nabla g(0), x > \leq g_R(x) \leq g(x)$ shows that $g_R, R \geq 0$ is uniformly integrable so that we can use dominated convergence theorem to show that the expectation of g_R converges to that of g.

2.7.2 Applications of Brascamp-Lieb inequalities

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 which is absolutly 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

$$d\mu_V^{N,\beta}(\mathbf{A}_1,\cdots,\mathbf{A}_m) := \frac{1}{Z_V^N} e^{-N\mathrm{tr}(V(\mathbf{A}_1,\cdots,\mathbf{A}_m))} d\mu_c^{N,\beta}(\mathbf{A}_1)\cdots d\mu_c^{N,\beta}(\mathbf{A}_m)$$

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 $\mathbf{A}$ with entries with law

$$\mu_c^{N,\beta}(d\mathbf{A}) = \frac{1}{Z_N^c} e^{-\frac{cN}{2}\mathrm{tr}(\mathbf{A}^2)} d\mathbf{A}$$

with $d\mathbf{A} = \prod_{i\le j} d\Re(A_{ij}) \prod_{i\le j} d(\Im A_{ij})$ when $\beta=2$ and $d\mathbf{A} = \prod_{i\le j} dA_{ij}$ if $\beta=1$.

We assume that V is convex in the sense that for any $N\in\mathbb{N}$,

$$(A_{ij})_{1\le i\le j\le N} \in \mathcal{E}_N^{(\beta)} \to \mathrm{tr}(V(\mathbf{A}_1,\cdots,\mathbf{A}_m))$$

is real valued and convex.

This hypothesis is satisfied if $V(X_1,\cdots,X_m) = \sum_{i=1}^k V_i(\sum_{j=1}^m \alpha_j^i X_j)$ when α_j^i are real variables and V_i are convex functions on $\mathbb{R}$ by Klein's Lemma 22.

Theorem 8 implies that for all convex function g on $(\mathbb{R})^{\beta mN(N-1)/2+mN}$,

$$\int g(\mathbf{A}-\mathbf{M}) d\mu_V^{N,\beta}(\mathbf{A}) \le \int g(\mathbf{A}) \prod_{i=1}^m d\mu_c^{N,\beta}(\mathbf{A}_i) \tag{31}$$

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 (31), $g(\mathbf{A})$ is a shorthand for a function of the (real and imaginary parts of the) entries of the matrices $\mathbf{A} = (\mathbf{A}_1,\cdots,\mathbf{A}_m)$.

By different choices of the function g we shall now obtain some a priori bounds on the random matrices $(\mathbf{A}_1,\cdots,\mathbf{A}_m)$ with law $\mu_c^{N,\beta}$.

Lemma 27. Assume that there exists $d>0$ such that for some finite $c(V)$,

$$V(X_1,\cdots,X_m) \le c(V)(1+\sum_{i=1}^m X_i^{2d}).$$

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

$$\limsup_N \mu_V^{N,\beta}(\frac{1}{N}\mathrm{tr}(\mathbf{A}_i^{2n})) \le C_0^n.$$

Moreover, C_0 depends continuously on $V(0), D_iV(0), c(V)$ 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 $[-\sqrt{C_0}, \sqrt{C_0}]$.
Proof. Let k be in $\{1, \cdots, m\}$. As $\mathbf{A} \to \mathrm{tr}(\mathbf{A}_k^{4d})$ is convex by Klein's lemma 22, Brascamp-Lieb inequality (31) implies that

$$\mu_V^{N,\beta}(\frac{1}{N}\mathrm{tr}(\mathbf{A}_k - \mathbf{M}_k)^{4d}) \le \mu_c^{N,\beta}(\frac{1}{N}\mathrm{tr}(\mathbf{A}_k)^{4d}) = \mu_c^{N,\beta}(\mathbf{L}_{\mathbf{A}_k}(x^{4d})) \tag{32}$$

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

$$\mathbf{M}_k = \mu_V^{N,\beta}[\mathbf{A}_k] = U \mu_V^{N,\beta}[\mathbf{A}_k] U^* \Rightarrow \mathbf{M}_k = \mu_V^{N,\beta}(\frac{1}{N}\mathrm{tr}(\mathbf{A}_k)) I. \tag{33}$$

Let us bound $\mu_V^{N,\beta}(\frac{1}{N}\mathrm{tr}(\mathbf{A}_k))$. Jensen's inequality implies

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

and so

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

According to Theorem 4, $\mu_c^{N,\beta}(\frac{1}{N}\mathrm{tr}(V))$ converges as N goes to infinity for any polynomial function V. Moreover, the limit, evaluated at a monomial $q = X_{\ell_1} \cdots X_{\ell k}$, is given by $\sigma^m(q(X/\sqrt{c}))$ which is bounded by $(2/\sqrt{c})^k$ according to Theorem 4. Hence,

$$|\sigma^m(V)| \le c(V)((2/\sqrt{c})^{2d} + 1) := C(V)/2.$$

Thus, for N sufficiently large, $Z_N^V \ge e^{-N^2 C(V)}$.

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

$$\mathrm{tr}(V(\mathbf{A})) \ge \mathrm{tr}(V(0) + \sum_{i=1}^m D_i V(0) \mathbf{A}_i)$$

with D_i the cyclic derivative introduced in Lemma 20. By Chebyshev's inequality, we therefore obtain, for all $\lambda \ge 0$,

$$\begin{aligned}
&\mu_V^{N,\beta}(|\mathbf{L}_{\mathbf{A}_k}(x)| \ge y) \le \mu_V^{N,\beta}(\mathbf{L}_{\mathbf{A}_k}(x) \ge y) + \mu_V^{N,\beta}(-\mathbf{L}_{\mathbf{A}_k}(x) \ge y) \\
\le\ & e^{N^2(C(V) - V(0) - \lambda y)} \left(\mu_c^{N,\beta}(e^{-N\mathrm{tr}(\sum_{i=1}^m D_i V(0)\mathbf{A}_i - \lambda \mathbf{A}_k)}) + \mu_c^{N,\beta}(e^{-N\mathrm{tr}(\sum_{i=1}^m D_i V(0)\mathbf{A}_i + \lambda \mathbf{A}_k)}) \right) \\
=\ & e^{N^2(C(V) - V(0) - \lambda y)} e^{\frac{N}{2c}\sum_{\ell \ne k} \mathrm{tr}(D_i V(0)^2)} \left(e^{\frac{N}{2c}\mathrm{tr}((D_k V(0) - \lambda)^2)} + e^{\frac{N}{2c}\mathrm{tr}((D_k V(0) + \lambda)^2)} \right).
\end{aligned}$$

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

$$\mu_V^N\left(|\mathbf{L}_{\mathbf{A}_k}(x)| \ge y\right) \le e^{BN^2 - \frac{N^2 c}{4} y^2}$$

so that for N large enough,

$$
\begin{aligned}
\mu_V^{N,\beta}(|\mathbf{L}_{\mathbf{A}_k}(x)|) &= \int_0^\infty \mu_V^N\left(|\mathbf{L}_{\mathbf{A}_k}(x)|\geq y\right) dy \\
&\leq 4\sqrt{c^{-1}B} + \int_{y\geq 4\sqrt{c^{-1}B}} e^{-\frac{N^2 c}{4}(y^2-4\frac{B}{c})} dy \leq 8\sqrt{Bc^{-1}}. \qquad (34)
\end{aligned}
$$

This, with (32), completes the proof. □

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 $\mathbf{A}_1,\ldots,\mathbf{A}_m$ under the law $\mu_V^{N,\beta}$.

Lemma 28. 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,

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

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

Proof. The spectral radius $\lambda_{\max}^N(\mathbf{A}) = \max_{1\leq i\leq m} \sup_{\|u\|_2=1} < u, \mathbf{A}_i\mathbf{A}_i^* u >^{\frac{1}{2}}$ is a convex function of the entries (see Lemma 23), so we can apply Brascamp-Lieb inequality (31) to obtain that for all $s \in [0, \frac{c}{10}]$,

$$\int e^{sN\lambda_{\max}^N(\mathbf{A}-\mathbf{M})} d\mu_V^{N,\beta}(\mathbf{A}) \leq \int e^{sN\lambda_{\max}^N(\mathbf{A})} d\mu_c^{N,\beta}(\mathbf{A}).$$

But, by Theorem 6 applied with a quadratic potential V, we know that

$$
\begin{aligned}
\int e^{sN\lambda_{\max}^N(\mathbf{A})} d\mu_c^{N,\beta}(\mathbf{A}) &\leq e^{sN\mu_c^{N,\beta}(\lambda_{\max}^N)} \int e^{sN(\lambda_{\max}^N - \mu_c^{N,\beta}(\lambda_{\max}^N))} d\mu_c^{N,\beta} \\
&= sNe^{sN\mu_c^{N,\beta}(\lambda_{\max}^N)} \int_{-\infty}^\infty e^{sNy} \mu_c^{N,\beta}\left(\lambda_{\max}^N - \mu_c^{N,\beta}(\lambda_{\max}^N) \geq y\right) dy \\
&\leq sNe^{sN\mu_c^{N,\beta}(\lambda_{\max}^N)} (1 + 2\int_0^\infty e^{sNy} e^{-\frac{Nc}{4}y^2} dy) \\
&\leq \sqrt{2\pi} sNe^{sN\mu_c^{N,\beta}(\lambda_{\max}^N)} (1 + 2e^{\frac{2s^2 N}{c}})
\end{aligned}
$$

Hence, since $\mu_c^{N,\beta}(\lambda_{\max}^N)$ is uniformly bounded (by using finite bounds in section 3.2 for instance), we deduce that for all $s \geq 0$, there exists a finite constant $C(s)$ such that

$$\int e^{sN\lambda_{\max}^N(\mathbf{A}-\mathbf{M})} d\mu_V^{N,\beta}(\mathbf{A}) \leq C(s)^N.$$

By (33) and (34), we know that

$$\lambda_{\max}^N(\mathbf{A}) \leq \lambda_{\max}^N(\mathbf{A}-\mathbf{M}) + \lambda_{\max}^N(\mathbf{M}) \leq \lambda_{\max}^N(\mathbf{A}-\mathbf{M}) + 8\sqrt{Bc^{-1}}$$

from which we deduce that $\int e^{sN\lambda_{\max}^N(\mathbf{A})} d\mu_V^{N,\beta}(\mathbf{A}) \leq C^N$ for a positive finite constant C. We conclude by a simple application of Chebyshev's inequality. □

Lemma 29. If $c > 0$, $\varepsilon \in]0, \frac{1}{2}[$, then there exists $C = C(c,\varepsilon) < \infty$ such that for all $d \le N^{\frac{1}{2}-\varepsilon}$,

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

Note that this control could be generalized to $d \le N^{2/3-\varepsilon}$, by using the refinements obtained by Soshnikov in [95, Theorem 2 p.17] 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 to insure that

$$\mu_V^{N,\beta}\left(|\lambda_{\max}^N(\mathbf{A} - \mu_V^{N,\beta}(\mathbf{A}))|^d\right) \le \mu_c^{N,\beta}\left(|\lambda_{\max}^N(\mathbf{A} - \mu_c^{N,\beta}(\mathbf{A}))|^d\right).$$

Now, we have seen in the proof of Lemma 27 that $\mu_V^{N,\beta}(\mathbf{A})$ has a uniformly bounded spectral radius, say by x.Hence, we find that

$$\mu_c^{N,\beta}\left(|\lambda_{\max}^N(\mathbf{A})|^{N^{\frac{1}{2}-\varepsilon/2}}\right) \le c(\varepsilon)\frac{N(2c^{-1})^{N^{\frac{1}{2}-\varepsilon/2}}}{\sqrt{\pi N^{3(\frac{1}{2}-\varepsilon/2)}}}.$$

Applying Jensen's inequality we therefore get, for $d \le N^{\frac{1}{2}-\varepsilon}$,

$$\mu_c^{N,\beta}\left(|\lambda_{\max}^N(\mathbf{A})|^d\right) \le c'(\varepsilon)(2c^{-1})^d.$$

Hence,

$$\mu_V^{N,\beta}\left(|\lambda_{\max}^N(\mathbf{A})|^d\right)^{\frac{1}{d}} \le x + c'(\varepsilon)^{\frac{1}{d}}2c^{-1}$$

which proves the claim. □

2.7.3 Coupling concentration inequalities and Brascamp-Lieb inequalities

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

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

We let

$$\tilde{\delta}^N(P) := \mathrm{tr}(P(\mathbf{A}_1,\cdots,\mathbf{A}_m)) - \mu_V^{N,\beta}\left(\mathrm{tr}(P(\mathbf{A}_1,\cdots,\mathbf{A}_m))\right).$$

Then, we have

Lemma 30. For all N in $\mathbb{N}$, all $M > 0$, there exists a finite constant $C(P,M)$ and $\varepsilon(P,M,N)$ such that for any $\varepsilon > 0$,

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

If P is a monomial of degree d we can choose

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

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

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

with C the constant of Lemma 29.

Proof. It is enough to consider the case where P is a monomial. By Corollary 26, we only need to control $\varepsilon(P,M,N)$.

$$\begin{aligned}
\varepsilon(P,M,N) &\leq \mu_V^{N,\beta}\left(1_{(\Lambda_M^N)^c}\left(|\mathrm{tr}(P)| + dM^{d-1}\sqrt{\sum_{i=1}^m \mathrm{tr}(\mathbf{A}_i\mathbf{A}_i^*)} + \sup_{\mathbf{A}\in\Lambda_M^N}|\mathrm{tr}(P(\mathbf{A}))|\right)\right) \\
&\leq N\mu_V^{N,\beta}\left(1_{(\Lambda_M^N)^c}\left(|\lambda_{\max}^N(\mathbf{A})|^d + \sqrt{C(P,M)}|\lambda_{\max}^N(\mathbf{A})|^2 + M^d\right)\right)
\end{aligned}$$

Now,

$$\mu_V^N\left(1_{(\Lambda_M^N)^c}|\lambda_{\max}^N(\mathbf{A})|^d\right) \leq \mu_V^N\left(1_{(\Lambda_M^N)^c}\right)^{\frac{1}{2}}\mu_V^N\left(|\lambda_{\max}^N(\mathbf{A})|^{2d}\right)^{\frac{1}{2}} \leq C^d e^{-\frac{\alpha}{2}NM}$$

where we used Lemmas 28 and 29. By the previous control on $C(P,M)$, we get

$$\varepsilon(P,M,N) \quad \leq \quad 3dN(CM)^{d+1}e^{-\frac{\alpha}{2}NM}.$$

Thus for $d \leq N^{\frac{1}{2}-\varepsilon}$ and M large enough, we obtain the announced bound. $\square$

For later purposes, we have to find a control on the variance of $\mathbf{L}$.

Lemma 31. For any $c > 0$ and $\varepsilon \in]0, \frac{1}{2}[$, there exists $B, C, M_0 > 0$ such that for all $\mathbf{t} \in \mathbf{B}_{\eta,\mathbf{c}}$, all $M \geq M_0$, and monomial P of degree less than $N^{\frac{1}{2}-\varepsilon}$,

$$\mu_V^{N,\beta}\left((\tilde{\delta}^N(P))^2\right) \leq BC(P,M) + C^{2d}N^4 e^{-\frac{\alpha MN}{2}}. \tag{35}$$

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

Proof. If P is a monomial of degree d, we write

$$\mu_V^{N,\beta}((\tilde{\delta}^N(P))^2) \leq \mu_V^{N,\beta}(\mathbf{1}_{Ł_M^N}(\tilde{\delta}^N(P))^2) + \mu_V^{N,\beta}(\mathbf{1}_{(Ł_M^N)^c}(\tilde{\delta}^N(P))^2) = I_1 + I_2. \tag{36}$$

For I_1, the previous Lemma implies that, for $d \leq N$,

$$\begin{aligned}
I_1 &= 2\int_0^\infty x\mu_V^{N,\beta}\left(\{|\mathrm{tr}(P) - \mu_V^{N,\beta}(\mathrm{tr}(P))| \geq x\} \cap Ł_M^N\right)dx \\
&\leq \varepsilon(P,N,M)^2 + 4\int_0^\infty xe^{-\frac{cx^2}{2C(P,M)}}dx \leq BC(P,M)
\end{aligned}$$

with a constant B which depends only on c. For the second term, we take $M \geq M_0$ with M_0 as in Lemma 28 (Exponential tail of the largest eigenvalue) to get

$$I_2 \leq \mu_V^{N,\beta}[(Ł_M^N)^c]^{\frac{1}{2}} \mu_V^{N,\beta}((\tilde{\delta}^N(P))^4)^{\frac{1}{2}} \leq e^{-\frac{\alpha MN}{2}} \mu_V^{N,\beta}((\tilde{\delta}^N(P))^4)^{\frac{1}{2}}.$$

By Cauchy-Schwartz inequality, we obtain the control

$$\mu_V^{N,\beta}[\tilde{\delta}^N(P)^4] \leq 2^4 \mu_V^{N,\beta}((\mathrm{tr}(P))^4).$$

Now, by non-commutative Hölder's inequality Theorem 24,

$$[\mathrm{tr}(P)]^4 \leq N^4 \max_{1 \leq i \leq m} \frac{1}{N} \mathrm{tr}(A_i^{4d})$$

so that we obtain the bound

$$\mu_V^{N,\beta}[\tilde{\delta}^N(P)^4] \leq 2^4 N^4 \max_{1 \leq i \leq m} \mu_V^{N,\beta}[\frac{1}{N}\mathrm{tr}(A_i^{4d})].$$

By Lemma 29, for $d \leq N^{\frac{1}{2}-\varepsilon}$,

$$\mu_V^{N,\beta}[\frac{1}{N}\mathrm{tr}(A_i^{4d})] \leq C^{2d}. \tag{37}$$

Plugging back this estimate into (36), we have proved that for N and M sufficiently large , all monomials P of degree $d \leq N^{\frac{1}{2}-\varepsilon}$, all $\mathbf{t} \in \mathbf{B}_{\eta,\mathbf{c}}$

$$\mu_V^{N,\beta}\left((\hat{\delta}^N(P))^2\right) \leq BC(P,M) + C^{2d} N^4 e^{-\frac{\alpha MN}{2}}$$

with a finite constant C depending only on ε, c and M_0. □

Bibliographical notes The basic notions of concentration inequalities can be found in [12]. We also used the lecture notes [64]. Concentration inequalities theory was first applied to traces of functions of random matrices in [65]. Concentration for the eigenvalues themselves were studied in [4] and [81]. Concentration of Haar measures were obtained by using mixing times of random walks in [38].

Chapter 3

One-Matrix models

In the case where the entries of the matrix $\mathbf{X}^{N,\beta}$ are Gaussian, and more precisely in the case of the so-called Gaussian ensembles, the law of the eigenvalues of $\mathbf{X}^{N,\beta}$ is simple, and given by a Gaussian law with a Coulomb gas interaction. We shall start this part by discussing this point and then will use the explicit formulae for the joint law of the eigenvalues to prove large deviations principles, first for the law of the empirical measure of the eigenvalues and second for law of the largest eigenvalue.

Since the results will now 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

$$X_{kl}^{N,\beta} = \frac{\sum_{i=1}^{\beta} g_{kl}^i e_\beta^i}{\sqrt{\beta N}}, \quad 1\le k<l\le N, \quad X_{kk}^{N,\beta} = \sqrt{\frac{2}{\beta N}} g_{kk} e_\beta^1, \quad 1\le k\le N$$

where $(e_\beta^i)_{1\le i\le\beta}$ is a basis of $\mathbb{R}^\beta$, that is $e_1^1=1$, $e_2^1=1, e_2^2=i$. This definition can be extended to the case $\beta=4$, named the Gaussian symplectic ensemble, when N is even by choosing $\mathbf{X}^{N,\beta}=\left(X_{ij}^{N,\beta}\right)_{1\le i,j\le\frac{N}{2}}$ with $X_{kl}^{N,\beta}$ a 2×2 matrix defined as above but with $(e_\beta^k)_{1\le k\le 4}$ the Pauli matrices

$$e_4^1=\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad e_4^2=\begin{pmatrix}0&-1\\1&0\end{pmatrix},\quad e_4^3=\begin{pmatrix}0&-i\\-i&0\end{pmatrix},\quad e_4^4=\begin{pmatrix}i&0\\0&-i\end{pmatrix}.$$

$(g_{kl}^i, k\le l, 1\le i\le\beta)$ 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**), $(\mathbf{X}^{N,1}, N\in\mathbb{N})$ as the Gaussian Orthogonal Ensemble (**GOE**) and $(\mathbf{X}^{N,4}, N\in\mathbb{N})$ as the Gaussian Symplectic Ensemble (**GSE**) since they can be characterized by the fact that their laws are invariant under the action of the unitary, orthogonal and symplectic group respectively (see [82]). We denote $P_N^{(\beta)}$ the law of $\mathbf{X}^{N,\beta}$.

We have the following key lemma about the joint distribution of the eigenvalues of these ensembles.

Lemma 32. Let $\mathbf{X} \in \mathcal{H}_N^{(\beta)}$ be random with law $P_N^{(\beta)}$. The joint distribution of the eigenvalues $\lambda_1(X) \leq \cdots \leq \lambda_N(X)$, has density proportional to

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

Let us give of the proof in the case $\beta = 1$. It is simply to write the decomposition $X = UDU^*$, with the eigenvalues matrix D which is diagonal and with real entries, and with the eigenvectors matrix U (which is unitary). Suppose this map was a bijection (which it is not, at least at the matrices X that do not possess all distinct eigenvalues) and that one can parametrize the 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 $i \neq j$, the Jacobian must vanish on that set. When $\beta = 1$, this imposes that the polynomial must be proportional to $\prod_{1 \leq i < j \leq N}(x_i - x_j)$. Further degree and symmetry considerations allow to generalize this to $\beta = 2$. We refer the reader to [9] for a full proof, which shows that the set of matrices for which the above manipulations are not permitted has Lebesgue measure zero.

3.1 Large deviations for the law of the spectral measure of Gaussian Wigner's matrices

In this section, we consider the law of N random variables $(\lambda_1, \cdots, \lambda_N)$ with law

$$P_{N,\beta}^{V;[a_-,a_+]}(d\lambda_1, \cdots, d\lambda_N) = (Z_{N,\beta}^{V,[a_-,a_+]})^{-1} \prod \mathbf{1}_{\lambda_i \in [a_-,a_+]} |\Delta(\lambda)|^\beta e^{-N \sum_{i=1}^N V(\lambda_i)} \prod_{i=1}^{N} d\lambda_i, \tag{2}$$

for a continuous function $V : \mathbb{R} \to \mathbb{R}$ such that, if a_- and/or a_+ are infinite,

$$\liminf_{|x| \to \infty} \frac{V(x)}{\beta \log |x|} > 1 \tag{3}$$

and a positive real number β. Here, $\Delta(\lambda) = \prod_{1 \leq i < j \leq N}(\lambda_i - \lambda_j)$ and $-\infty \leq a_- < a_+ \leq +\infty$.

When $V(x) = 4^{-1}\beta x^2$ and $a_- = -a_+ = -\infty$, we have seen in Lemma 32 that $P^N_{4^{-1}\beta x^2,\beta}$ is the law of the eigenvalues of a $N \times N$ GOE matrix when $\beta = 1$, and of a GUE matrix when $\beta = 2$. The case $\beta = 4$ corresponds to another matrix ensemble, namely the GSE. In view of this remark and other applications discussed in Chapter 4, we consider in this section the slightly more general model with a potential V. We emphasize however that the distribution (2) precludes us from considering random matrices with independent non Gaussian entries.

We have proved already at the beginning of these notes that the empirical measure

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

converges almost surely towards the semi-circular law. Moreover, we studied its fluctuations around its mean, both by central limit theorem and by concentration inequalities. Such results did not depend much on the Gaussian nature of the entries.

We address here a different type of question. Namely, we study the probability that L_N takes a very unlikely value. This was already considered in our discussion of concentration inequalities, c.f. Section 2.6.4, where the emphasis was put on obtaining upper bounds on the probability of deviation. In contrast, the purpose of the analysis here is to exhibit a precise estimate on these probabilities, or at least on their logarithmic asymptotics. The appropriate tool for handling such questions is large deviations theory.

Endow $\mathcal{P}(\mathbb{R})$ with the usual weak topology. Our goal is to estimate the probability $P^{V;[a_-,a_+]}_{N,\beta}(L_N \in A)$, for measurable sets $A \subset \mathcal{P}(\mathbb{R})$. Of particular interest is the case where A does not contain the limiting distribution of L_N.

Define the *non-commutative entropy* $\Sigma : \mathcal{P}(\mathbb{R}) \to [-\infty, \infty]$, as

$$\Sigma(\mu) = \int\int \log|x-y| d\mu(x) d\mu(y) . \tag{4}$$

Set next

$$I^V_\beta(\mu) = \begin{cases} \int V(x)d\mu(x) - \frac{\beta}{2}\Sigma(\mu) - c^V_\beta , & \text{if } \int V(x)d\mu(x) < \infty \text{ and } \mu([a_-,a_+]) = 1 \\ \infty, & \text{otherwise} , \end{cases}$$
$$\text{with } c^V_\beta = \inf_{\nu\in\mathcal{P}([a_-,a_+])}\{\textstyle\int V(x)d\nu(x) - \frac{\beta}{2}\Sigma(\nu)\}. \tag{5}$$

Theorem 9. *Let $L_N = N^{-1}\sum_{i=1}^N \delta_{\lambda_i^N}$ where the random variables $\{\lambda_i^N\}_{i=1}^N$ are distributed according to the law $P^{V;[a_-,a_+]}_{N,\beta}$, see (2). Then, the family of random measures L_N satisfies, in $\mathcal{P}(\mathbb{R})$ equipped with the weak topology, a full large deviation principle with good rate function I^V_β in the scale N^2. That is, $I^V_\beta : \mathcal{P}(\mathbb{R}) \to [0,\infty]$ possesses compact level sets $\{\nu : I^V_\beta(\nu) \le M\}$ for all $M \in \mathbb{R}_+$, and*

$$\begin{aligned} &\text{For any open set } O \subset \mathcal{P}(\mathbb{R}) , \\ &\liminf_{N\to\infty} \frac{1}{N^2} \log P^{V,[a_-,a_+]}_{N,\beta} (L_N \in O) \ge -\inf_O I^V_\beta , \end{aligned} \tag{6}$$

and

$$\begin{aligned} &\text{For any closed set } F \subset \mathcal{P}(\mathbb{R}) , \\ &\limsup_{N\to\infty} \frac{1}{N^2} \log P^{V,[a_-,a_+]}_{N,\beta} (L_N \in F) \le -\inf_F I^V_\beta . \end{aligned} \tag{7}$$

The proof of Theorem 9 relies on the properties of the function I^V_β collected in Lemma 33 below. Define the *logarithmic capacity* of a measurable set $A \subset \mathbb{R}$ as

$$\gamma(A) := \exp\left\{ -\inf_{\nu\in M_1(A)} \int\int \log \frac{1}{|x-y|} d\nu(x) d\nu(y) \right\} .$$

Lemma 33.

a. I_β^V is well defined on $\mathcal{P}(\mathbb{R})$ and takes its values in $[0,+\infty]$.

b. $I_\beta^V(\mu)$ is infinite as soon as μ satisfies one of the following conditions

  b.1 $\int V(x)d\mu(x) = +\infty$.

  b.2 There exists a set $A \subset \mathbb{R}$ of positive μ mass but null logarithmic capacity, i.e. a set A such that $\mu(A) > 0$ but $\gamma(A) = 0$.

c. I_β^V is a good rate function.

d. I_β^V is a strictly convex function on $\mathcal{P}(\mathbb{R})$.

e. I_β^V achieves its minimum value at a unique probability measure σ_β^V on $[a_-,a_+]$ characterized by

$$V(x) - \beta\int \log|y-x|d\sigma_\beta^V(y) = \inf_{\nu\in\mathcal{P}(\mathbb{R})}\left(\int V d\nu - \beta\Sigma(\nu)\right),\ \sigma_\beta^V\ a.s., \tag{8}$$

and, for all x except possibly on a set with null logarithmic capacity,

$$V(x) - \beta\int \log|y-x|d\sigma_\beta^V(y) \geq \inf_{\nu\in\mathcal{P}(\mathbb{R})}\left(\int V d\nu - \beta\Sigma(\nu)\right). \tag{9}$$

As an immediate corollary of Theorem 9 and of part e. of Lemma 33 we have the following.

Corollary 34 (Second proof of Wigner's theorem)**.** *Under $P_{N,\beta}^{V;[a_-,a_+]}$, L_N converges almost surely towards σ_β^V.*

Proof of Lemma 33 If $I_\beta^V(\mu) < \infty$, since V is bounded below by assumption (3), $\Sigma(\mu) > -\infty$ and therefore also $\int V d\mu < \infty$. This proves that $I_\beta^V(\mu)$ is well defined (and by definition non negative), yielding point a.

Set

$$f(x,y) = \frac{1}{2}V(x) + \frac{1}{2}V(y) - \frac{\beta}{2}\log|x-y|. \tag{10}$$

Note that $f(x,y)$ goes to $+\infty$ when x,y do by (3). Indeed, $\log|x-y| \leq \log(|x|+1) + \log(|y|+1)$ implies

$$f(x,y) \geq \frac{1}{2}(V(x) - \beta\log(|x|+1)) + \frac{1}{2}(V(y) - \beta\log(|y|+1)) \tag{11}$$

as well as when x,y approach the diagonal $\{x=y\}$; for all $L>0$, there exist constants $K(L)$ (going to infinity with L) such that

$$\begin{aligned}&\{(x,y):\ f(x,y) \geq K(L)\} \subset B_L\,,\\ &B_L := \{(x,y):\ |x-y| < L^{-1}\} \cup \{(x,y):\ |x| > L\} \cup \{(x,y):\ |y| > L\}.\end{aligned} \tag{12}$$

Since f is continuous on the compact set B_L^c, we conclude that f is bounded below, and denote $b_f > -\infty$ a lower bound. Therefore, since for any measurable subset A of $\mathbb{R}$,

$$\begin{aligned}
I_\beta^V(\mu) &= \int\int (f(x,y) - b_f) d\mu(x) d\mu(y) + b_f - c_\beta^V \\
&\geq \int_A \int_A (f(x,y) - b_f) d\mu(x) d\mu(y) + b_f - c_\beta^V \\
&\geq \frac{\beta}{2} \int_A \int_A \log|x-y|^{-1} d\mu(x) d\mu(y) + \inf_{x\in\mathbb{R}} V(x) \mu(A)^2 - |b_f| - c_\beta^V \\
&\geq -\frac{\beta}{2} \mu(A)^2 \log(\gamma(A)) - |b_f| - c_\beta^V + \inf_{x\in\mathbb{R}} V(x) \mu(A)^2
\end{aligned}$$

one concludes that if $I_\beta^V(\mu) < \infty$, and A is a measurable set with $\mu(A) > 0$, then $\gamma(A) > 0$. This completes the proof of point b.

We now show that I_V^β is a good rate function, and first that its level sets $\{I_V^\beta \leq M\}$ are closed, that is that I_V^β is lower semi-continuous. Indeed, by the monotone convergence theorem,

$$\begin{aligned}
I_\beta^V(\mu) &= \int\int f(x,y) d\mu(x) d\mu(y) - c_\beta^V \\
&= \sup_{M\geq 0} \int\int (f(x,y) \wedge M) d\mu(x) d\mu(y) - c_\beta^V
\end{aligned}$$

But $f^M = f \wedge M$ is bounded continuous and so for $M < \infty$,

$$I_\beta^{V,M}(\mu) = \int\int (f(x,y) \wedge M) d\mu(x) d\mu(y)$$

is bounded continuous on $\mathcal{P}(\mathbb{R})$. As a supremum of the continuous functions $I_\beta^{V,M}$, I_β^V is lower semi-continuous. Hence, by Theorem 29, to prove that $\{I_\beta^V \leq L\}$ is compact, it is enough to show that $\{I_\beta^V \leq L\}$ is included in a compact subset of $\mathcal{P}(\mathbb{R})$ of the form

$$K_\varepsilon = \cap_{B\in\mathbb{N}} \{\mu \in \mathcal{P}(\mathbb{R}) : \mu([-B,B]^c) \leq \varepsilon(B)\}$$

with a sequence $\varepsilon(B)$ going to zero as B goes to infinity.

Arguing as in (12), there exist constants $K'(L)$ going to infinity as L goes to infinity, such that

$$\{(x,y) : \ |x| > L, |y| > L\} \subset \{(x,y) : \ f(x,y) \geq K'(L)\}. \tag{13}$$

Hence, for any $L > 0$ large,

$$\begin{aligned}
\mu(|x| > L)^2 &= \mu \otimes \mu(|x| > L, |y| > L) \\
&\leq \mu \otimes \mu\big(f(x,y) \geq K'(L)\big) \\
&\leq \frac{1}{K'(L) - b_f} \int\int (f(x,y) - b_f) d\mu(x) d\mu(y) \\
&= \frac{1}{K'(L) - b_f} (I_\beta^V(\mu) + c_\beta^V - b_f)
\end{aligned}$$

Hence, with $\varepsilon(B) = [\sqrt{(M + c^V_\beta - b_f)_+}/\sqrt{(K'(B) - b_f)_+}] \wedge 1$ going to zero when B goes to infinity, one has that $\{I^V_\beta \le M\} \subset K_\varepsilon$. This completes the proof of point c.

Since I^V_β is a good rate function, it achieves its minimal value. Let σ^V_β be a minimizer. Then, for any signed measure $\bar{\nu}(dx) = \varphi(x)\sigma^V_\beta(dx) + \psi(x)dx$ with two bounded measurable compactly supported functions (φ, ψ) such that $\psi \ge 0$ and $\bar{\nu}(\mathbb{R}) = 0$, for $\varepsilon > 0$ small enough, $\sigma^V_\beta + \varepsilon\bar{\nu}$ is a probability measure so that

$$I^V_\beta(\sigma^V_\beta + \varepsilon\bar{\nu}) \ge I^V_\beta(\sigma^V_\beta)$$

which implies

$$\int \left(V(x) - \beta \int \log|x-y| d\sigma^V_\beta(y) \right) d\bar{\nu}(x) \ge 0.$$

Taking $\psi = 0$, we deduce by symmetry that there is a constant C^V_β such that

$$V(x) - \beta \int \log|x-y| d\sigma^V_\beta(y) = C^V_\beta\,, \quad \sigma^V_\beta \text{ a.s.,} \tag{14}$$

which implies that σ^V_β is compactly supported (as $V(x) - \beta \int \log|x-y| d\sigma^V_\beta(y)$ goes to infinity when x does). Taking $\varphi(x) = -\int \psi(y)dy$, we then find that

$$V(x) - \beta \int \log|x-y| d\sigma^V_\beta(y) \ge C^V_\beta \tag{15}$$

Lebesgue almost surely, and then everywhere outside of the support of σ^V_β by continuity. By (14) and (15) we deduce that

$$C^V_\beta = \inf_{\nu \in \mathcal{P}(\mathbb{R})} \{ \int (V(x) - \beta \int \log|x-y| d\sigma^V_\beta(y)) d\nu(x) \}\,.$$

This completes the proof of (8) and (9). The claimed uniqueness of σ^V_β, and hence the completion of the proof of part e., will then follow from the strict convexity claim (point d. of the lemma), which we turn to next.

Note first that we can rewrite I^V_β as

$$I^V_\beta(\mu) = -\frac{\beta}{2}\Sigma(\mu - \sigma^V_\beta) + \int \left(V - \beta \int \log|x-y| d\sigma^V_\beta(y) - C^V_\beta \right) d\mu(x).$$

The fact that I^V_β is strictly convex comes from the observation that Σ is strictly concave, as can be checked from the formula

$$\log|x-y| = \int_0^\infty \frac{1}{2t} \left(\exp\{-\frac{1}{2t}\} - \exp\{-\frac{|x-y|^2}{2t}\} \right) dt \tag{16}$$

which entails that for any $\mu \in \mathcal{P}(\mathbb{R})$,

$$\Sigma(\mu - \sigma^V_\beta) = -\int_0^\infty \frac{1}{2t} \left(\int \int \exp\{-\frac{|x-y|^2}{2t}\} d(\mu - \sigma^V_\beta)(x) d(\mu - \sigma^V_\beta)(y) \right) dt.$$

Indeed, one may apply Fubini's theorem when μ_1, μ_2 are supported in $[-\frac{1}{2}, \frac{1}{2}]$ since then $\mu_1 \otimes \mu_2(\exp\{-\frac{1}{2t}\} - \exp\{-\frac{|x-y|^2}{2t}\} \leq 0) = 1$. One then deduces the claim for any compactly supported probability measures by scaling and finally for all probability measures by approximations. The fact that for all $t \geq 0$,

$$\int\int \exp\{-\frac{|x-y|^2}{2t}\} d(\mu - \sigma_\beta^V)(x) d(\mu - \sigma_\beta^V)(y)$$

$$= \sqrt{\frac{t}{2\pi}} \int_{-\infty}^{+\infty} \left| \int \exp\{i\lambda x\} d(\mu - \sigma_\beta^V)(x) \right|^2 \exp\{-\frac{t\lambda^2}{2}\} d\lambda$$

therefore entails that Σ is concave since $\mu \to \left| \int \exp\{i\lambda x\} d(\mu - \sigma_\beta^V)(x) \right|^2$ is convex for all $\lambda \in \mathbb{R}$. Strict convexity comes from the fact by the Cauchy-Schwarz inequality, $\Sigma(\alpha\mu + (1-\alpha)\nu) = \alpha\Sigma(\mu) + (1-\alpha)\Sigma(\nu)$ if and only if $\Sigma(\nu - \mu) = 0$ which implies that all the Fourier transforms of $\nu - \mu$ are null, and hence $\mu = \nu$. This completes the proof of part d and hence of the lemma. $\square$

Proof of Theorem 9: To begin, let us remark that with f as in (10),

$$P_{N,\beta}^{V;[a_-,a_+]}(d\lambda_1, \cdots, d\lambda_N) = (Z_{N,\beta}^{V,[a_-,a_+]})^{-1} \prod_{i=1}^N 1_{\lambda_i \in [a_-,a_+]} e^{-N^2 \int_{x \neq y} f(x,y) dL_N(x) dL_N(y)} \prod_{i=1}^N e^{-V(\lambda_i)} d\lambda_i.$$

Hence, if $\mu \to \int_{x \neq y} f(x,y) d\mu(x) d\mu(y)$ was a bounded continuous function, the proof would follow from a standard Laplace method. The main point will be therefore to overcome the singularity of this function, with the most delicate part being overcoming the singularity of the logarithm.

Following Appendix 7.1 (see Corollary 112), a full large deviation principle can be proved by showing that exponential tightness holds, as well as estimating the probability of small balls. We follow these steps below.

- *Exponential tightness* Of course this is clear if a_- and a_+ are finite. Observe that by Jensen's inequality,

$$\begin{aligned} \log Z_{N,\beta}^{V,[a_-,a_+]} &\geq N \log \int_{a_-}^{a_+} e^{-V(x)} dx \\ &\quad - N^2 \int \left(\int_{x \neq y} f(x,y) dL_N(x) dL_N(y) \right) \prod_{i=1}^N \frac{1_{\lambda_i \in [a_-,a_+]} e^{-V(\lambda_i)} d\lambda_i}{\int_{a_-}^{a_+} e^{-V(x)} dx} \geq -CN^2 \end{aligned}$$

with some finite constant C. Moreover, by (11) and (3), there exist constants $a > 0$ and $c > -\infty$ so that

$$f(x,y) \geq a|V(x)| + a|V(y)| + c$$

from which one concludes that for all $M \geq 0$,

$$P_{N,\beta}^{V;[a_-,a_+]}\left(\int |V(x)| dL_N \geq M \right) \leq e^{-2aN^2M + (C-c)N^2} \left(\int_{a_-}^{a_+} e^{-V(x)} dx \right)^N.$$

Since V goes to infinity at infinity, $K_M = \{\mu \in \mathcal{P}(\mathbb{R}) : \int |V| d\mu \leq M\}$ is a compact set for all $M < \infty$, so that we have proved that the law of L_N under $P_{N,\beta}^{V;[a_-,a_+]}$ is exponentially tight.

-*Large deviation upper bound* d will denote the Dudley metric. Note first that $L_N \in \mathcal{P}([a_-,a_+])$ so that the rate function has to be infinite on $\mathcal{P}(\mathbb{R}) \backslash \mathcal{P}([a_-,a_+])$. We next consider $\mu \in \mathcal{P}([a_-,a_+])$, and prove that if we set $\overline{P}_{N,\beta}^{V;[a_-,a_+]} = Z_{N,\beta}^{V,[a_-,a_+]} P_{N,\beta}^{V;[a_-,a_+]}$

$$\lim_{\varepsilon \to 0} \limsup_{N \to \infty} \frac{1}{N^2} \log \overline{P}_{N,\beta}^{V;[a_-,a_+]} (d(L_N,\mu) \leq \varepsilon) \leq -\int f(x,y) d\mu(x) d\mu(y). \tag{17}$$

For any $M \geq 0$, the following bound holds

$$\overline{P}_{N,\beta}^{V;[a_-,a_+]} (d(L_N,\mu) \leq \varepsilon)$$

$$\leq \int_{d(L_N,\mu) \leq \varepsilon} e^{-N^2 \int_{x \neq y} f(x,y) \wedge M dL_N(x) dL_N(y)} \prod_{i=1}^{N} e^{-V(\lambda_i)} d\lambda_i.$$

Since under the product Lebesgue measure, the λ_i's are almost surely distinct, it holds that $L_N \otimes L_N(x=y) = N^{-1}$, $\overline{P}_{N,\beta}^{V;[a_-,a_+]}$ almost surely. Thus, we deduce for all $M \geq 0$, with $f_M(x,y) = f(x,y) \wedge M$,

$$\int f_M(x,y) dL_N(x) dL_N(y) = \int_{x \neq y} f_M(x,y) dL_N(x) dL_N(y) + MN^{-1},$$

and so

$$\overline{P}_{N,\beta}^{V;[a_-,a_+]} (d(L_N,\mu) \leq \varepsilon)$$

$$\leq e^{MN} \int_{d(L_N,\mu) \leq \varepsilon} e^{-N^2 \int f_M(x,y) dL_N(x) dL_N(y)} \prod_{i=1}^{N} e^{-V(\lambda_i)} d\lambda_i.$$

Since $I_\beta^{V,M}(\nu) = \int f_M(x,y) d\nu(x) d\nu(y)$ is bounded continuous, we deduce that

$$\lim_{\varepsilon \to 0} \limsup_{N \to \infty} \frac{1}{N^2} \log \overline{P}_{N,\beta}^{V;[a_-,a_+]} (d(L_N,\mu) \leq \varepsilon) \leq -I_\beta^{V,M}(\mu).$$

We finally let M go to infinity and conclude by the monotone convergence theorem. Note that the same argument shows that

$$\limsup_{N \to \infty} \frac{1}{N^2} \log Z_{N,\beta}^{V,[a_-,a_+]} \leq -\inf_{\mu \in \mathcal{P}(\mathbb{R})} \int f(x,y) d\mu(x) d\mu(y). \tag{18}$$

- *Large deviation lower bound.* We prove here that for any $\mu \in \mathcal{P}(\mathbb{R})$

$$\lim_{\varepsilon \to 0} \liminf_{N \to \infty} \frac{1}{N^2} \log \overline{P}_{N,\beta}^{V;[a_-,a_+]} (d(L_N,\mu) \leq \varepsilon) \geq -\int f(x,y) d\mu(x) d\mu(y). \tag{19}$$

Note that we can assume without loss of generality that $I_\beta^V(\mu) < \infty$, since otherwise the bound in trivial, and so in particular, we may and will assume that μ has no atoms. We can also assume that μ is compactly supported since if we consider $\mu_M = \mu([-M,M])^{-1}1_{|x|\le M}d\mu(x)$, clearly μ_M converges towards μ and by the monotone convergence theorem, one checks that, since f is bounded below,

$$\lim_{M\uparrow\infty}\int f(x,y)d\mu_M(x)d\mu_M(y) = \int f(x,y)d\mu(x)d\mu(y)$$

which insures that it is enough to prove the lower bound for $(\mu_M, M\in\mathbb{R}, I_\beta^V(\mu)<\infty)$, and so for compactly supported probability measures with finite entropy.

The idea is to localize the eigenvalues $(\lambda_i)_{1\le i\le N}$ in small sets and to take advantage of the fast speed N^2 of the large deviations to neglect the small volume of these sets. To do so, we first remark that for any $\nu \in \mathcal{P}([a_-,a_+])$ with no atoms if we set

$$\begin{aligned} x^{1,N} &= \inf\left\{x |\ \nu(]-\infty,x]) \ge \frac{1}{N+1}\right\} \\ x^{i+1,N} &= \inf\left\{x \ge x^{i,N} |\ \nu\left(]x^{i,N},x]\right) \ge \frac{1}{N+1}\right\} \qquad 1\le i\le N-1, \end{aligned}$$

for any real number η, there exists an integer number $N(\eta)$ such that, for any N larger than $N(\eta)$,

$$d\left(\nu, \frac{1}{N}\sum_{i=1}^{N}\delta_{x^{i,N}}\right) < \eta.$$

In particular, for $N \ge N(\frac{\delta}{2})$,

$$\left\{(\lambda_i)_{1\le i\le N} \mid |\lambda_i - x^{i,N}| < \frac{\delta}{2}\ \forall i\in[1,N]\right\} \subset \{(\lambda_i)_{1\le i\le N} \mid d(L_N,\nu)<\delta\}$$

Moreover if we take

$$\Omega = \left\{(\lambda_i)_{1\le i\le N} \mid x^{i,N} \le \lambda_i < x_i + \frac{\delta}{2}\ \forall i\in[1,[N/2]], x^{i,N} - \frac{\delta}{2} \le \lambda_i \le x^{i,N} \forall i \in [[N/2]+1,N]\right\}$$

then since $x_{1,N} \ge a_-$ and $x_{N,N} \le a_+$ we deduce that as

$$\Omega \subset \{(\lambda_i)_{1\le i\le N} \in [a_-,a_+] \mid d(L_N,\nu)<\delta\}$$

so that we have the lower bound

$$
\begin{aligned}
&\overline{P}_{N,\beta}^{V;[a_-,a_+]}(d(L_N,\mu)\le\varepsilon)\\
\ge\;& \int_\Omega e^{-N^2\int_{x\neq y}f(x,y)dL_N(x)dL_N(y)}\prod_{i=1}^N e^{-V(\lambda_i)}d\lambda_i\\
=\;& \int\prod_{i\le[N/2]}1_{\{0\le\lambda_i\le\frac{\delta}{2}\}}\prod_{i\ge[N/2]+1}1_{\{-\frac{\delta}{2}\le\lambda_i\le 0\}}\prod_{i<j}|x^{i,N}-x^{j,N}+\lambda_i-\lambda_j|^\beta e^{-N\sum_{i=1}^N V(x^{i,N}+\lambda_i)}\prod_{i=1}^N d\lambda_i\\
\ge\;& \left(\prod_{i+1<j}|x^{i,N}-x^{j,N}|^\beta\prod_i|x^{i,N}-x^{i+1,N}|^{\frac{\beta}{2}}e^{-N\sum_{i=1}^N V(x^{i,N})}\right)\\
&\times\left(\int\prod_{i\le[N/2]}1_{\{0\le\lambda_i\le\frac{\delta}{2}\}}\prod_{i\ge[N/2]+1}1_{\{-\frac{\delta}{2}\le\lambda_i\le 0\}}1_{\lambda_i\ge\lambda_{i+1}}\prod_i|\lambda_i-\lambda_{i+1}|^{\frac{\beta}{2}}e^{-N\sum_{i=1}^N[V(x^{i,N}+\lambda_i)-V(x^{i,N})]}\prod_{i=1}^N d\lambda_i\right)\\
=:\;& P_{N,1}\times P_{N,2}
\end{aligned}
\tag{20}
$$

where we used that $|x^{i,N}-x^{j,N}+\lambda_i-\lambda_j|\ge|x^{i,N}-x^{j,N}|\vee|\lambda_i-\lambda_j|$ when $\lambda_i\ge\lambda_j$ and $x^{i,N}\ge x^{j,N}$. To estimate $P_{N,2}$, note that since we assumed that μ is compactly supported, the $(x^{i,N},1\le i\le N)_{N\in\mathbb{N}}$ are uniformly bounded and so by continuity of V

$$
\lim_{N\to\infty}\sup_{N\in\mathbb{N}}\sup_{1\le i\le N}\sup_{|x|\le\delta}|V(x^{i,N}+x)-V(x^{i,N})|=0.
$$

Moreover, writing $u_{i+1}=\lambda_i-\lambda_{i+1}$, $u_1=\lambda_1$

$$
\begin{aligned}
&\int\prod_{i\le[N/2]}1_{\{0\le\lambda_i\le\frac{\delta}{2}\}}\prod_{i\ge[N/2]+1}1_{\{-\frac{\delta}{2}\le\lambda_i\le 0\}}1_{\lambda_i\ge\lambda_{i+1}}\prod_i|\lambda_i-\lambda_{i+1}|^{\frac{\beta}{2}}\prod_{i=1}^N d\lambda_i\\
\ge\;& \int\prod_{i=2}^N 1_{0<u_i<\frac{\delta}{N}}1_{u_1\in[\frac{\delta}{2}-\frac{\delta}{N},\frac{\delta}{2}]}\prod_{i=2}^N u_i^{\frac{\beta}{2}}\prod_{i=1}^N du_i\\
\ge\;& \left(\frac{\delta}{(\beta+2)N}\right)^{N(\frac{\beta}{2}+1)}.
\end{aligned}
$$

Therefore,

$$
\lim_{\delta\to 0}\liminf_{N\to\infty}\frac{1}{N^2}\log P_{N,2}\ge 0. \tag{21}
$$

To handle the term $P_{N,1}$, the uniform boundness of the $x^{i,N}$'s and the convergence of their empirical measure towards μ imply that

$$
\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N V(x^{i,N})=\int V(x)d\mu(x). \tag{22}
$$

Finally since $x \to \log(x)$ increases on $\mathbb{R}^+$, we notice that

$$\begin{aligned}
&\int_{x^{1,N} \le x < y \le x^{N,N}} \log(y-x) d\mu(x) d\mu(y) \\
\le \quad & \sum_{1 \le i \le j \le N-1} \log(x^{j+1,N} - x^{i,N}) \int_{\substack{x \in [x^{i,N}, x^{i+1,N}] \\ y \in [x^{j,N}, x^{j+1,N}]}} 1_{x<y} d\mu(x) d\mu(y) \\
= \quad & \frac{1}{(N+1)^2} \sum_{i<j} \log|x^{i,N} - x^{j+1,N}| + \frac{1}{2(N+1)^2} \sum_{i=1}^{N-1} \log|x^{i+1,N} - x^{i,N}|.
\end{aligned}$$

Since $\log|x-y|$ is bounded when x, y are in the support of the compactly supported measure μ, the monotone convergence theorem implies that the left side in the last display converges towards $\int\int \log|x-y| d\mu(x) d\mu(x)$. Thus, with (22), we have proved

$$\liminf_{N \to \infty} \frac{1}{N^2} \log P_{N,1} \ge \int_{x<y} \log(y-x) d\mu(x) d\mu(y) - \int V(x) d\mu(x)$$

which concludes, with (20) and (21), the proof of (19).

-*Conclusion* By (19), for all $\mu \in \mathcal{P}([a_-, a_+])$,

$$\begin{aligned}
\liminf_{N \to \infty} \frac{1}{N^2} \log Z^N_{\beta,V} &\ge \lim_{\varepsilon \to 0} \liminf_{N \to \infty} \frac{1}{N^2} \log \overline{P}^{V;[a_-,a_+]}_{N,\beta} (d(L_N, \mu) \le \varepsilon) \\
&\ge -\int f(x,y) d\mu(x) d\mu(y)
\end{aligned}$$

and so optimizing with respect to $\mu \in \mathcal{P}(\mathbb{R})$ and with (18),

$$\lim_{N \to \infty} \frac{1}{N^2} \log Z^N_{\beta,V} = -\inf_{\mu \in \mathcal{P}(\mathbb{R})} \{ \int f(x,y) d\mu(x) d\mu(y) \} = -c^V_\beta.$$

Thus, (19) and (17) imply the weak large deviation principle, i.e. that for all $\mu \in \mathcal{P}(\mathbb{R})$,

$$\begin{aligned}
&\lim_{\varepsilon \to 0} \liminf_{N \to \infty} \frac{1}{N^2} \log P^{V;[a_-,a_+]}_{N,\beta} (d(L_N, \mu) \le \varepsilon) \\
&= \lim_{\varepsilon \to 0} \limsup_{N \to \infty} \frac{1}{N^2} \log P^{V;[a_-,a_+]}_{N,\beta} (d(L_N, \mu) \le \varepsilon) = -I^V_\beta(\mu).
\end{aligned}$$

This, together with the exponential tightness property proved above completes the proof of the full large deviation principle stated in Theorem 9. □

Bibliographical Notes The proof of Theorem 9 is a slight generalization of the techniques introduced in [16] to more general potentials.

3.2 Large deviations of the maximum eigenvalue

We let

$$\mathcal{J}^{V;[a_-,a_+]}(x) = \frac{V(x)}{2} - \int_{a_-}^{a_+} \mathrm{d}\mu^{V;[a_-,a_+]}_{\mathrm{eq}}(\xi) \ln|x-\xi| \tag{23}$$

when $x \in [a_-, a_+]$, and $+\infty$ otherwise. Suppose that $[a_-, a_+] \neq [\alpha_-, \alpha_+]$, and set:

$$\widetilde{\mathcal{J}}^{V;[a_-,a_+]}(x) = \mathcal{J}^{V;[a_-,a_+]}(x) - \inf_{\xi \in [a_-,a_+]} \mathcal{J}^{V;[a_-,a_+]}(\xi) \tag{24}$$

In this section we assume for simplicity that the (closed) support of the minimizing measure σ_β^V is connected; we denote it by $[\alpha_-, \alpha_+]$. We define also $\widetilde{\mathcal{J}}_{\max}^{V;[a_-,a_+]}(x)$ (resp. $\widetilde{\mathcal{J}}_{\min}^{V;[a_-,a_+]}(x)$) which is equal to $\widetilde{\mathcal{J}}^{V;[a_-,a_+]}(x)$, except when $x \in]-\infty, \alpha_-]$ (resp. $[\alpha_+, +\infty[$) where we set its value to $+\infty$.

Proposition 35. *Let $V : [a_-, a_+] \to \mathbb{R}$ be a continuous function, and if $a_\tau = \tau\infty$, assume that:*

$$\liminf_{x \to \tau\infty} \frac{V(x)}{2 \ln |x|} > 1 \tag{25}$$

Assume that $\widetilde{\mathcal{J}}^{V;[a_-,a_+]}$ does not vanish outside $[\alpha_-, \alpha_+]$. Then:

(*i*) *$\beta\widetilde{\mathcal{J}}_{\max}^{V;[a_-,a_+]}$ (resp. $\beta\widetilde{\mathcal{J}}_{\min}^{V;[a_-,a_+]}$) is a good rate function on $[a_-, a_+]$, which vanishes at α_+ (resp. α_-).*

(*ii*) *The law of $\lambda_{\max}$ (resp. $\lambda_{\min}$) under $P_{N,\beta}^{V;[a_-,a_+]}$ satisfies a large deviation principle with speed N and rate function equal to $\beta\widetilde{\mathcal{J}}_{\max}^{V;[a_-,a_+]}$ (resp. $\beta\widetilde{\mathcal{J}}_{\min}^{V;[a_-,a_+]}$) on $[a_-, a_+]$. In other words, for any closed subset F, or open subset Ω, of $[a_-, a_+]$:*

$$\begin{aligned} \limsup_{N \to \infty} \frac{1}{N} \ln P_{N,\beta}^{V;[a_-,a_+]}(\lambda_{\max} \in F) &\leq -\beta \inf_{x \in F} \widetilde{\mathcal{J}}_{\max}^{V;[a_-,a_+]}(x) \\ \liminf_{N \to \infty} \frac{1}{N} \ln P_{N,\beta}^{V;[a_-,a_+]}(\lambda_{\max} \in \Omega) &\geq -\beta \inf_{x \in \Omega} \widetilde{\mathcal{J}}_{\max}^{V;[a_-,a_+]}(x) \end{aligned}$$

and similar statements hold for $\lambda_{\min}$. In particular, for any $\varepsilon > 0$,

$$\begin{aligned} \limsup_{N \to \infty} \frac{1}{N} \ln P_{N,\beta}^{V;[a_-,a_+]}(\lambda_{\min} \leq \alpha_- - \varepsilon) &< 0 \\ \limsup_{N \to \infty} \frac{1}{N} \ln P_{N,\beta}^{V;[a_-,a_+]}(\lambda_{\max} \geq \alpha_+ + \varepsilon) &< 0 \end{aligned}$$

3.2.1 $\widetilde{\mathcal{J}}^{V;[a_-,a_+]}$ is a good rate function

$\widetilde{\mathcal{J}}^{V;[a_-,a_+]}$ is lower semicontinuous as a supremum of the continuous functions

$$\widetilde{\mathcal{J}}_\varepsilon^{V;[a_-,a_+]}(x) := \frac{V(x)}{2} - \int_{a_-}^{a_+} d\mu_{\mathrm{eq}}^{V;[a_-,a_+]}(\xi) \ln\left[\max(|x-\xi|, \varepsilon)\right] - \inf_{\xi \in [a_-,a_+]} \mathcal{J}^{V;[a_-,a_+]}(\xi)$$

Moreover, by the assumption of Eqn. 25, it goes to infinity at infinity. Hence, $\widetilde{\mathcal{J}}^{V;[a_-,a_+]}$ has compact level sets. Since it is non-negative, it is a good rate function.

3.2.2 The law of the extreme eigenvalues is exponentially tight

Exponential tightness of the extreme eigenvalues means:

$$\limsup_{M\to\infty}\limsup_{N\to\infty}\frac{1}{N}\ln P_{N,\beta}^{V;[a_-,a_+]}(\lambda_{\max}\geq M \text{ or } \lambda_{\min}\leq -M) = -\infty \tag{26}$$

By [9, Lemma 2.6.7], it is enough to show that:

$$\limsup_{N\to\infty}\frac{1}{N}\ln\frac{Z_{N-1,\beta}^{V;[a_-,a_+]}}{Z_{N,\beta}^{V;[a_-,a_+]}} < \infty \tag{27}$$

For this purpose, observe that by Jensen's inequality

$$\begin{aligned}\frac{Z_{N,\beta}^{V;[a_-,a_+]}}{Z_{N-1,\beta}^{V;[a_-,a_+]}} &= P_{N-1,\beta}^{V;[a_-,a_+]}\left[\int_{a_-}^{a_+}\mathrm{d}\lambda_N\exp\Big(\beta\sum_{i=1}^{N-1}\ln|\lambda_N-\lambda_i| - \frac{\beta N}{2}V(\lambda_N) - \frac{\beta}{2}\sum_{i=1}^{N-1}V(\lambda_i)\Big)\right] \\ &\geq \kappa\exp\left\{\frac{\beta}{2}(P_{N-1,\beta}^{V;[a_-,a_+]}\otimes\chi)\left[2\sum_{i=1}^{N-1}\ln|\lambda_N-\lambda_i| - (N-1)V(\lambda_N) - \sum_{i=1}^{N-1}V(\lambda_i)\right]\right\}\end{aligned}$$

where we denoted χ the law on λ_N given by:

$$\mathrm{d}\chi(x) = \frac{\mathbf{1}_{[a_-,a_+]}(x)\mathrm{d}x}{\kappa}\,e^{-\frac{\beta}{2}V(x)} \qquad \kappa = \int_{a_-}^{a_+}\mathrm{d}\xi\, e^{-\frac{\beta}{2}V(\xi)} \tag{28}$$

The function $\xi\mapsto\int_{\mathbb{R}}\mathrm{d}\chi(\lambda_N)\ln|\lambda_N-\xi|$ is bounded on compact sets and going to infinity like $\ln|\xi|$, so is bounded from below, by a constant $\frac{\kappa_1}{2}$. We can rewrite:

$$\frac{Z_{N,\beta}^{V;[a_-,a_+]}}{Z_{N-1,\beta}^{V;[a_-,a_+]}} \geq \kappa\exp\left\{\beta(N-1)\left[\kappa_1 - \chi[V] - P_{N-1,\beta}^{V;[a_-,a_+]}[L_{N-1}(V)]\right]\right\} \tag{29}$$

By exponential tightness [9, Eqn. 2.6.21], we know that there exists a constant $\kappa_2 > 0$ so that

$$-P_{N-1,\beta}^{V;[a_-,a_+]}\big[L_{N-1}(V)\big] \quad\geq\quad -P_{N-1,\beta}^{V;[a_-,a_+]}\big[L_{N-1}(|V|)\big] \geq -\kappa_2$$

So, if we set $\kappa_3 = \chi[V]$ and choose κ_2 large enough, we have:

$$\frac{Z_{N,\beta}^{V;[a_-,a_+]}}{Z_{N-1,\beta}^{V;[a_-,a_+]}} \geq \kappa\, e^{-\beta(N-1)\delta} \tag{30}$$

with a positive constant $\delta = -\kappa_1 + \kappa_2 + \kappa_3$. This justifies Eqn. 9 and completes the proof of Eqn. 26.

3.2.3 Upper bound for large deviation of the extreme eigenvalues

We give the argument for the minimal eigenvalue, the case of the maximal eigenvalue being similar. By exponential tightness (Eqn. 26), it is enough to prove a weak large deviation upper bound, that is control the probability of small balls. First, observe that for any $x - \alpha_- \geq 2\varepsilon > 0$,

$$P_{N,\beta}^{V;[a_-,a_+]}[\lambda_{\min} \geq x] \leq P_{N,\beta}^{V;[a_-,a_+]}[L_N(\mathbf{1}_{[\alpha_-,\alpha_-+\varepsilon]}) = 0]$$

is of order $e^{-N^2\kappa_\varepsilon}$ for some $\kappa_\varepsilon > 0$ by the large deviation principle for the law of L_N under $P_{N,\beta}^{V;[a_-,a_+]}$, see e.g. [16] or [9, Theorem 2.6.1]. Moreover, the probability that $\lambda_{\min}$ is smaller than a_- vanishes and therefore we have

$$\limsup_{\varepsilon\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\ln P_{N,\beta}^{V;[a_-,a_+]}(\lambda_{\min} \in]-\infty, a_- - \varepsilon] \cup [\alpha_- + \varepsilon, +\infty[) = -\infty$$

Hence, we may and shall concentrate on probability of deviating on $[a_-, \alpha_-]$, and actually we may restrict ourselves to the case where a_- and a_+ are finite by Eqn. 26. We let F be a closed subset of $[a_-, \alpha_-]$. We then have:

$$P_{N,\beta}^{V;[a_-,a_+]}[\lambda_{\min} \in F] = Y_N \int_F \mathrm{d}\xi\, e^{-\frac{\beta}{2}V(\xi)}\, \Xi_N(\xi) \tag{31}$$

where we introduced:

$$\begin{aligned} Y_N &= \frac{Z_{N-1,\beta}^{\frac{N}{N-1}V;[a_-,a_+]}}{Z_{N,\beta}^{V;[a_-,a_+]}} \\ \Xi_N(\xi) &= P_{N-1,\beta}^{\frac{N}{N-1}V;[a_-,a_+]}\left(e^{\beta\sum_{i=1}^{N-1}\ln|\xi-\lambda_i| - \frac{\beta}{2}(N-1)V(\xi)}\prod_{i=1}^{N-1}\mathbf{1}_{[a_-,\lambda_i]}(\xi)\right) \end{aligned}$$

Upper bound for $\Xi_N(\xi)$

Notice that the logarithm is uniformly bounded from above on compacts so that the exponent is at most of order N. Therefore, we may and shall assume that under $\mu_{N-1,\beta}^{\frac{N}{N-1}V;[a_-,a_+]}$, L_{N-1} is at a distance smaller than $\kappa > 0$ from the equilibrium measure $\mu_{\mathrm{eq}} := \mu_{\mathrm{eq}}^{V;[a_-,a_+]}$, since the opposite event has probability smaller than $e^{-\Gamma_\kappa(N-1)^2}$ for some $\Gamma_\kappa > 0$, see e.g. [9, Theorem 2.6.1]. Thus, we have for large N:

$$\Xi_N(\xi) \leq e^{-\Gamma_\kappa N^2/2} + e^{\beta(N-1)\sup_{d(\mu,\mu_{\mathrm{eq}})<\kappa}\left(-\frac{V(\xi)}{2} + \int \ln|\xi-\eta|\mathrm{d}\mu_{\mathrm{eq}}(\eta)\right)} \tag{32}$$

where we take the supremum over probability measures on $[a_-, a_+]$. We observe also that for all probability measures μ on $[a_-, a_+]$, and for any $\zeta > 0$:

$$\int_{a_-}^{a_+} \ln|\xi - \eta|\mathrm{d}\mu(\eta) \leq \varphi_\zeta(\mu, \xi) = \int_{a_-}^{a_+} \ln\left[\max(|\xi - \eta|, \zeta)\right]\mathrm{d}\mu(\eta) \tag{33}$$

where $\varphi_\zeta(\mu,\xi)$ is continuous in μ and ξ, and $\varphi_\zeta(\mu_{\mathrm{eq}},\xi)$ converges towards $\varphi_0(\mu_{\mathrm{eq}},\xi)$ as ζ goes to zero. We deduce that:

$$\limsup_{\kappa\downarrow 0}\sup_{\xi\in F}\sup_{d(\mu,\mu_{\mathrm{eq}})<\kappa}\beta\left(\int \ln|\xi-\eta|d\mu(\eta)-\frac{V(\xi)}{2}\right)\leq -\beta\inf_{\xi\in F}\mathcal{J}^{V;[a_-,a_+]}(\xi) \tag{34}$$

Therefore, for any $\eta'>0$, and N large enough, we conclude that:

$$\sup_{\xi\in F}\Xi_N(\xi)\leq e^{N\left(\eta'-\beta\inf_{\xi\in F}\mathcal{J}^{V;[a_-,a_+]}(\xi)\right)} \tag{35}$$

Lower bound for Y_N

We observe that, for any $\varepsilon>0$ small enough, and any $x\in[a_-+\varepsilon,a_+-\varepsilon]$, there exists δ_ε going to zero with ε so that

$$\begin{aligned}
\frac{1}{Y_N} &= \frac{Z_{N,\beta}^{V;[a_-,a_+]}}{Z_{N-1,\beta}^{\frac{N}{N-1}V;[a_-,a_+]}}\\
&= \mu_{N-1,\beta}^{\frac{N}{N-1}V;[a_-,a_+]}\left(\int_{a_-}^{a_+}\mathrm{d}\xi\, e^{-\frac{\beta N}{2}V(\xi)}\prod_{i=1}^{N-1}|\xi-\lambda_i|^\beta\right)\\
&\geq \mu_{N-1,\beta}^{\frac{N}{N-1}V;[a_-,a_+]}\left(\int_{x-\varepsilon}^{x+\varepsilon}\mathrm{d}\xi\, e^{-\frac{\beta N}{2}V(\xi)}\prod_{i=1}^{N-1}|\xi-\lambda_i|^\beta\right)\\
&\geq 2\varepsilon\, e^{-\frac{\beta N}{2}V(x)-N\delta_\varepsilon}\mu_{N-1,\beta}^{\frac{N}{N-1}V;[a_-,a_+]}\left(e^{\sum_{i=1}^{N-1}\frac{\beta}{2\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}\ln|\xi-\lambda_i|\,\mathrm{d}\xi}\right)
\end{aligned}$$

where we have finally used Jensen's inequality. But $\lambda\to\frac{1}{2\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}\ln|\xi-\lambda|d\xi$ is bounded continuous on $[a_-,a_+]$ and therefore by the large deviation principle for the law of the empirical measure L_{N-1} under $\mu_{N-1,\beta}^{\frac{N}{N-1}V;[a_-,a_+]}$ (with rate function which vanishes only at μ_{eq}) we deduce that:

$$\frac{1}{Y_N}\geq 2\varepsilon\, e^{-\frac{\beta N}{2}V(x)-2N\delta_\varepsilon}\, e^{(N-1)\int\frac{\beta}{2\varepsilon}\left(\int_{x-\varepsilon}^{x+\varepsilon}\ln|\xi-\lambda|\mathrm{d}\xi\right)\mathrm{d}\mu_{\mathrm{eq}}(\lambda)} \tag{36}$$

Hence, by taking ε sufficiently small independently of N, and optimizing over the choice of $x\in]a_-,a_+[$, we conclude that for any $\eta''>0$, and N large enough,

$$\frac{1}{Y_N}\geq e^{-N\left(\eta''+\beta\inf_{\xi\in[a_-,a_+]}\mathcal{J}^{V;[a_-,a_+]}(\xi)\right)} \tag{37}$$

Putting Eqn. 35 and 37 together, we deduce that for all $\delta>0$ and N large enough:

$$P_{N,\beta}^{V;[a_-,a_+]}(\lambda_{\min}\in F)\leq e^{N\beta\left(-\inf_{x\in F}\mathcal{J}^{V;[a_-,a_+]}(x)+\inf_{\xi\in[a_-,a_+]}\mathcal{J}^{V;[a_-,a_+]}(\xi)+\delta\right)} \tag{38}$$

which provides the announced upper bound.

Conclusion

As a consequence, since we assumed that the rate function only vanishes at α_-, α_+ we deduce that for any $\varepsilon > 0$, there exists $\delta_\varepsilon > 0$ so that:

$$\mu_{N,\beta}^{V,[a_-,a_+]}(\lambda_{\min} \le \alpha_- - \varepsilon) \le e^{-\delta_\varepsilon N} \tag{39}$$

as well as a similar result for the largest eigenvalue.

3.2.4 Lower bound for large deviation of extreme eigenvalues

To establish a lower bound, we start again from Eqn. 107 with an open ball $B =]x - \varepsilon, x + \varepsilon[\subset [a_-, \alpha_-]$:

$$P_{N,\beta}^{V,[a_-,a_+]}(\lambda_{\min} \in B) = Y_N \int_B \mathrm{d}\xi e^{-\frac{\beta}{2}V(\xi)} \Xi_N(\xi) \tag{40}$$

but replace the role of Y_N and Ξ_N in the bounds. Namely, we first have by Jensen's inequality:

$$\int_B \mathrm{d}\xi\, e^{-\frac{\beta}{2}V(\xi)} \Xi_N(\xi) \ge \kappa_N e^{\int \mathrm{d}\widetilde{\chi}(\xi,\lambda)\left(\beta \sum_{i=1}^{N-1} \ln|\xi-\lambda_i| - \frac{\beta}{2}(N-1)V(\xi)\right)} \tag{41}$$

with

$$\begin{aligned} \mathrm{d}\widetilde{\chi}(\xi,\lambda) &= \frac{\mathbf{1}_B(\xi)\,\mathbf{1}_{\lambda_{\min}\ge\xi}}{\kappa_N} \mathrm{d}\xi\, e^{-\frac{\beta}{2}V(\xi)}\, \mathrm{d}P_{N-1,\beta}^{\frac{N}{N-1}V;[a_-,a_+]}(\lambda) \\ \kappa_N &= \int_B \mathrm{d}\xi\, e^{-\frac{\beta}{2}V(\xi)} P_{N-1,\beta}^{\frac{N}{N-1}V;[a_-,a_+]}[\mathbf{1}_{\lambda_{\min}\ge\xi}] \end{aligned}$$

Thanks to Eqn. 39 (note that it applies similarly to $NV/(N-1)$ as the assumptions does not depend on the fine asymptotics of V), we know that κ_N converges towards a non vanishing constant. Moreover, the logarithm, once integrated against $\mathrm{d}\xi$, produces a smooth bounded function and therefore we can use the convergence of L_{N-1} towards μ_{eq} under $P_{N-1,\beta}^{\frac{N}{N-1}V;[a_-,a_+]}$ to conclude that:

$$\liminf_{N\to\infty} \frac{1}{N} \ln \int_B \mathrm{d}\xi\, e^{-\frac{\beta}{2}V(\xi)} \Xi_N(\xi) \ge -\frac{\beta}{2} \frac{\int_B \mathrm{d}\xi\, e^{-\frac{\beta}{2}V(\xi)} \left(V(\xi) - 2\int \mathrm{d}\mu_{\mathrm{eq}}(\eta) \ln|\xi-\eta|\right)}{\int_B \mathrm{d}\xi\, e^{-\frac{\beta}{2}V(\xi)}}. \tag{42}$$

Letting now ε going to zero in $B =]x-\varepsilon, x+\varepsilon[$ proves that:

$$\liminf_{\varepsilon\to 0} \liminf_{N\to\infty} \frac{1}{N} \ln \int_{B(x,\varepsilon)} \mathrm{d}\xi\, e^{-\frac{\beta}{2}V(\xi)}\, \Xi_N(\xi) \ge \beta\Big(\int \mathrm{d}\mu_{\mathrm{eq}}(\eta) \ln|\xi-\eta| - \frac{V(\eta)}{2}\Big) \tag{43}$$

To bound Y_N from below, it is enough to bound $1/Y_N$ from above, which can be done in the same way we bounded Ξ_N from above in the argument for the upper bound. We finally conclude:

$$\liminf_{\varepsilon\to 0} \liminf_{N\to\infty} \frac{1}{N} \ln \mu_{N,\beta}^{V,[a_-,a_+]}(\lambda_{\min} \in]x-\varepsilon, x+\varepsilon[) \ge -\beta \widetilde{\mathcal{J}}^{V;[a_-,b_+]}(x) \tag{44}$$

which completes the proof of the large deviation principle.

Exercise 36 (suggested by B. Collins)**.** Generalize the proof to obtain the large deviation principle for the joint law of the k^{th} largest eigenvalues (k finite) with good rate function given by

$$I^*(x_1,\cdots,x_k) = \sum_{l=1}^{k} I^*(x_k) - \beta \sum_{1\leq \ell \leq p \leq k} \log(x_\ell - x_k) + \text{constant}.$$

if $x_1 \geq x_2 \cdots \geq x_k \geq 2$ and $+\infty$ otherwise.

Bibliographical notes This proof is taken from [15].

3.3 Topological models of one-matrix models

Our goal is to prove an asymptotic expansion in powers of $1/N$ when $N \to \infty$ for the partition function $Z_{N,\beta}^{V;[b_-,b_+]}$ and the correlators $W_n^{V;[b_-,b_+]}(x_1,\ldots,x_n)$. This is not always expected. In particular it is false when the support of $\mu_{\text{eq}}^{V;[b_-,b_+]}$, the limiting eigenvalue distribution, is not connected: corrections to the leading order feature a quasi periodic behavior with N (see [53] for a general heuristic argument). Our proof uses a priori bounds on the correlators, and what we really need is to establish that $W_n \in O(1)$ for $n \geq 2$. We shall prove this condition either based on a result of Boutet de Monvel, Pastur and Shcherbina [43] (also used recently in the context of β ensembles by Kriecherbauer and Shcherbina [75]), or under the additional assumption that V is strictly convex. In the convex setting, such a priori bounds can be derived from concentration of measures properties, in which case our article is self-contained. Our basic assumptions and main results are:

Hypothesis 37.

- *(Regularity)* $V : [b_-,b_+] \to \mathbb{R}$ *is continuous, and if* V *depends on* N*, it has a limit* $V^{\{0\}}$ *in the space of continuous functions over* $[b_-,b_+]$ *for the sup norm.*
- *(Confinement) If* $b_\tau = \tau\infty$, $\liminf_{x\to\tau\infty} \frac{V(x)}{2\ln|x|} > 1$.
- *(One-cut regime) The support of* $\mu_{\text{eq}}^{V;[b_-,b_+]}$ *consists in a unique interval* $[\alpha_-,\alpha_+] \subseteq [b_-,b_+]$.
- *(Control of large deviations) The function* $x \in [b_-,b_+] \setminus]\alpha_-,\alpha_+[\mapsto \frac{1}{2}V(x) - \int \ln|x - \xi| d\mu_{\text{eq}}(\xi)$ *achieves its minimum value at* α_- *and* α_+ *only.*
- *(Offcriticality)* $S(x) > 0$ *whenever* $x \in [\alpha_-,\alpha_+]$*, where:*

$$S(x) = \pi \frac{\mathrm{d}\mu_{\text{eq}}}{\mathrm{d}x} \sqrt{\left| \frac{\prod_{\tau'\in\mathsf{Hard}}(x-\alpha_{\tau'})}{\prod_{\tau\in\mathsf{Soft}}(x-\alpha_\tau)} \right|} \tag{45}$$

  where $\tau \in \mathsf{Hard}$ *(resp.* $\tau \in \mathsf{Soft}$*) iff* $b_\tau = \alpha_\tau$ *(resp.* $\tau(b_\tau - \alpha_\tau) > 0$*).*

- *(Analyticity) V can be extended as a holomorphic function in some open neighborhood of $[\alpha_-, \alpha_+]$.*
- *V has a $1/N$ expansion in this neighborhood, in the sense of Hyp. 52.*

Notice that the "one-cut regime", "offcriticality" and "control of large deviations" assumptions automatically hold when V is strictly convex (see [73, Proposition 3.1], which extends easily to analytic functions instead of polynomials).

Proposition 38. *Assume Hyp. 37. Then, $W_n^{V;[b_-,b_+]}$ admits an asymptotic expansion when $N \to \infty$:*

$$W_n^{V;[b_-,b_+]}(x_1,\dots,x_n) = \sum_{k\geq n-2} N^{-k}\, W_n^{V;\{k\}}(x_1,\dots,x_n) \tag{46}$$

which has the precise meaning that, for all $K \geq n-2$:

$$W_n^{V;[b_-,b_+]}(x_1,\dots,x_n) = \sum_{k=n-2}^{K} N^{-k}\, W_n^{V;\{k\}}(x_1,\dots,x_n) + o(N^{-K}) \tag{47}$$

The $o(N^{-K})$ is uniform for $x_1,\dots,x_n$ in any compact of $(\mathbb{C}\setminus[b_-,b_+])^n$, but not uniform in n and K. Moreover, if $(b_\tau - \alpha_\tau)\tau > 0$ (meaning that α_τ is a soft edge), the functions $W_n^{V;\{k\}}$ are independent of b_τ chosen such that $(b_\tau - \alpha_\tau)\tau > 0$ and Hypotheses 37 hold.

Proposition 39. *Assume Hyp. 37, and $b_- < \alpha_- < \alpha_+ < b_+$ (all edges are soft). Then, $Z_{N,\beta}^{V;[b_-,b_+]}$ admits an asymptotic expansion when $N \to \infty$:*

$$Z_{N,\beta}^{V;[b_-,b_+]} = Z_{N,\mathrm{G}\beta\mathrm{E}} \left(\frac{\alpha_+ - \alpha_-}{4}\right)^{N+\beta\frac{N(N-1)}{2}} \exp\left(\sum_{k\geq -2} N^{-k} F_\beta^{V;\{k\}}\right) \tag{48}$$

In other words:

$$\forall K \geq -2 \quad Z_{N,\beta}^{V;[b_-,b_+]} = Z_{N,\mathrm{G}\beta\mathrm{E}} \left(\frac{\alpha_+ - \alpha_-}{4}\right)^{N+\beta\frac{N(N-1)}{2}} \exp\left(\sum_{k=-2}^{K} N^{-k} F_\beta^{V;\{k\}} + o(N^{-K})\right) \tag{49}$$

Moreover, the coefficients $F_\beta^{V;\{k\}}$ are independent of b_- and b_+ chosen such that $b_- < \alpha_- < \alpha_+ < b_+$ and Hypotheses 37 hold.

$Z_{N,\mathrm{G}\beta\mathrm{E}}$ is the partition function of the Gaussian β ensemble, defined by the quadratic potential $V_{\mathrm{G}}(x) = \frac{x^2}{2}$. It is given by a Selberg integral [93] :

$$Z_{N,\mathrm{G}\beta\mathrm{E}} = (2\pi)^{N/2}\,\big(N\beta/2\big)^{-\beta N^2/4 + (\beta/4 - 1/2)N}\, \frac{\prod_{j=1}^{N}\Gamma\big(1 + j\beta/2\big)}{\Gamma\big(1+\beta/2\big)^N} \tag{50}$$

For hard edges (i.e. $b_- = \alpha_-$ or $b_+ = \alpha_+$), one may still interpolate between $Z_{N,\beta}^{V;[b_-,b_+]}$ and a Gaussian β ensemble restricted to some interval (Corollary 64), but the partition function of the latter is not a Selberg integral and thus not known in closed form.

Commentary

When V does not depend on N and β, $W_n^{V;\{k\}}$ has a very simple dependence in β:

$$W_n^{V;\{k\}} = \sum_{g=0}^{\lfloor (k-n+2)/2 \rfloor} \left(\frac{\beta}{2}\right)^{1-g-n} \left(1-\frac{2}{\beta}\right)^{k+2-2g-n} \mathcal{W}_n^{V;(g;k+2-2g-n)} \tag{51}$$

and likewise:

$$F^{V;\{k\}} = C_{N,\beta} + \sum_{g=0}^{\lfloor k/2 \rfloor+1} \left(\frac{\beta}{2}\right)^{1-g} \left(1-\frac{2}{\beta}\right)^{k+2-2g} \mathcal{F}^{V;(g;k+2-2g)} \tag{52}$$

Assuming existence of the $1/N$ expansion, or at the level of formal matrix integrals, the recursive computation of the $\mathcal{W}_n^{V;(g;l)}$ and $\mathcal{F}^{V;(g;l)}$ was developed by Chekhov and Eynard in [39]. For $\beta = 2$, it is well-known that Eqn. 48 is an expansion in even powers of N, i.e. $F_{\beta=2}^{V;\{2k+1\}} = 0$. Such a result goes back to the so-called topological expansion of t'Hooft, shown in the context of matrix models by Brézin, Itzykson, Parisi and Zuber [32]. Indeed when $\beta = 2$, the sum in Eqn. 51 has only one term, namely $k = 2g-2+n$, which is present only when $k = n \bmod 2$, and likewise for Eqn. 52 which can be considered as the case $n = 0$.

At the asymptotic level, the case $\beta = 2$ was tackled in [3]. For $\beta = 1,2,4$, the partition function and the correlators can be computed with the help of orthogonal polynomials [83]. These are solutions of a Riemann-Hilbert problem [56], for which the large N asymptotics have been intensively studied [24, 46, 45, 44] with the steepest descent method introduced in [47]. As a consequence, Ercolani and McLaughlin [51] were able to prove the existence of a $1/N^2$ expansion of $\ln Z_{N,\beta=2}^V$. However, the topological expansions in the cases $\beta = 1$ and 4 are technically more involved in this framework, and have resisted to analysis up to now.

Integrability properties of β matrix models are unraveled for general $\beta > 0$, in particular there is no known orthogonal polynomials techniques to evaluate the partition function $Z_{N,\beta}^V$ and the correlation functions $W_n(x_1,\ldots,x_n)$. Yet, it is always possible to study the Cauchy-Stieltjes transform of the empirical measure of the eigenvalues and the "loop equations", also called Schwinger-Dyson equations or Pastur equations [86], that govern its expectations and cumulants. Thanks to the rough bounds for W_1^V and W_2^V established in [43], Johansson [73] proved a central limit theorem and obtained the first correction to W_1 when V is an even polynomial satisfying Hyp. 37. This was also the subject of a recent work by Kriecherbauer and Shcherbina [75], with Hyp. 37 only. These authors have obtained in particular the expansion of $\ln Z_{N,\beta}^V$ up to a $O(1)$ when $N \to \infty$ (see their Theorem 2).

The determination of $W_1^{V;\{-1\}}$ [104, 32, 16] and $W_2^{V;\{0\}}$ [8, 14, 73] has been known for long, in β ensembles or many other matrix models. It was also observed long ago [8] that, if a $1/N$ expansion is assumed to exist, the loop equations turn into a system of recursive linear equations determining fully the decaying orders. To solve it, one just has to invert a linear operator $\mathcal{K}$. Recursiveness is a consequence of the assumption or the fact that $W_n \in O(N^{2-n})$, which allows the determination of the leading order of W_{n-1} without knowledge of W_n (for $n \geq 3$). These techniques found their origin in [8, 7, 5, 6] and culminated with the formalism of the "topological recursion" of [52, 55] for $\beta = 2$, and [39] for any fixed $\beta > 0$.

In this article, we observe that $\mathcal{K}^{-1}$ is a continuous operator on some appropriate space of analytic functions. Combining with the a priori control on correlators which dates back to [43], we prove the existence of the full expansion.

For strictly convex potentials, concentration inequalities also provide rough bounds on the correlators and therefore allow us to give self-contained proofs,independent from [43]-[75]. In this framework, loop equations were used in [61] to establish the asymptotic expansion of models of several hermitian random matrices ($\beta = 2$) with strictly convex interactions. Maurel-Segala [78] also studied models of several symmetric random matrices ($\beta = 1$) with strictly convex interactions. In order to prove the asymptotic expansion, the main step of [61] was to show that some operator on non-commutative polynomials could be inverted, with bounded appropriate norm, and this was only done in a perturbative regime. Here, thanks to complex analysis, the potential need not be a small perturbation of the quadratic potential.

Our techniques could also be applied to other matrix models. For instance, the convergent β, $O(\mathfrak{n})$ matrix model:

$$\mathrm{d}P^{V;\mathbb{R}_+}_{N,\beta,O(\mathfrak{n})}(\lambda) = \frac{1}{Z^{V;\mathbb{R}_+}_{N,\beta,O(\mathfrak{n})}} \prod_{i=1}^{N} \mathrm{d}\lambda_i \, e^{-\frac{N\beta}{2}V(\lambda_i)} \, \frac{\prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j|^{\beta}}{\prod_{1 \leq i,j \leq N} (\lambda_i + \lambda_j)^{\mathfrak{n}/2}} \tag{53}$$

An important point is that the corresponding quadratic functional:

$$\mathcal{E}[\rho] = \iint_{\mathbb{R}_+^2} \mathrm{d}\rho(\xi)\mathrm{d}\rho(\eta)\Big[-\frac{\beta}{2}\ln|\xi-\eta| + \frac{\mathfrak{n}}{2}\ln|\xi+\eta|\Big] + \frac{\beta}{2}\int_{\mathbb{R}_+} \mathrm{d}\rho(\xi)\, V^{\{0\}}(\xi) \tag{54}$$

is strictly convex in the regime $|n| < \beta$, therefore ensuring uniqueness of its minimizer. Besides, the analytic tools for the recursive determination of the one-cut solution to the loop equations of the $O(\mathfrak{n})$ model in this regime were clarified in [27]. The existence of a full $1/N$ expansion for convergent $O(\mathfrak{n})$ matrix models under Hyp. 37 could probably be established by following the lines we are presenting for the β matrix models.

Outline

We first study in Section 3.4 the weak dependence in the bounds of integration under weak assumptions on V. In particular, we may trade the initial interval $[b_-, b_+]$ for a finite interval $[a_-, a_+]$. We then write in Section 3.5 the corresponding loop equations for the correlators. Section 3.6 is devoted to the proof of the asymptotic expansion of the correlators with slightly stronger assumptions (Prop. 53). They are weakened in Section 3.7 to complete the proof of our main results for the correlators (Prop. 38) and the free energy (Prop. 39). We also remind how early steps of our proof imply the central limit theorem of Johansson (Prop. 66).

3.4 Weak dependence on the soft edges

In this section we show that the partition function and the correlators depend very weakly on the boundary points of the interval of integration $[b_-, b_+]$ if they are soft, i.e. do not coincide with the boundary points of the support $[\alpha_-, \alpha_+]$ of the equilibrium measure. We show more precisely that this dependence yields only exponentially small corrections, by deriving a large deviation principle for the law of the extreme eigenvalues. This point was already studied in [9, section 2.6.2] under a technical assumption [9, Assumption 2.6.5] that we replace here by assuming that the rate function of our large deviation principle vanishes only at α_- and α_+. This result is not new in essence and not specific to the one-cut regime, see for instance [2, Proposition 2] which is proved with the extra assumption that V has bounded second derivatives in a neighborhood of $\mathrm{supp}\,\mu^{\mathrm{eq}}$, or [87, Proposition 11.1.4] which is proved with the extra assumption that V satisfies a Lipschitz condition in $[b_-, b_+]$.

Weak dependence on the soft edges

We first state the global version of the result:

Proposition 40. *Let $V : [b_-, b_+] \to \mathbb{R}$ be a continuous function, and if $b_\tau = \tau\infty$, assume that:*

$$\liminf_{x \to \tau\infty} \frac{V(x)}{2\ln|x|} > 1 \tag{55}$$

Suppose $b_- < \alpha_-$, and assume furthermore that the minimum value of $\mathcal{J}^{V;[b_-,b_+]}$ is achieved only on $[\alpha_-, \alpha_+]$. Then, for any $\varepsilon > 0$, there exists $\eta_\varepsilon > 0$ so that:

$$Z_{N,\beta}^{V;[b_-,b_+]} = Z_{N,\beta}^{V;[\alpha_- - \varepsilon, b_+]}\big(1 + O(e^{-N\eta_\varepsilon})\big), \tag{56}$$

and there exists a universal constant $\gamma_n > 0$ such that, for any $x_1, \ldots, x_n \in (\mathbb{C} \setminus [b_-, b_+])^n$:

$$\big|W_n^{V;[b_-,b_+]}(x_1, \ldots, x_n) - W_n^{V;[\alpha_- - \varepsilon, b_+]}(x_1, \ldots, x_n)\big| \leq \frac{\gamma_n\, e^{-N\eta_\varepsilon}}{\prod_{i=1}^n d(x_i, [b_-, b_+])} \tag{57}$$

A similar result holds for the upper edge.

We also have a local version:

Proposition 41. *Let $V : [b_-, b_+] \to \mathbb{R}$ be a continuous function, and if $b_\tau = \tau\infty$, assume that:*

$$\liminf_{x \to \tau\infty} \frac{V(x)}{2\ln|x|} > 1 \tag{58}$$

Suppose $b_- < \alpha_+$, and assume furthermore that the minimum value of $\mathcal{J}^{V;[b_-,b_+]}$ is achieved only on $[\alpha_-, \alpha_+]$. For any $\varepsilon > 0$ small enough, there exists $\eta_\varepsilon > 0$ so that, for any $a_- \in]b_-, \alpha_- - \varepsilon[$:

$$\Big|\partial_{a_-} \ln Z_N^{V;[a_-,b_+]}\Big| \leq e^{-N\eta_\varepsilon} \tag{59}$$

and, for any $x_1,\ldots,x_n \in (\mathbb{C}\setminus[a_-,b_+])^n$:

$$\forall x_1,\ldots,x_n \in \mathbb{C}\setminus[a_-,b_+],\quad \Big|\partial_{a_-} W_n^{V;[a_-,b_+]}(x_1,\ldots,x_n)\Big| \leq \frac{\gamma_n N^n}{\prod_{i=1}^n d(x_i,[a_-,b_+])}\, e^{-N\eta_\varepsilon} \tag{60}$$

A similar statement holds for derivatives with respect to the upper bound.

Proof. If $b_- \neq \alpha_-$, let $a_- \in]b_-,\alpha_-[$. Notice that:

$$\left(1-\frac{Z_{N,\beta}^{V;[a_-,b_+]}}{Z_{N,\beta}^{V;[b_-,b_+]}}\right) = P_{N,\beta}^{V;[b_-,b_+]}[\lambda_{\min} \leq a_-] \tag{61}$$

If now $\varphi\,:\,[b_-,b_+]^N \to \mathbb{C}$ is a bounded continuous function, we can write:

$$P_{N,\beta}^{V;[b_-,b_+]}[\varphi(\lambda)] - P_{N,\beta}^{V;[a_-,b_+]}[\varphi(\lambda)] = P_{N,\beta}^{V;[b_-,b_+]}\big[\varphi(\lambda)\,\mathbf{1}_{\lambda_{\min}\leq a_-}\big] + \left(\frac{Z_{N,\beta}^{V;[a_-,b_+]}}{Z_{N,\beta}^{V;[b_-,b_+]}} - 1\right) P_{N,\beta}^{V;[a_-,b_+]}[\varphi(\lambda)] \tag{62}$$

Thus, we find:

$$\big|P_{N,\beta}^{V;[b_-,b_+]}[\varphi(\lambda)] - P_{N,\beta}^{V;[a_-,b_+]}[\varphi(\lambda)]\big| \quad\leq\quad 2\left(\sup_{\lambda\in[b_-,b_+]^N}|\varphi(\lambda)|\right) P_{N,\beta}^{V;[b_-,b_+]}[\lambda_{\min}\leq a_-]$$

This can be applied for the disconnected correlators:

$$\overline{W}_n^{V;[b_-,b_+]}(x_1,\ldots,x_n) = P_{N,\beta}^{V;[b_-,b_+]}\Big[\prod_{j=1}^n \sum_{i_j=1}^N \frac{1}{x_j-\lambda_{i_j}}\Big] \tag{63}$$

and we obtain:

$$\big|\overline{W}_n^{V;[b_-,b_+]}(x_1,\ldots,x_n) - \overline{W}_n^{V;[a_-,b_+]}(x_1,\ldots,x_n)\big| \leq \frac{2N^n}{\prod_{j=1}^n d(x_j,[b_-,b_+])}\, P_{N,\beta}^{V;[b_-,b_+]}[\lambda_{\min}\leq a_-] \tag{64}$$

Similarly, one finds:

$$\big|\partial_{a_-} W_n^{V;[a_-,b_+]}(x_1,\ldots,x_n)\big| \leq \frac{2N^n}{\prod_{j=1}^n d(x_j,[a_-,b_+])}\,\partial_{a_-}\ln Z_{N,\beta}^{V;[a_-,b_+]} \tag{65}$$

The correlators $W_n^{V;[b_-,b_+]}$ are just sums of monomials of the form $W_{n_1}^{V;[b_-,b_+]}(I_1)\cdots W_{n_m}^{V;[b_-,b_+]}(I_m)$ where $I_1,\ldots,I_m$ is a partition of $\{x_1,\ldots,x_n\}$. So, it is enough to establish the weak dependence at the level of the partition function. The global version is a direct consequence of Eqn. 26 applied to Eqn. 61:

$$\limsup_{N\to\infty}\frac{1}{N}\ln\left(1-\frac{Z_{N,\beta}^{V;[a_-,b_+]}}{Z_{N,\beta}^{V;[b_-,b_+]}}\right) < 0 \tag{66}$$

For the local version, we rather need to bound:

$$\partial_{a_-}\ln Z_{N,\beta}^{V;[a_-,b_+]} = N\,\frac{Z_{N-1,\beta}^{\frac{NV}{N-1};[a_-,b_+]}}{Z_{N,\beta}^{V;[a_-,b_+]}}\, P_{N-1,\beta}^{\frac{NV}{N-1};[a_-,b_+]}\left[e^{\beta\left(-\frac{NV(a_-)}{2}+\sum_{i=1}^{N-1}\ln|\lambda_i-a_-|\right)}\right] \tag{67}$$

If $a_- \in]b_-, \alpha_-[$ is fixed, by the large deviation principle for L_{N-1} under $P_{N-1}^{\frac{NV}{N-1};[a_-,b_+]}$, since the logarithm is a lower semicontinuous function, there exists $\gamma > 0$ such that, for any $\varepsilon > 0$, for N large enough:

$$P_{N-1,\beta}^{\frac{NV}{N-1};[a_-,b_+]}\left[e^{\beta\left(-\frac{NV(a_-)}{2}+\sum_{j=1}^{N-1}\ln|a_- - \lambda_j|\right)}\right] \leq \gamma e^{-\beta N(1-\varepsilon)\mathcal{J}^{V;[b_-,b_+]}(a_-)} \tag{68}$$

Moreover, we have seen in Eqn. 37 that for N large enough:

$$\frac{Z_{N-1,\beta}^{\frac{NV}{N-1};[a_-,b_+]}}{Z_{N,\beta}^{V;[a_-,b_+]}} \leq e^{\beta N(1-\varepsilon)\inf_{\xi\in[b_-,\alpha_-]}\mathcal{J}_{\min}^{V;[a_-,b_+]}(\xi)} \tag{69}$$

By assumption, $\widetilde{\mathcal{J}}_{\min}^{V;[b_-,b_+]}(a_-) = \mathcal{J}_{\min}^{V;[b_-,b_+]}(a_-) - \inf_{\xi\in[b_-,\alpha_-]}\mathcal{J}^{V;[b_-,b_+]}(\xi) > 0$, leading to:

$$\limsup_{N\to\infty}\frac{1}{N}\ln\left|\partial_{a_-}\ln Z_{N,\beta}^{V;[a_-,b_+]}\right| < 0 \tag{70}$$

which is the bound we sought. The arguments at the upper edge are similar. $\diamond$

3.5 Loop equations

We shall assume in this Section and also in Section 3.6 that the λ_i are integrated over a segment $[a_-, a_+]$ with:

Hypothesis 42. $-\infty < a_- < a_+ < +\infty$.

Indeed, considering finite intervals $[a_-, a_+]$ is convenient to ensure from the beginning that the Cauchy-Stieltjes transform yields functions which are holomorphic outside $[a_-, a_+]$. We also assume in this section:

Hypothesis 43. $V : [a_-, a_+] \to \mathbb{C}$ *can be extended as a holomorphic function in some open neighborhood of* $[a_-, a_+]$.

This will allow us to use complex analysis (Cauchy residue formula, moving the contours, etc.)

We shall derive the ”loop equations”, also called Schwinger-Dyson equations or Pastur equations [86] in this context. These equations express the invariance by change of variable of an integration, up to boundary terms. We stress that these equations are exact for finite N. Although the technique is well-known, we recall the derivation here for the β matrix models with edges a_-, a_+ in order to have a self-contained presentation.

3.5.1 First version

Theorem 44. *Loop equation at rank* 1. *For any $x \in \mathbb{C} \setminus [a_-, a_+]$:*

$$\begin{aligned} & W_2(x,x) + \big(W_1(x)\big)^2 + \Big(1 - \frac{2}{\beta}\Big)\frac{\mathrm{d}}{\mathrm{d}x}\Big(W_1(x)\Big) \\ + \frac{N(1-2/\beta) - N^2}{(x-a_-)(x-a_+)} - N\left(\oint_{\mathcal{C}([a_-,a_+])} \frac{\mathrm{d}\xi}{2i\pi}\,\frac{1}{x-\xi}\,\frac{(\xi-a_-)(\xi-a_+)}{(x-a_-)(x-a_+)}\,V'(\xi)\,W_1(\xi)\right) & = 0 \end{aligned}$$

$\mathcal{C}([a_-,a_+])$ is a contour surrounding $[a_-,a_+]$ in positive orientation, and included in the domain where V' is holomorphic.

Theorem 45. *Loop equation at rank n. Let $x_I = (x_i)_{i \in I}$ be a $(n-1)$-uple of spectator variables in $(\mathbb{C} \setminus [a_-,a_+])^{n-1}$. For any $x \in \mathbb{C} \setminus [a_-,a_+]$:*

$$\begin{aligned} & W_{n+1}(x,x,x_I) + \sum_{J \subseteq I} W_{|J|+1}(x,x_J)\,W_{n-|J|}(x,x_{I\setminus J}) + \Big(1-\frac{2}{\beta}\Big)\frac{\mathrm{d}}{\mathrm{d}x}\Big(W_n(x,x_I)\Big) \\ & -N\left(\oint_{\mathcal{C}([a_-,a_+])} \frac{\mathrm{d}\xi}{2i\pi}\,\frac{1}{x-\xi}\,\frac{(\xi-a_-)(\xi-a_+)}{(x-a_-)(x-a_+)}\,V'(\xi)\,W_n(\xi,x_I)\right) \\ & + \frac{2}{\beta}\sum_{i \in I}\frac{\mathrm{d}}{\mathrm{d}x_i}\Big(\frac{W_{n-1}(x,x_{I\setminus\{i\}}) - \frac{(x_i-a_-)(x_i-a_+)}{(x-a_-)(x-a_+)}\,W_{n-1}(x_I)}{x-x_i}\Big) \quad = \quad 0 \end{aligned}$$

Proof of Theorem 44. For any smooth real-valued function h, and $\varepsilon > 0$ small enough,

$$\psi_{h,\varepsilon} : \lambda \mapsto \lambda + \varepsilon h(\lambda) \tag{71}$$

defines a differentiable family of diffeomorphisms from $[a_-,a_+]$ to some interval $\psi_{h,\varepsilon}([a_-,a_+])$. We assume hereafter that $h(a_-) = h(a_+) = 0$ so that $\psi_{h,\varepsilon}([a_-,a_+]) = [a_-,a_+]$ for ε small enough. We have:

$$1 = \int_{[a_-,a_+]^N} \mathrm{d}\mu_{N,\beta}^{V}\big(\psi_{h,\varepsilon}(\lambda_1),\ldots,\psi_{h,\varepsilon}(\lambda_N)\big) \tag{72}$$

When $\varepsilon \to 0$, the first subleading order of the right hand side must vanish. It can be computed in three parts. A first term comes from the variation of the Lebesgue measure $\prod_i \mathrm{d}\lambda_i$, which is given by the Jacobian of the change of variable:

$$\Big(\prod_{i=1}^{N}\mathrm{d}\psi_{h,\varepsilon}(\lambda_i)\Big) = \Big(\prod_{i=1}^{N}\mathrm{d}\lambda_i\Big)\Big(1 + \varepsilon\int h'(\xi)\mathrm{d}M_N(\xi) + o(\varepsilon)\Big) \tag{73}$$

A second term comes from the variation of the Vandermonde:

$$\begin{aligned} |\Delta\big(\psi_{h,\varepsilon}(\lambda)\big)|^{\beta} &= |\Delta(\lambda)|^{\beta}\Big[1 + \varepsilon\beta\sum_{1 \le i < j \le N}\frac{h(\lambda_i)-h(\lambda_j)}{\lambda_i-\lambda_j} + o(\varepsilon)\Big] \\ = \quad & |\Delta(\lambda)|^{\beta}\Big\{1 + \varepsilon\frac{\beta}{2}\Big(\iint\frac{h(\xi)-h(\eta)}{\xi-\eta}\mathrm{d}M_N(\xi)\mathrm{d}M_N(\eta) - \int h'(\xi)\mathrm{d}M_N(\xi)\Big) + o(\varepsilon)\Big\} \end{aligned}$$

The last term comes from the variation of the Boltzmann weight:

$$\prod_{i=1}^{N} e^{-\frac{N\beta}{2}V[\psi_{h,\varepsilon}(\lambda_i)]} = \Big(\prod_{i=1}^{N} e^{-\frac{N\beta}{2}V(\lambda_i)}\Big)\Big(1-\varepsilon\frac{N\beta}{2}\int V'(\xi)\,h(\xi)\mathrm{d}M_N(\xi)+o(\varepsilon)\Big) \quad (74)$$

Summing all terms up, the first order in ε in Eqn. 72 vanishes iff:

$$\begin{aligned} & P_{N,\beta}^{V;[a_-,a_+]}\Big[\iint \frac{h(\xi)-h(\eta)}{\xi-\eta}\,\mathrm{d}M_N(\xi)\mathrm{d}M_N(\eta) - N\int V'(\xi)h(\xi)\,\mathrm{d}M_N(\xi)\Big] \\ = \quad & \Big(1-\frac{2}{\beta}\Big)P_{N,\beta}^{V;[a_-,a_+]}\Big[\int h'(\xi)\,\mathrm{d}M_N(\xi)\Big] \end{aligned}$$

Note that even though this equation was obtained for real-valued functions h, we can at this point remove this condition by linearity. To obtain an equation involving correlators, one can take for $x \in \mathbb{C}\setminus[a_-,a_+]$ the function h defined by:

$$h(\xi) = \frac{(\xi-a_-)(\xi-a_+)}{x-\xi} = \frac{(x-a_-)(x-a_+)}{x-\xi} + a_- + a_+ - x - \xi \quad (75)$$

thus preserving $[a_-,a_+]$. We recall that V' is holomorphic in a neighborhood of $[a_-,a_+]$. So, by Cauchy formula, for any contour $\mathcal{C}([a_-,a_+])$ surrounding $[a_-,a_+]$ inside this neighborhood and not enclosing x:

$$\int \frac{V'(\xi)(\xi-a_-)(\xi-a_+)}{x-\xi}\,\mathrm{d}M_N(\xi) = \int \mathrm{d}M_N(\xi)\oint_{\mathcal{C}([a_-,a_+])}\frac{\mathrm{d}\eta}{2i\pi}\,\frac{V'(\eta)(\eta-a_-)(\eta-a_+)}{(\eta-\xi)(x-\eta)} \quad (76)$$

Hence, we obtain:

$$\begin{aligned} & W_2(x,x) + \big(W_1(x)\big)^2 - \frac{N^2}{(x-a_-)(x-a_+)} \\ & -N\oint_{\mathcal{C}([a_-,a_+])}\frac{\mathrm{d}\xi}{2i\pi}\,\frac{1}{x-\xi}\,\frac{(\xi-a_-)(\xi-a_+)}{(x-a_-)(x-a_+)}\,V'(\xi)\,W_1(\xi) \\ = \quad & \Big(1-\frac{2}{\beta}\Big)\Big(-\frac{\mathrm{d}}{\mathrm{d}x}\big(W_1(x)\big) - \frac{N}{(x-a_-)(x-a_+)}\Big) \end{aligned}$$

Proof of Theorem 45. By definition of the cumulants, if we define a shifted potential $V_{(x;\varepsilon)}(\xi) = V(\xi) + \frac{\varepsilon}{x-\xi}$:

$$W_n^V(x,x_2,\ldots,x_n) = -\frac{2}{\beta N}\partial_\varepsilon\Big(W_{n-1}^{V_{(x;\varepsilon)}}(x_2,\ldots,x_n)\Big)\Big|_{\varepsilon=0} \quad (77)$$

Notice that the matrix integral with this shifted potential is still convergent, because the eigenvalues live on the finite interval $[a_-,a_+]$. Therefore, we can obtain the loop equations at rank n by taking a perturbed potential in Thm. 44:

$$V_{(x_2;\varepsilon_2),\ldots,(x_n;\varepsilon_n)}(\xi) = V(\xi) + \sum_{i=2}^{n}\frac{\varepsilon_i}{x_i-\xi} \quad (78)$$

and identifying the term in $\Big[\prod_{i=2}^{n}\big(\frac{-2}{\beta N}\big)\varepsilon_i\Big]$ when $\varepsilon_i \to 0$.

3.5.2 Second version

Here is another equivalent form of the loop equations. All W_n depend implicitly on the interval of integration $[a_-,a_+]$.

Theorem 46. *Loop equation at rank* 1. *For any $x\in\mathbb{C}\setminus[a_-,a_+]$:*

$$W_2(x,x)+\big(W_1(x)\big)^2+\Big(1-\frac{2}{\beta}\Big)\frac{\mathrm{d}}{\mathrm{d}x}\Big(W_1(x)\Big)$$
$$-N\left(\oint_{C([a_-,a_+])}\frac{\mathrm{d}\xi}{2i\pi}\,\frac{V'(\xi)\,W_1(\xi)}{x-\xi}\right)-\frac{2}{\beta}\left(\frac{\partial_{a_-}\ln Z}{x-a_-}+\frac{\partial_{a_+}\ln Z}{x-a_+}\right)\quad=\quad 0$$

$C([a_-,a_+])$ is a contour surrounding $[a_-,a_+]$ in positive orientation, and included in the domain where V is holomorphic.

Theorem 47. *Loop equation at rank n. Let $x_I=(x_i)_{i\in I}$ a $(n-1)$-uple of spectator variables in $(\mathbb{C}\setminus[a_-,a_+])^{n-1}$. For any $x\in\mathbb{C}\setminus[a_-,a_+]$:*

$$W_{n+1}(x,x,x_I)+\sum_{J\subseteq I}W_{|J|+1}(x,x_J)\,W_{n-|J|}(x,x_{I\setminus J})+\left(1-\frac{2}{\beta}\right)\frac{\mathrm{d}}{\mathrm{d}x}\Big(W_n(x,x_I)\Big)$$
$$-N\left(\oint_{C([a_-,a_+])}\frac{\mathrm{d}\xi}{2i\pi}\,\frac{V'(\xi)\,W_n(\xi,x_I)}{x-\xi}\right)$$
$$+\frac{2}{\beta}\sum_{i\in I}\frac{\mathrm{d}}{\mathrm{d}x_i}\left(\frac{W_{n-1}(x,x_{I\setminus\{i\}})-W_{n-1}(x_I)}{x-x_i}\right)-\frac{2}{\beta}\left(\frac{\partial_{a_-}W_{n-1}(x_I)}{x-a_-}+\frac{\partial_{a_+}W_{n-1}(x_I)}{x-a_+}\right)\quad=\quad 0$$

Proof In the former proof, if we use a change of variable h which does not preserve $[a_-,a_+]$, the partition function becomes (to first order in ε):

$$Z_N^{V;\psi_{h,\varepsilon}([a_-,a_+])}\to Z_N^{V;[a_-,a_+]}\left[1+\varepsilon\left(h(a_-)\,\partial_{a_-}\ln Z_N^{V;[a_-,a_+]}+h(a_+)\,\partial_{a_+}\ln Z_N^{V;[a_-,a_+]}\right)+o(\varepsilon)\right]\tag{79}$$

Thus, Eqn. 75 receives those extra terms, and becomes:

$$P_{N,\beta}^{V;[a_-,a_+]}\Big[\iint\frac{h(\xi)-h(\eta)}{\xi-\eta}\,\mathrm{d}M_N(\xi)\mathrm{d}M_N(\eta)-N\int V'(\xi)h(\xi)\,\mathrm{d}M_N(\xi)\Big]$$
$$=\quad\Big(1-\frac{2}{\beta}\Big)P_{N,\beta}^{V;[a_-,a_+]}\Big[\int h'(\xi)\,\mathrm{d}M_N(\xi)\Big]+\frac{2}{\beta}\left(h(a_-)\,\partial_{a_-}\ln Z_N^{V;[a_-,a_+]}+h(a_+)\,\partial_b\ln Z_N^{V;[a_-,a_+]}\right)$$

In particular, when we choose $h(\xi)=\frac{1}{x-\xi}$, we obtain:

$$W_2(x,x)+\big(W_1(x)\big)^2-N\oint_{C([a_-,a_+])}\frac{\mathrm{d}\xi}{2i\pi}\,\frac{V'(\xi)W_1(\xi)}{x-\xi}$$
$$=\quad-\left(1-\frac{2}{\beta}\right)\frac{\mathrm{d}}{\mathrm{d}x}\Big(W_1(x)\Big)+\frac{2}{\beta}\,\frac{\partial_{a_-}\ln Z}{x-a_-}+\frac{2}{\beta}\,\frac{\partial_{a_+}\ln Z}{x-a_+}$$

The loop equation at higher rank can be deduced as before by perturbing the potential.
◇

3.5.3 Remark

If we compare those expressions to the first version of the loop equations, we find by consistency:

$$\partial_{a_\tau} \ln Z = \frac{1}{a_{-\tau} - a_\tau} \left\{ -\frac{N^2\beta}{2} + N\left(\frac{\beta}{2} - 1\right) + \frac{N\beta}{2} \oint_{\mathcal{C}([a_-,a_+])} \frac{\mathrm{d}\xi}{2i\pi} (\xi - a_{-\tau}) V'(\xi) W_1(\xi) \right\} \tag{80}$$

and for higher correlators $\partial_{a_\tau} W_{n-1}(x_I)$ equals

$$\frac{1}{a_{-\tau} - a_\tau} \left\{ \frac{N\beta}{2} \oint_{\mathcal{C}([a_-,a_+])} \frac{\mathrm{d}\xi}{2i\pi} (\xi - a_{-\tau}) V'(\xi) W_n(\xi) + \sum_{i \in I} \frac{\mathrm{d}}{\mathrm{d}x_i} \Big((x_i - a_{-\tau}) W_{n-1}(x_I) \Big) \right\} \tag{81}$$

for $\tau \in \{\pm\}$.

3.6 The $1/N$ expansion

3.6.1 Notations, assumptions, proposition

This section relies on complex analysis and inequalities for probability measures. We make four assumptions on the potential V, which are valid only in this section. The link with our main theorem will be done in Section 3.7.

We keep on with the assumption:

Hypothesis 48. $-\infty < a_- < a_+ < +\infty$.

Since V is smooth, the equilibrium measure $\mathrm{d}\mu_{\mathrm{eq}}^{V;[a_-,a_+]}(\xi)$ will in fact be a density $\rho(\xi)\mathrm{d}\xi$, where $\rho : [a_-,a_+] \to [0,+\infty]$ is a continuous function. We call $\mathrm{supp}\,\rho = \{x \in [a_-,a_+] \quad \rho(x) > 0\}$ its support. In the hermitian case ($\beta = 2$), a $1/N$ expansion is expected only when $\mathrm{supp}\,\rho$ is connected. We assume here also:

Hypothesis 49. *V leads to a one-cut regime, i.e. the support of $\mu_{\mathrm{eq}}^{V;[a_-,a_+]}$ is an interval $[\alpha_-,\alpha_+] \subseteq [a_-,a_+]$.*

In order to write the loop equations as in Section 3.5, we assume:

Hypothesis 50. *V is real-valued on $[a_-,a_+]$, and can be extended as a holomorphic function on some open neighborhood U of $[a_-,a_+]$.*

We justify in Remark 60 later that there exists a unique analytic function $y : U \to \mathbb{C} \cup \{\infty\}$ such that, for any $x \in [\alpha_-,\alpha_+]$, we have:

$$\rho(x) = \frac{1}{i\pi} \lim_{\varepsilon \to 0^+} y(x + i\varepsilon) \tag{82}$$

This function can be written $y(x) = S(x)\sigma(x)$, where S is now a holomorphic function defined on U, and σ is of the form:

$$\sigma(x) = \sqrt{\frac{\prod_{\tau \in \mathsf{Soft}}(x - \alpha_\tau)}{\prod_{\tau' \in \mathsf{Hard}}(x - \alpha_{\tau'})}} \tag{83}$$

The lower edge a_- is

- either a hard edge, meaning that $a_- = \alpha_-$. Then, $\rho(x) \in O\big((x - \alpha_-)^{-1/2}\big)$ when $x \to \alpha_-$.
- or a soft edge, meaning that $a_- < \alpha_-$. Then, $\rho(x) \in O\big((x - \alpha_-)^{1/2}\big)$ when $x \to \alpha_-$.

and the same distinction exists independently for the upper edge a_+. Our discussion holds for both hard and soft cases. However, a key technical assumption is:

Hypothesis 51. *V is offcritical on $[a_-, a_+]$, in the sense that $S(x)$ remains positive on $[a_-, a_+]$.*

For instance, Hyp. 49 and 51 automatically hold when V is strictly convex. For a generic V satisfying Hyp. 49, we have $S(\alpha_-) > 0$ and $S(\alpha_+) > 0$, so we can always find an interval $[a_-, a_+]$ which is a strict enlargment of $[\alpha_-, \alpha_+]$, such that Hyp. 51 holds on $[a_-, a_+]$. We call "critical point on $[a_-, a_+]$", the situation corresponding to a choice of V such that S has a zero on $[a_-, a_+]$. In this article, we do not tackle the question of the double scaling limit for β matrix models ($N \to +\infty$ and coefficients of V finely tuned with N to achieve a critical point when $N = \infty$). Though, this would be a very interesting regime in relation with universality questions, considering the absence of Riemann-Hilbert techniques when $\beta \neq 1, 2, 4$.

We shall allow V itself to depend on N and have a $1/N$ expansion. To give precise statements about those expansions, we need some notations. For any Jordan curve Γ, we note $\mathrm{Ext}(\Gamma)$ (resp. $\mathrm{Int}(\Gamma)$) the unbounded (resp. bounded) connected component of $\mathbb{C} \setminus \Gamma$. In the following, we fix once for all a Jordan curve Γ_E, and a sequence of nested Jordan curves $(\Gamma_l)_{l \in \mathbb{N}}$, which all live in $\mathbb{C} \setminus [a_-, a_+]$, and such that

(i) $\Gamma_E \subseteq U$.

(ii) $\{x \in U \quad S(x) = 0\} \cap \mathrm{Int}(\Gamma_E) = \emptyset$.

(iii) $\forall l \in \mathbb{N} \quad \Gamma_l \subseteq \mathrm{Int}(\Gamma_{l+1})$.

(iv) $\forall l \in \mathbb{N} \quad \Gamma_l \subseteq \mathrm{Int}(\Gamma_E)$.

The contour configuration is depicted in Fig. 3.1, where the zeroes of S were called s_i. In the remaining of the text, Γ will refer to a Jordan curve in $\mathrm{Int}(\Gamma_E) \setminus [a_-, a_+]$. We will use the following norm on the space $\mathcal{H}^{(1)}_{n;[a_-,a_+]}$ of holomorphic functions on $(\mathbb{C} \setminus [a_-, a_+])^n$, which behave as $O(1/x_i)$ when $x_i \to \infty$.

$$\varphi f \varphi_\Gamma = \sup_{x_i \in \mathrm{Ext}(\Gamma)} |f(x_1, \dots, x_n)| = \sup_{x_i \in \Gamma} |f(x_1, \dots, x_n)| \tag{84}$$

The second equality is a consequence of the maximum principle. One can easily derive the following useful inequalities:

$$\forall f \in \mathcal{H}^{(1)}_{1;[a_-,a_+]} \quad \forall x_0 \in \mathrm{Ext}(\Gamma_{l+1}) \quad \forall l \in \mathbb{N} \qquad \left\| \frac{f(\bullet) - f(x_0)}{\bullet - x_0} \right\|_{\Gamma_l} \leq \varphi f' \varphi_{\Gamma_l+1} \leq \zeta_l \, \varphi f \varphi_{\Gamma_l} \tag{85}$$

where $\zeta_l = \frac{\ell(\Gamma_l)}{2\pi d^2(\Gamma_l, \Gamma_{l+1})}$ is a finite constant depending only on the relative position of Γ_l and Γ_{l+1}.

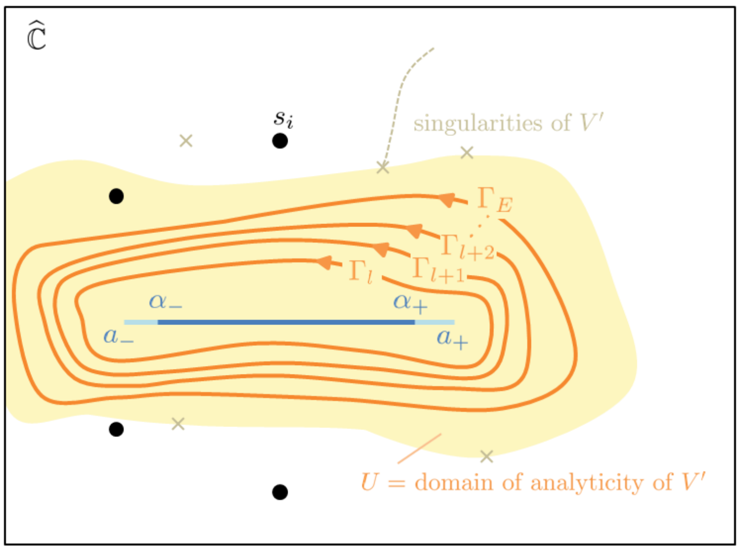


Figure 3.1: Hypothesis on the location of s_i and contour configurations

Now, we can state our last assumption:

Hypothesis 52. *V admits a $1/N$ asymptotic expansion:*

$$V(x) \underset{\Gamma_E}{=} \sum_{k \geq 0} N^{-k} V^{\{k\}}(x) \tag{86}$$

with functions $V^{\{k\}}$ independent of N, such that $V^{\{k\}}$ is holomorphic in U. The $=_{\Gamma_E}$ equality means that, for any positive integer K, there exists a positive constant v_K such that, for N large enough:

$$\sup_{\xi \in \Gamma_E} \left| V(\xi) - \sum_{k=0}^{K} N^{-k} V^{\{k\}}(\xi) \right| \leq N^{-(K+1)} v_K \tag{87}$$

(The maximum principle implies automatically the same bound with Γ replacing Γ_E as V is analytic in $\mathrm{Int}(\Gamma_E)$).

In many applications, V is independent of N (i.e. $V \equiv V^{\{0\}}$). There is however no difficulty in our reasoning to consider potentials which depend on N within Hyp. 52.

Our intermediate result is:

Theorem 53. *If Hyp. 48-52 hold, the correlators admit an asymptotic expansion when $N \to \infty$ with respect to the norm $\varphi \cdot \varphi_{\Gamma_E}$, of the form:*

$$\forall n \geq 1, \qquad W_n = \sum_{k \geq n-2} N^{-k}\, W_n^{\{k\}} \tag{88}$$

where $W_n^{\{k\}} \in \mathcal{H}^{(1)}_{n;[\alpha_-,\alpha_+]}$.

3.6.2 Relevant linear operators

The operator $\mathcal{K}$

We introduce the following linear operator defined on the space $\mathcal{H}^{(2)}_{1;[a_-,a_+]}$ of holomorphic functions on $\mathbb{C}\setminus[a_-,a_+]$ which behave as $O(1/x^2)$ when $x \to \infty$:

$$(\mathcal{K}f)(x) = 2W_1^{\{-1\}}(x)f(x) - \oint_{\mathcal{C}([a_-,a_+])} \frac{\mathrm{d}\xi}{2i\pi}\,\frac{L(\xi)}{L(x)}\Big(\frac{1}{x-\xi}+c\Big)\,(V^{\{0\}})'(\xi)\,f(\xi) \tag{89}$$

This operator for an appropriate choice of L and c appears in the loop equations. We have found the following choice convenient:

$$\begin{aligned} L(x) &= \prod_{\tau \in \mathsf{Hard}} (x - a_\tau) \\ c &= \begin{cases} 0 & \text{if } \mathsf{Soft} = \{\pm\} \text{ or } \mathsf{Hard} = \{\pm\} \\ \frac{1}{a_\tau - a_{-\tau}} & \text{if } \tau \in \mathsf{Soft} \text{ and } (-\tau) \in \mathsf{Hard} \end{cases} \end{aligned}$$

We may also rewrite:

$$(\mathcal{K}f)(x) = -2y(x)\,f(x) + \frac{(\mathcal{Q}f)(x)}{L(x)} \tag{90}$$

with:

$$(\mathcal{Q}f)(x) = -\oint_{\mathcal{C}([a_-,a_+])\cup\mathcal{C}(x)} \frac{\mathrm{d}\xi}{2i\pi}\,L(\xi)\Big(\frac{1}{x-\xi}+c\Big)(V^{\{0\}})'(\xi)\,f(\xi) \tag{91}$$

where $\mathcal{C}(x)$ is a contour surrounding x only (computing a residue at x). For any $f \in \mathcal{H}^{(1)}_{1;[a_-,a_+]}$, $(\mathcal{Q}f)$ is analytic, with singularities only where $(V^{\{0\}})'$ has singularities, in particular is holomorphic in the neighborhood of $[a_-,a_+]$. We have set:

$$y(x) = -W_1^{\{-1\}}(x) + \frac{(V^{\{0\}})'(x)}{2} \tag{92}$$

y is discontinuous on the support of $\mu_{\mathrm{eq}}^{V;[a_-,a_+]}$, i.e. on $[\alpha_-,\alpha_+] \subseteq [a_-,a_+]$, but analytic on $\mathbb{C}\setminus[\alpha_-,\alpha_+]$. We justify in Remark 60 that $y(x) = S(x)\sigma(x)$ where $\sigma(x)$ was introduced in Eqn. 83 and the squareroot is chosen with its usual discontinuity on $\mathbb{R}_-$. Let us call $s_i \neq \alpha_-, \alpha_+$ the zeroes of $S(x)$ in the complex plane, and we assume that they do not lie in $[a_-, a_+]$ (Hyp. 51).

It is clear that $\mathrm{Im}\,\mathcal{K} \subseteq \mathcal{H}^{(1)}_{1;[a_-,a_+]}$. Here, $W_1^{\{-1\}}$ (hence y) has only cut $[\alpha_-, \alpha_+]$, and this operator is invertible[1]. Its inverse can be explicitly written, it is given by Tricomi formula [98]:

$$\forall x \in \mathbb{C} \setminus [a_-, a_+], \quad \forall g \in \mathrm{Im}\,\mathcal{K}, \qquad \big(\mathcal{K}^{-1} g\big)(x) = \oint_{\mathcal{C}([a_-,a_+])} \frac{\mathrm{d}\xi}{2i\pi} \frac{1}{\xi - x} \frac{\widetilde{\sigma}(\xi)}{\widetilde{\sigma}(x)} \frac{g(\xi)}{2y(\xi)}$$

where $\widetilde{\sigma}(x) = \sqrt{(x-\alpha_-)(x-\alpha_+)}$, and where we integrate over a contour surrounding $[a_-, a_+]$ but not x. Indeed, if $g \in \mathrm{Im}\,\mathcal{K}$, we can write for any $x \in \mathbb{C} \setminus [a_-, a_+]$:

$$\begin{aligned}
\widetilde{\sigma}(x)\, f(x) &= \operatorname*{Res}_{\xi \to x} \frac{\mathrm{d}\xi}{\xi - x} \widetilde{\sigma}(\xi)\, f(\xi) \\
&= -\oint_{\mathcal{C}([a_-,a_+])} \frac{\mathrm{d}\xi}{2i\pi} \frac{\widetilde{\sigma}(\xi)\, f(\xi)}{\xi - x} \\
&= -\oint_{\mathcal{C}([a_-,a_+])} \frac{\mathrm{d}\xi}{2i\pi} \frac{\widetilde{\sigma}(\xi)}{\xi - x} \frac{1}{2y(\xi)} \left(-g(\xi) + \frac{(\mathcal{Q}f)(\xi)}{L(\xi)} \right) \\
&= \oint_{\mathcal{C}([a_-,a_+])} \frac{\mathrm{d}\xi}{2i\pi} \frac{\widetilde{\sigma}(\xi)}{\xi - x} \frac{g(\xi)}{2y(\xi)}
\end{aligned}$$

In the second line, we moved the contour from a neighborhood of x to a neighborhood of $[a_-, a_+]$, and used the fact that $\widetilde{\sigma}(\xi) \in O(\xi)$ and $f(\xi) \in O(1/\xi^2)$, so that the residue at ∞ of the integrand vanishes. In the fourth line, we use the fact that L is chosen such that $\frac{\widetilde{\sigma}(\xi)}{y(\xi)L(\xi)} = \frac{1}{S(\xi)}$, which is holomorphic in a neighborhood of $[a_-, a_+]$ thanks to Hyp. 51. Since, $\big(\mathcal{Q}f\big)$ is also holomorphic in a neighborhood of $[a_-, a_+]$, the contour integral of this term vanishes. For our purposes, it is not necessary to describe the vector space $\mathrm{Im}\,\mathcal{K}$. Notice that if we apply $\mathcal{K}^{-1}$ to a function $g \in \mathrm{Im}\,\mathcal{K}$ which is furthermore holomorphic outside $\mathbb{C} \setminus [\alpha_-, \alpha_+]$, we can contract the contour $\mathcal{C}([a_-, a_+])$ to a contour $\mathcal{C}([\alpha_-, \alpha_+])$.

Continuity of $\mathcal{K}$ and $\mathcal{K}^{-1}$

The key fact in this article is that $\mathcal{K}^{-1}$ is a continuous operator in $\big(\mathrm{Im}\,\mathcal{K}, \varphi \cdot \varphi_\Gamma\big)$:

Lemma 54. $\mathrm{Im}\,\mathcal{K}$ *is closed subspace of* $\mathcal{H}^{(1)}_{1;[a_-,a_+]}$ *for the topology induced by the norm* $\varphi \cdot \varphi_\Gamma$*, and there exists a constant* $k > 0$*, such that:*

$$\forall g \in \mathrm{Im}\,\mathcal{K}, \qquad \varphi \mathcal{K}^{-1} g \varphi_\Gamma \leq k \varphi g \varphi_\Gamma \tag{93}$$

We call $\varphi \mathcal{K}^{-1} \varphi_\Gamma$*, the infimum of such constants k.*

[1] In general, on the space of holomorphic functions with $g+1$ cuts, $\dim \mathrm{Ker}\,\mathcal{K} = g$, and one has to prescribe g cycle integrals in order to define an inverse operator.

Proof. Let us prove first that $\mathcal{K}$, as a endomorphism of $\mathcal{H}^{(1)}_{1;[a_-,a_+]}$, is continuous. For any $f \in \mathcal{H}^{(1)}_{1;[a_-,a_+]}$ in formula 90, if x runs along Γ, we can move the contour $\mathcal{C}([a_-,a_+]) \cup \mathcal{C}(x)$ to Γ_E and find the bound:

$$\begin{aligned}\forall f \in \varphi\mathcal{K}f\varphi_\Gamma \quad &\leq \quad 2\big(\max_{x\in\Gamma}|y(x)|\big)\varphi f\varphi_\Gamma + \frac{\ell(\Gamma_E)}{2\pi}\,\frac{\max_{\xi\in\Gamma_E}|L(\xi)|}{\min_{x\in\Gamma}|L(x)|}\Big(\frac{1}{d(\Gamma_E,\Gamma)}+c\Big)\max_{\xi\in\Gamma_E}|(V^{\{0\}})'(\xi)|\,\varphi f\varphi_{\Gamma_E} \\ &\leq \quad \left[2\max_{x\in\Gamma}|y(x)| + \frac{\ell(\Gamma_E)}{2\pi}\,\frac{\max_{\xi\in\Gamma_E}|L(\xi)|}{\min_{x\in\Gamma}|L(x)|}\Big(\frac{1}{d(\Gamma,\Gamma_E)}+c\Big)\max_{\xi\in\Gamma_E}|(V^{\{0\}})'(\xi)|\right]\varphi f\varphi_\Gamma\end{aligned}$$

We have used again the maximum principle for f to find the second line. Likewise, we can show that $\mathcal{K}^{-1}$: $\mathrm{Im}\,\mathcal{K} \to \mathcal{H}^{(1)}_{1;[a_-,a_+]}$ is continuous. In formula 93, we put x on Γ, and move the contour from $\mathcal{C}([a_-,a_+])$ to Γ_E in Eqn. 93. Doing so, we pick up a simple pole at $\xi = x$, and we find:

$$\big(\mathcal{K}^{-1}g\big)(x) = -\frac{g(x)}{2y(x)} + \frac{1}{\widetilde{\sigma}(x)}\oint_{\Gamma_E}\frac{\mathrm{d}\xi}{2i\pi}\,\frac{1}{\xi-x}\,\frac{L(\xi)\,g(\xi)}{2S(\xi)} \tag{94}$$

We find the bound:

$$\begin{aligned}\varphi\mathcal{K}^{-1}g\varphi_\Gamma \quad &\leq \quad \frac{\varphi g\varphi_\Gamma}{2\,\min_{x\in\Gamma}|y(x)|} + \frac{\ell(\Gamma_E)}{4\pi d(\Gamma,\Gamma_E)}\,\frac{\max_{\xi\in\Gamma_E}|L(\xi)|}{\min_{x\in\Gamma}|\widetilde{\sigma}(x)|}\,\frac{\varphi g\varphi_{\Gamma_E}}{\min_{\xi\in\Gamma_E}|S(\xi)|} \\ &\leq \quad \left(\frac{1}{2\,\min_{x\in\Gamma}|y(x)|} + \frac{\ell(\Gamma_E)}{4\pi d(\Gamma,\Gamma_E)}\,\frac{\max_{\xi\in\Gamma_E}|L(\xi)|}{\min_{x\in\Gamma}|\widetilde{\sigma}(x)|\,\min_{\xi\in\Gamma}|S(\xi)|}\right)\varphi g\varphi_\Gamma\end{aligned}$$

where we used the maximum principle in the last line. Eventually, let us show that $\mathrm{Im}\,\mathcal{K}$ is a closed subspace of $\mathcal{H}^{(1)}_{1;[a_-,a_+]}$. We pick up a sequence $(g_n)_n$ in $\mathrm{Im}\,\mathcal{K}$ converging towards $g \in \mathcal{H}^{(1)}_{1;[a_-,a_+]}$ for a norm $\varphi\cdot\varphi_{\Gamma_0}$ on a given contour Γ_0. Let $(f_n)_n$ be a sequence in $\mathcal{H}^{(1)}_{1;[a_-,a_+]}$ such that $g_n = \mathcal{K}f_n$, or equivalently $f_n = \mathcal{K}^{-1}g_n$. Using Eqn. 95 for any contour Γ, we know that $\varphi f_n\varphi_\Gamma \leq k\,\varphi g_n\varphi_\Gamma$ for some constant $k > 0$. So, f_n is a locally bounded subsequence of holomorphic functions in $\mathbb{C}\setminus[a_-,a_+]$. By Montel's theorem, it admits a subsequence $(f_{\varphi(n)})_n$ converging to some $f \in \mathcal{H}^{(1)}_{1;[a_-,a_+]}$ uniformly on any compact of $\mathbb{C}\setminus[a_-,a_+]$. Then using Eqn. 94, $g_{\varphi(n)} = \mathcal{K}f_{\varphi(n)} \to \mathcal{K}f$ for the norm $\varphi\cdot\varphi_{\Gamma_0}$. In particular, $g(x) = \mathcal{K}f(x)$ for all $x \in \mathrm{Ext}(\Gamma_0)$. Since g and f are both analytic in $\mathbb{C}\setminus[a_-,a_+]$, they must coincide on $\mathbb{C}\setminus[a_-,a_+]$. Hence, $g \in \mathrm{Im}\,\mathcal{K}$, showing that $\mathrm{Im}\,\mathcal{K}$ is closed. ⋄

$\varphi\mathcal{K}^{-1}\varphi_\Gamma$ is controlled by the distance of the zeroes s_i to the support $[a_-,a_+]$. This motivates Hyp. 51.

The endomorphism "negative part"

Let g be a holomorphic function at least in a neighborhood of $[a_-,a_+]$. The following endomorphism of $\mathcal{H}^{(1)}_{1;[a_-,a_+]}$ often appears in the loop equations:

$$\mathcal{N}_g(f)(x) = \oint_{\mathcal{C}([a_-,a_+])}\frac{\mathrm{d}\xi}{2i\pi}\,\frac{L(\xi)}{L(x)}\Big(\frac{1}{x-\xi}+c\Big)g(\xi)\,f(\xi) \tag{95}$$

We may write sometimes $\mathcal{N}_g[f(x)]$ as an abuse of notation. The analyticity assumption on g ensures that $\mathcal{N}_g$ is a continuous operator with respect to the norm $\varphi\cdot\varphi_\Gamma$. Indeed, let us put x on Γ and move the contour $\mathcal{C}([a_-,a_+])$ to Γ_E:

$$\mathcal{N}_g(f)(x) = g(x)f(x) + \oint_{\Gamma_E} \frac{\mathrm{d}\xi}{2i\pi}\,\frac{L(\xi)}{L(x)}\Big(\frac{1}{x-\xi}+c\Big)\,g(\xi)\,f(\xi) \tag{96}$$

Thus, the maximum principle implies:

$$\begin{aligned}
\varphi\mathcal{N}_g(f)\varphi_\Gamma \quad &\leq \quad \varphi g\varphi_\Gamma\varphi f\varphi_\Gamma + \frac{\ell(\Gamma_E)}{2\pi}\Big(\frac{1}{d(\Gamma_E,\Gamma)}+|c|\Big)\,\frac{\max_{\xi\in\Gamma_E}|L(\xi)|}{\min_{x\in\Gamma}|L(x)|}\,\varphi g\varphi_{\Gamma_E}\,\varphi f\varphi_{\Gamma_E} \\
&\leq \quad \Big[\varphi g\varphi_\Gamma + \frac{\ell(\Gamma_E)}{2\pi}\Big(\frac{1}{d(\Gamma_E,\Gamma)}+|c|\Big)\,\frac{\max_{\xi\in\Gamma_E}|L(\xi)|}{\min_{x\in\Gamma}|L(x)|}\,\varphi g\varphi_{\Gamma_E}\Big]\,\varphi f\varphi_\Gamma
\end{aligned}$$

3.6.3 Order of magnitude of W_n

If there exists a $1/N$ expansion, W_n ought to be of order of magnitude N^{2-n}. Let us start with a lemma explaining how this can be infered from rough bounds on W_n. Hereafter, $O_l(\cdots)$ or $o_l(\cdots)$ mean $O(\cdots)$ or $o(\cdots)$ with respect to the norm $\varphi\cdot\varphi_{\Gamma_l}$. Since the contours Γ_l are ordered from the interior to the exterior, being a $o_{l+1}(\cdots)$ is weaker than being a $o_l(\cdots)$. When the index l is not precised, it is understood that the bound holds for any integer l.

Lemma 55. *Let $\delta_{-1}W_1 := N^{-1}W_1 - W_1^{\{-1\}}$ and $l \geq 0$. Assume $\delta_{-1}W_1 \in o_l(1)$, and for all integer $n \geq 2$, assume $W_n \in O_l(N)$. Then:*

$$\forall n \geq 2 \qquad \|W_n\|_{\Gamma_{4n-6+l}} \in O(N^{2-n}) \tag{97}$$

Proof. Let $\delta_0 V = V - V^{\{0\}}$. Firstly, as $\delta_{-1}W_1$ and $(\delta_0 V)'$ goes to 0 uniformly on Γ_{-1} when $N\to\infty$, we observe that for any fixed integer k, and N large enough:

$$\begin{aligned}
(1-\varepsilon_{N,k+1})\varphi W_n\varphi_{\Gamma_{k+1}} \quad &\leq \quad \varphi\mathcal{K}^{-1}\varphi_{\Gamma_{k+1}}\,\Big\|\Big[\mathcal{K}+\delta\mathcal{K}+\frac{1}{N}\Big(1-\frac{2}{\beta}\Big)\frac{\mathrm{d}}{\mathrm{d}x}\Big]\,W_n\Big\|_{\Gamma_{k+1}} \\
&\qquad + \frac{1}{N}\Big|1-\frac{2}{\beta}\Big|\,\zeta_k\,\varphi W_n\varphi_{\Gamma_k}
\end{aligned}$$

where

$$\begin{aligned}
[\delta\mathcal{K}](f)(x) \quad &= \quad -\mathcal{N}_{(\delta_0 V)'}[f(x)] + 2\big(\delta_{-1}W_1\big)(x)\,f(x) \\
\varepsilon_{N,k+1} \quad &= \quad \varphi\mathcal{K}^{-1}\varphi_{\Gamma_{k+1}}\big(\|\mathcal{N}_{(\delta_0 V)'}\varphi_{\Gamma_{k+1}} + 2\|\delta_{-1}W_1\|_{\Gamma_{k+1}}\big)
\end{aligned}$$

goes to zero as N goes to infinity for $k+1 \geq l$ by assumption. ζ_k is defined in Eqn. 85. We assume hereafter that N is large enough so that $\varepsilon_{N,k+1}$ is smaller than $1/2$.

Secondly, the first version of the loop equation at rank $n \geq 2$ (Thm. 45) can be rewritten:

$$\Big[\mathcal{K}+\delta\mathcal{K}+\frac{1}{N}\Big(1-\frac{2}{\beta}\Big)\frac{\mathrm{d}}{\mathrm{d}x}\Big]W_n(x,x_I) = A_{n+1}+B_n+C_{n-1}+D_{n-1} \tag{98}$$

where:

$$\begin{aligned}
A_{n+1} &= -\frac{1}{N}\,W_{n+1}(x,x,x_I) \\
B_n &= -\frac{1}{N}\sum_{\substack{n_1,n_2\geq 1\\ n_1+n_2=n-1}}\ \sum_{\substack{J\subseteq I\\ |J|=n_1}} W_{n_1+1}(x,x_J)\,W_{n_2+1}(x,x_{I\setminus J}) \\
C_{n-1} &= -\frac{1}{N}\,\frac{2}{\beta}\sum_{i\in I}\frac{\mathrm{d}}{\mathrm{d}x_i}\left\{\frac{W_{n-1}(x,x_{I\setminus\{i\}})}{x-x_i} - \frac{L(x_i)}{L(x)}\Big(\frac{1}{x-x_i}+c\Big)W_{n-1}(x_I)\right\} \\
D_{n-1} &= \frac{1}{N}\,\frac{2}{\beta}\sum_{\tau\in\mathsf{Soft}}\frac{\partial_{a_\tau}W_{n-1}(x_I)}{x-a_\tau}
\end{aligned}$$

We know from Proposition 41 that $D_{n-1}\in O(e^{-N\eta})$, so this term does not contribute at any order of magnitude N^{-k}. Now, if we assume that $W_n\in O_l(N)$ for all $n\geq 2$ (this is obviously true for $n=1$), we always have $A_{n+1}\in O_l(1)$ and $C_{n-1}\in O_{l+2}(1)$, whereas the last last term in Eqn. 98 is bounded by hypothesis for $k\geq l$.

Now, we want to bound W_n by induction on n. At rank $n=2$, we have $B_2=0$, and we deduce from Eqn. 98 that $W_2\in O_{l+2}(1)$. Then at rank $n=3$, the product term B_3 is $O_{l+2}(1/N)$ and C_2 is $O_{l+4}(1/N)$, thus $W_3\in O_{l+4}(A_4)=O_6(1)$. Then similarly at rank $n=4$, the product term B_4 is $O_{l+4}(1/N)$ and C_3 is $O_{l+6}(1/N)$, thus $W_4\in O_{l+6}(1)$. This implies in return that $A_4\in O_{l+6}(1/N)$, thus $W_3\in O_{l+6}(1/N)$. And so on … The result can be proved by a triangular induction, as depicted in Fig. 3.2. At each vertical step, we are forced to trade the contour Γ_k with the exterior contour Γ_{k+2} in order to control the C terms. So, to go from $W_n\in O_{k_n}(N^{2-n})$ (in the n^{th} column) to $W_n\in O_{k_{n+1}}(N^{2-(n+1)})$ (in the $(n+1)^{\mathrm{th}}$ column), we must reach W_{n+2} in the n^{th} column. This is done by two vertical steps, thus $k_{n+1}=k_n+4$. Since $k_2=l+2$, we have $k_n=4n-6+l$ for all $n\geq 2$. ◇

Lemma 56. *If there exists* $\gamma\in[0,1[$ *and* $\delta\in[0,\infty[$ *such that* $W_n\in O_0(N^{\gamma n-\delta})$ *for all* $n\geq 2$*, then:*

$$\varphi W_n\varphi_{\Gamma_{4n-6+l}}\in O(N^{2-n}) \tag{99}$$

where $l=2\lceil(\gamma^{-1}-1)^{-1}\rceil$.

Proof. Now, let us rather assume the existence of $\gamma\in[0,1[$ and $\delta\geq 0$ such that, for all $n\geq 2$, $W_n\in O_0(N^{\gamma n-\delta})$. D_{n-1} being always exponentially small, it does not matter in our discussion. At rank $n=2$, as $B_2=0$ and $C_1\in O_2(1)$, we have $W_2\in O_2\big(\max[\frac{1}{N}W_3,1]\big)$. We also have for all n:

$$\begin{aligned}
A_{n+1} &\in O_0\big(\max[N^{\gamma n-\delta-(1-\gamma)},1]\big) \\
B_n &\in O_0\big(\max[N^{\gamma n-2\delta-(1-\gamma)},1]\big) \\
C_{n-1} &\in O_2\big(\max[N^{\gamma n-\delta-(1+\gamma)},1]\big)
\end{aligned}$$

When these $O(\cdots)$ decay, it does not hurt to consider them as $O(1)$. So, our bounds are upgraded at least to $W_n\in O_2\big(\max[N^{\gamma n-\delta'},1]\big)$ with $\delta'=\delta+1-\gamma>\delta$. By repeating

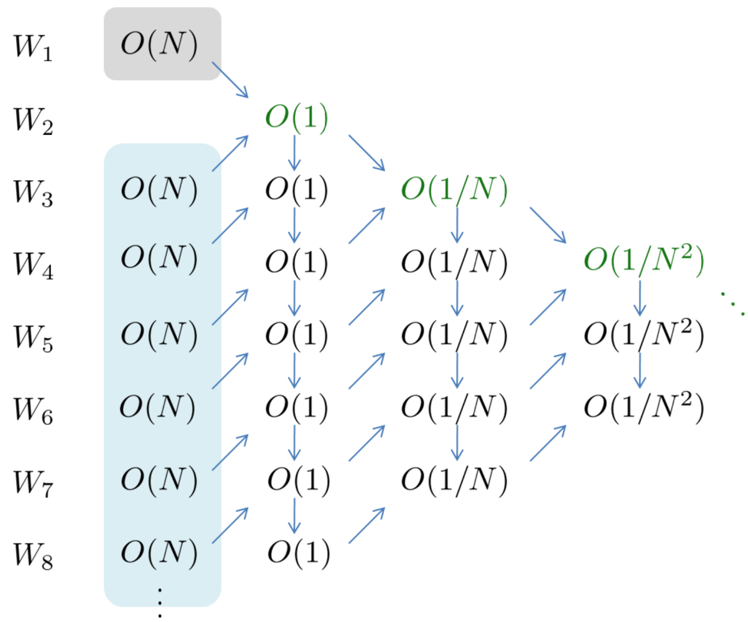


Figure 3.2: Triangular recursion for Lemma 55

the argument k times, we obtain for all $n \geq 2$, $W_n \in O_{2k}\big(\max[N^{((k+1)\gamma-k)n-\delta}, 1]\big)$. Since $\gamma < 1$, by choosing an integer $k \geq \frac{1}{1/\gamma - 1}$, we deduce that $W_n \in O_l(1)$ for all $n \geq 2$ with $l = 2k$, and we apply Lemma 55 to conclude. $\diamond$

3.6.4 Initialization

We now establish a priori control on the correlators. We shall use:

Lemma 57. *Let $w_N = N^{\varepsilon}$ for some $\varepsilon > 0$. Assume that for any integer p, there exists $C_p > 0$ and independent of N, such that for all $x \in \mathbb{C} \setminus [a_-, a_+]$:*

$$P_{N,\beta}^{V}\Big\{\Big|\int \frac{\mathrm{d}M_N(\xi)}{x-\xi} - P_{N,\beta}^{V}\Big[\int \frac{\mathrm{d}M_N(\xi)}{x-\xi}\Big]\Big|^p\Big\} \leq \frac{C_p w_N^p}{\big(d(x,[a_-,a_+])\big)^{2p}} \tag{100}$$

Then, for all $n \geq 2$, $W_n \in O(w_N^n)$ for the norm $\varphi \cdot \varphi_\Gamma$, when $N \to \infty$.

Proof. For $n \geq 2$, $W_n(x_1, \ldots, x_n)$ is a polynomial in:

$$P_{N,\beta}^{V;[a_-,a_+]}\Big\{\prod_{j \in J}\Big(\int \frac{\mathrm{d}M_N(\xi)}{x_j-\xi} - P_{N,\beta}^{V}\Big[\int \frac{\mathrm{d}M_N(\xi)}{x_j - \xi}\Big]\Big)\Big\} \tag{101}$$

with $J \subseteq \{1, \ldots, n\}$, and the coefficients of this polynomial are independent of N. Thus by Eqn. 100 and Hölder inequality, there exists $D_n \in \mathbb{R}_+^*$ independent of N such that:

$$\big|W_n(x_1,\ldots,x_n)\big| \leq \frac{D_n}{\big(\min_{1\leq i\leq n} d(x_i,[a_-,a_+])\big)^{2n}} \tag{102}$$

Hence, taking the sup for $x_i \in \Gamma$, we find $W_n \in O(w_N^n)$.

Lemma 58. *Under the five assumptions of Section 3.6.1, Eqn. 100 holds for any* $\varepsilon > 0$.

Proof. Our starting point comes from a result of Boutet de Monvel, Pastur and Shcherbina [43], developed by Johansson[2][73, (3.49)] and more recently in [75, (2.26)]. Let $\Gamma' \subseteq \mathrm{Int}\,\Gamma$ be a contour surronding $[a_-,a_+]$. For any $\varphi : \mathrm{Int}(\Gamma) \to \mathbb{C}$ which is a continuous function, and real-valued on $[a_-,a_+]$, there exists a positive constant C such that:

$$P_{N,\beta}^{V}\left[\exp\left(\frac{1}{2\big(\sup_{z'\in\Gamma'}|\varphi(z')|\big)w_N}\Big(\int\varphi(\xi)\mathrm{d}M_N(\xi) - N\int\varphi(\xi)\mathrm{d}L(\xi)\Big)\right)\right] \leq 3 \tag{103}$$

where $w_N = C\ln N$. By Chebychev's inequality, we deduce that:

$$\forall t\in[0,+\infty[,\quad P_{N,\beta}^{V}\left\{\Big|\int\varphi(\xi)\mathrm{d}M_N(\xi) - N\int\varphi(\xi)\mathrm{d}L(\xi)\Big| \geq t\big(\sup_{z'\in\Gamma'}|\varphi(z')|\big)w_N\right\} \leq 6e^{-t} \tag{104}$$

and therefore:

$$\forall p\in\mathbb{N},\quad P_{N,\beta}^{V}\left[\Big|\int\varphi(\xi)\mathrm{d}M_N(\xi) - N\int\varphi(\xi)\mathrm{d}L(\xi)\Big|^p\right] \leq p!\big(\sup_{z'\in\Gamma'}|\varphi(z')|\big)^p w_N^p \tag{105}$$

In particular, we can apply this discussion to $\varphi(z) = \mathrm{Re}\,\frac{1}{x-z}$ and $\varphi(z) = \mathrm{Im}\,\frac{1}{x-z}$ where x is a point of Γ. This leads to Eqn. 100. $\diamond$

In the case of a strictly convex potential, we may use instead concentration of measure:

Lemma 59. *If $V^{\{0\}}$ is strictly convex on $[a_-,a_+]$, then Eqn. 100 holds with* $\varepsilon = 0$.

Proof. Since $V^{\{0\}}$ is strictly convex on $[a_-,a_+]$, V is also strictly convex on $[a_-,a_+]$ for N large enough. By concentration of measure, see [66] or [9, Section 2.3 and Exercise 4.4.33], there exists $c > 0$ such that, for all $x \in \mathbb{C}\backslash[a_-,a_+]$, for all $\varepsilon > 0$ and $N \in \mathbb{N}$:

$$P_{N,\beta}^{V}\left\{\Big|\int\frac{\mathrm{d}M_N(\xi)}{x-\xi} - P_{N,\beta}^{V}\Big[\int\frac{\mathrm{d}M_N(\xi)}{x-\xi}\Big]\Big| \geq \frac{\varepsilon}{\big(d(x,[a_-,a_+])\big)^2}\right\} \leq 2e^{-c\varepsilon^2} \tag{106}$$

This entails Eqn. 100. $\diamond$

[2] Johansson's has written his proof in the framework $[a_-,a_+] = \mathbb{R}$, but there is no difficulty adapting it to $[a_-,a_+]$ finite. $\mathrm{Im}\,z$ should be replaced by $d(z,[a_-,a_+])$, and its powers in the bound of his Lemma 3.10 and 3.11 may differ, but the order of magnitude ω_N (our w_N) is the same.

3.6.5 Leading order of W_1

Afterwards, all steps only rely on the analysis of loop equations. Although we already know the characterization of the equilibrium measure μ_{eq}, and thus of its Stieltjes transform $W_1^{\{-1\}}$, let us recall how $W_1^{\{-1\}}$ is characterized by the loop equations. We write the loop equation at rank 1 (Thm. 44):

$$
\begin{aligned}
& \frac{1}{N^2}\,W_2(x,x) \\
& +\big(W_1^{\{-1\}}(x)\big)^2-\frac{1}{(x-a_-)(x-a_+)} \\
& -\oint_{\mathcal{C}([a_-,a_+])}\frac{\mathrm{d}\xi}{2i\pi}\,\frac{1}{x-\xi}\,\frac{(\xi-a_-)(\xi-a_+)}{(x-a_-)(x-a_+)}\,(V^{\{0\}})'(\xi)\,W_1^{\{-1\}}(\xi) \\
& +\mathcal{K}\big[\delta_{-1}W_1\big](x)+\frac{1}{N}\Big(1-\frac{2}{\beta}\Big)\Big(W_1^{\{-1\}}(x)+\frac{1}{(x-a_-)(x-a_+)}\Big)-\frac{1}{N}\,\mathcal{N}_{V^{\{1\}}}\big[W_1^{\{-1\}}\big](x) \\
& +\big((\delta_{-1}W_1)(x)\big)^2-\mathcal{N}_{(\delta_0 V)'}\big[\delta_{-1}W_1\big](x)+\frac{1}{N}\,\mathcal{N}_{(\delta_1 V)'}\big[W_1^{\{-1\}}\big](x) \quad = \quad 0
\end{aligned}
$$

We already know that the 4th and the 5th line are $o(1)$. Since $W_2\in o(N^2)$, $W_1^{\{-1\}}$ satisfy the loop equation at leading order:

$$
\begin{aligned}
\big(W_1^{\{-1\}}(x)\big)^2 \quad = \quad & \frac{1}{(x-a_-)(x-a_+)} \\
& +\oint_{\mathcal{C}([a_-,a_+])}\frac{\mathrm{d}\xi}{2i\pi}\,\frac{1}{x-\xi}\,\frac{(\xi-a_-)(\xi-a_+)}{(x-a_-)(x-a_+)}\,(V^{\{0\}})'(\xi)\,W_1^{\{-1\}}(\xi)
\end{aligned}
$$

Remark 60. Recall that $\mathrm{supp}\,\rho=[\alpha_-,\alpha_+]$ is the discontinuity locus of $W_1^{\{-1\}}$. By the properties of the Stieltjes transform:

$$
y(x)=\frac{(V^{\{0\}})'(x)}{2}-W_1^{\{-1\}}(x) \tag{107}
$$

defines a holomorphic function on $U\setminus[\alpha_-,\alpha_+]$, and:

$$
\forall x_0\in[\alpha_-,\alpha_+],\qquad \lim_{\varepsilon\to 0^+}y(x_0+i\varepsilon)=i\pi\,\rho(x_0) \tag{108}
$$

We state that there exists $M(x)$, continuous in some open neighborhood of $[\alpha_-,\alpha_+]$, such that:

$$
y(x)=\frac{M(x)}{\sqrt{(x-\alpha_-)(x-\alpha_+)}} \tag{109}
$$

Proof. In Eqn. 107, we may first deform the contour $\mathcal{C}([a_-,a_+])$ to $\mathcal{C}([\alpha_-,\alpha_+])$. Secondly, we can rewrite:

$$
\begin{aligned}
& \big(W_1^{\{-1\}}(x)\big)^2-(V^{\{0\}})'(x)W_1^{\{-1\}}(x)+\frac{U(x)}{(x-a_-)(x-a_+)}=0 \\
& U(x)=-1+\oint_{\mathcal{C}([\alpha_-,\alpha_+])\cup\mathcal{C}(x)}\frac{\mathrm{d}\xi}{2i\pi}\,\frac{(\xi-a_-)(\xi-a_+)}{x-\xi}\,(V^{\{0\}})'(\xi)\,W_1^{\{-1\}}(\xi)
\end{aligned}
$$

where $U(x)$ is now holomorphic in some open neighborhood of $[\alpha_-,\alpha_+]$. So:

$$y(x)=\sqrt{\frac{R(x)}{(x-a_-)(x-a_+)}},\quad R(x)=\frac{1}{4}(x-a_-)(x-a_+)(V^{\{0\}})'(x)-U(x) \quad (110)$$

This equation tells us that the discontinuity of y is of squareroot type. If $\alpha_-=a_-$ and $\alpha_+=a_+$, we have Eqn. 109. If say $a_-<\alpha_-$, the fact that $y(x)$ has no discontinuity on $[a_-,\alpha_-[$ but a discontinuity on $[\alpha_-,\alpha_+]$ forces $R(x)$ to have a simple zero at $x=a_-$ and at $x=\alpha_-$, so that $y(x)$ is finite when $x=a_-$ and vanishes as $O(\sqrt{x-\alpha_-})$ when $x\to\alpha_-$. A similar statement holds if $a_+>\alpha_+$. Then, Eqn. 109 holds a fortiori. ◇

3.6.6 First correction to W_1

Let us reconsider Eqn. 107 (or the equivalent relation taking Remark 3.5.3 into account) after removing the 2nd and the 3rd line which has just been identified as the leading order. We can write as in Section3.6.3:

$$\left[\mathcal{K}+\widetilde{\delta\mathcal{K}}+\frac{1}{N}\Big(1-\frac{2}{\beta}\Big)\frac{\mathrm{d}}{\mathrm{d}x}\right]\delta_{-1}W_1(x)=A_2+C_0+D_0 \quad (111)$$

where:

$$\begin{aligned}
\widetilde{\delta\mathcal{K}}[f](x) &= -\mathcal{N}_{(\delta_0 V)'}[f(x)]+\delta_{-1}W_1(x)f(x)\\
A_2 &= -\frac{1}{N^2}W_2(x,x)\\
C_0 &= -\frac{1}{N}\Big(1-\frac{2}{\beta}\Big)\left(\sum_{\tau\in\mathsf{Hard}}\frac{1}{a_\tau-a_{-\tau}}\,\frac{1}{x-a_\tau}\right)\\
D_0 &= \sum_{\tau\in\mathsf{Soft}}\frac{\partial_{a_\tau}\ln Z}{x-a_\tau}
\end{aligned}$$

By an argument similar to Eqn. 98, knowing that $W_2\in O_l(N)$ implies that $\delta_{-1}W_1\in O_{l+1}(1/N)$. Assuming further $W_3\in O_{l'}(N)$ implies after Section 3.6.3 that $W_2\in O_{l'+2}(1)$, so the 1st line of Eqn. 107 is subleading compared to the 3rd line. These two bounds are provided by Section 3.6.4 (the values of l and l' do not matter here). Hence:

Lemma 61. *There exists $W_1^{\{0\}}\in\mathcal{H}^{(1)}_{1;[\alpha_-,\alpha_+]}$ such that $W_1=NW_1^{\{-1\}}+W_1^{\{0\}}+o(1)$. Explicitly:*

$$W_1^{\{0\}}(x)=\mathcal{K}^{-1}\left\{-\Big(1-\frac{2}{\beta}\Big)\left[\frac{\mathrm{d}}{\mathrm{d}x}\big(W_1^{\{-1\}}(x)\big)+\sum_{\tau\in\mathsf{Hard}}\frac{1}{a_\tau-a_{-\tau}}\,\frac{1}{x-a_\tau}\right]+\mathcal{N}_{(V^{\{1\}})'}(W_1^{\{-1\}})(x)\right\} \quad (112)$$

This order was also obtained by [75] with similar arguments.

3.6.7 Recursion hypothesis at order k_0

Let $k_0 \geq -1$. We assume that the correlators W_n (for all $n \geq 1$) are determined up to a $o(N^{-k_0})$ for the norm $\varphi \cdot \varphi_{\Gamma_{l(k_0;n)}}$.

$$W_n(x_1,\ldots,x_n) = \sum_{k=n-2}^{k_0} N^{-k}\, W_n^{\{k\}}(x_1,\ldots,x_n) + N^{-k_0}\, \delta_{k_0} W_n(x_1,\ldots,x_n) \qquad (113)$$

Here, $W_n^{\{k\}}(x_1,\ldots,x_n)$ are already known (they depend on β but not on N), and we call:

$$\omega_n^{\{k\}} = \sup_{\substack{1\leq n'\leq n \\ -1\leq k'\leq k}} \varphi W_n^{\{k\}} \varphi_{\Gamma_{l(k_0;n)}} \qquad (114)$$

a bound for their norm. We can always assume that $l(k,n)$ defined for $-1 \leq k \leq k_0$ and $n \geq 1$ is an increasing function of k and n. Though the errors $\delta_{k_0} W_n$ are not supposed to be known, we assume that they are small:

$$\forall n \geq 1 \qquad \varphi \delta_{k_0} W_n \varphi_{\Gamma_{l(k_0;n)}} \leq \varepsilon_N^{\{k_0\}}\, \Delta_n^{\{k_0\}} \qquad (115)$$

Here, $\varepsilon_N^{\{k_0\}}$ depends only on N and k_0, and $\varepsilon_N^{\{k_0\}} \to 0$ when $N \to \infty$, and $\Delta_n^{\{k_0\}}$ is a constant independent of N. We may assume that $\Delta_n^{\{k_0\}}$ increases with $n \geq 1$, upon replacement by $\sup_{1\leq n'\leq n} \Delta_n^{\{k_0\}}$. When $n > k_0 + 2$, we assume that Eqn. 113 reduces to:

$$W_n = N^{-k_0}\, \delta_{k_0} W_n \qquad (116)$$

Lemma 57 and Section 61 ensure that the initial ($k_0 = -1$) recursion hypothesis is satisfied. Moreover, we can take $\varepsilon_N^{\{-1\}} = 1/N$, and up to a redefinition $\Gamma_k \to \Gamma_{k-m}$ for some integer m, we can take $l(-1;n) = 4(n-1)$.

3.6.8 Determination of $\delta_{k_0} W_{n_0}$

Let $n_0 \geq 1$. We now turn to the determination of the leading order of $\delta_{k_0} W_{n_0}(x,x_I)$. The case $(n_0,k_0) = (1,-1)$ is a bit special (because of the second term of the second line in Eqn. 107) and is given by Lemma 61. In all other cases, we consider the loop equation at rank n_0 (Thm. 45). Up to $o(N^{-(k_0-1)})$, the equation is true and involves quantities which are already known from the recursion hypothesis. The equality of the $o(N^{-(k_0-1)})$ involves the unknown $\delta_{k_0} W_{n_0}(x,x_I)$. The operator $\mathcal{K}$ introduced in Section 3.6.2 plays a special role. When the potential V has a $1/N$ expansion, the operator $\mathcal{N}$ introduced in Section 3.6.2 also appears, and we denote:

$$V = \sum_{k=0}^{k_0+1} N^{-k}\, V^{\{k\}} + N^{-(k_0+1)}\, \delta_{k_0+1} V$$

We find:

$$N^{-(k_0-1)}\, \mathcal{K}\big(\delta_{k_0} W_{n_0}\big)(x,x_I) = -N^{-k_0}\, E_{n_0}^{\{k_0\}}(x,x_I) - N^{-k_0}\, R_{n_0}^{\{k_0\}}(x,x_I) \qquad (117)$$

with

$$
\begin{aligned}
E_{n_0}^{\{k_0\}}(x,x_I) \quad := \quad & W_{n_0+1}^{\{k_0\}}(x,x,x_I) - \sum_{k=1}^{k_0+1} \mathcal{N}_{(V^{\{k\}})'}\left[W_{n_0}^{\{k_0+1-k\}}(x,x_I)\right] \\
& + \sum_{J\subseteq I}\sum_{k=0}^{k_0} W_{|J|+1}^{\{k\}}(x,x_J)\, W_{n_0-|J|}^{\{k_0-k\}}(x,x_{I\setminus J}) + \Big(1-\frac{2}{\beta}\Big)\frac{\mathrm{d}}{\mathrm{d}x}\Big(W_{n_0}^{\{k_0\}}(x,x_I)\Big) \\
& + \frac{2}{\beta}\sum_{i\in I}\frac{\mathrm{d}}{\mathrm{d}x_i}\left\{\frac{W_{n_0-1}^{\{k_0\}}(x,x_{I\setminus\{i\}})}{x-x_i} - \frac{L(x_i)}{L(x)}\Big(\frac{1}{x-x_i}+c\Big)W_{n_0-1}^{\{k_0\}}(x_I)\right\}
\end{aligned}
$$

and the remaining

$$
\begin{aligned}
R_{n_0}^{\{k_0\}}(x,x_I) \quad := \quad & \delta_{k_0} W_{n_0+1}(x,x,x_I) + \sum_{k=1}^{k_0} N^{-k} \sum_{k'=0}^{k_0}\sum_{J\subseteq I} W_{|J|+1}^{\{k'\}}(x,x_J)\, W_{n_0-|J|}^{\{k_0+k-k'\}}(x,x_{I\setminus J}) \\
& + \sum_{k=0}^{k_0} N^{-k} \sum_{J\subseteq I}\big(\delta_{k_0} W_{|J|+1}\big)(x,x_J)\, W_{n_0-|J|}^{\{k\}}(x,x_{I\setminus J}) \\
& + N^{-k_0}\sum_{J\subseteq I}\big(\delta_{k_0} W_{|J|+1}\big)(x,x_J)\,\big(\delta_{k_0} W_{n_0-|J|}\big)(x,x_{I\setminus J}) \\
& + \Big(1-\frac{2}{\beta}\Big)\frac{\mathrm{d}}{\mathrm{d}x}\Big(\big(\delta_{k_0} W_{n_0}\big)(x,x_I)\Big) - \big(\delta_{k_0} W_{n_0+1}\big)(x,x,x_I) \\
& - \sum_{k=0}^{k_0} N^{-k}\, \mathcal{N}_{\left(\delta_{k_0+1}V\right)'}\left[W_{n_0}^{\{k\}}(x,x_I)\right] - \sum_{k=0}^{k_0} N^{-k}\, \mathcal{N}_{(V^{\{k+1\}})'}\left[\big(\delta_{k_0} W_{n_0}\big)(x,x_I)\right] \\
& - N^{-k_0}\, \mathcal{N}_{\left(\delta_{k_0+1}V\right)'}\left[\big(\delta_{k_0} W_{n_0}\big)(x,x_I)\right] \\
& + \frac{2}{\beta}\sum_{i\in I}\frac{\mathrm{d}}{\mathrm{d}x_i}\left\{\frac{\big(\delta_{k_0} W_{n_0-1}\big)(x,x_{I\setminus\{i\}})}{x-x_i} - \frac{L(x_i)}{L(x)}\Big(\frac{1}{x-x_i}+c\Big)\big(\delta_{k_0} W_{n_0-1}\big)(x_I)\right\} \\
& + \frac{2}{\beta}\sum_{\tau\in\mathsf{Soft}} N^{k_0}\,\frac{\partial_{a_\tau} W_{n-1}(x_I)}{x-a_\tau}
\end{aligned}
$$

It is understood that $\mathcal{K}$ and $\mathcal{N}_g$ operate on the x variable. The variables x_I are spectators. Notice that this equation is linear in $\delta_{k_0} W_{n_0}$, up to a small quadratic term.

Looking naively at this equation, we see that the leading term of $\delta_{k_0} W_{n_0}$ happens to be of order $1/N$ (giving a $N^{-(k_0+1)}$ contribution to W_{n_0}), and is obtained by applying $\mathcal{K}^{-1}$ to $E_{n_0}^{\{k_0\}}(x,x_I)$. To make this idea rigorous, let us bound $R_{n_0}^{\{k_0\}}$. Even if some terms in the right hand side have not been determined yet (like $\delta_{k_0} W_n$ that we are just considering), we already know a bound for each of them from the recursion hypothesis. Very rough bounds are enough, we just need to show that the right hand side is small when $N \to \infty$. When $k_0 = -1$, we must pay special attention at the terms involving N^{-k_0} directly, i.e. the 3rd line and the 6th line in Eqn. 118. In the 6th line, $(\delta_{k_0+1} V)'$ is of order N^{-1}, so we obtain a term of order $\varepsilon_N^{\{k_0\}}$, which is always small. The 3rd line is of order $N(\varepsilon_N^{\{-1\}})^2$, which is also small since we have here $\varepsilon_N^{\{-1\}} = 1/N$ (Lemma 61). For N large enough, we have:

$$\begin{aligned}
\varphi R_{n_0}^{\{k_0\}} \varphi_{\Gamma_{l(k_0+1;n_0)}} \quad \leq \quad & \varepsilon_N^{\{k_0\}}\,\Delta_{n_0+1}^{\{k_0\}} + N^{-1}\,(k_0+1)2^{n_0-1}\,\big(\omega_{n_0}^{\{k_0\}}\big)^2 \\
& + \varepsilon_N^{\{k_0\}}\,2^{n_0-1}\,\Delta_{n_0}^{\{k_0\}}\,\omega_{n_0}^{\{k_0\}} + (\varepsilon_N^{\{k_0\}})^2\,N^{-k_0}\,\big(\Delta_{n_0}^{\{k_0\}}\big)^2 \\
& + \varepsilon_N^{\{k_0\}}\left|1 - \frac{2}{\beta}\right|\zeta_{l(k_0;n_0)}\,\Delta_{n_0}^{\{k_0\}} + N^{-1}\sum_{k=0}^{k_0}\varphi\mathcal{N}_{\left(\delta_{k_0+1}V\right)'}\varphi_{\Gamma_{l(k_0;n_0)}}\,\Delta_{n_0}^{\{k_0\}} \\
& + \varepsilon_N^{\{k_0\}}\sum_{k=0}^{k_0}\varphi\mathcal{N}_{(V^{\{k\}})'}\varphi_{\Gamma_{l(k_0;n_0)}}\,\Delta_{n_0}^{\{k_0\}} + \varepsilon_N^{\{k_0\}}\,N^{-k_0}\,\varphi\mathcal{N}_{\left(\delta_{k_0+1}V\right)'}\varphi_{\Gamma_{l(k_0;n_0)}}\,\Delta_{n_0}^{\{k_0\}} \\
& + \varepsilon_N^{\{k_0\}}\,\frac{2}{\beta}\,\zeta_{l(k_0;n_0-1)}\left(|c| + \zeta_{l(k_0;n_0-1)+1}\,\frac{\sup_{\xi\in\Gamma_{l(k_0;n_0-1)}}|L(\xi)|}{\inf_{x\in\Gamma_{l(k_0;n_0-1)}}|L(x)|}\right)\Delta_{n_0-1}^{\{k_0\}} \\
& + \frac{2}{\beta}\,\frac{\gamma_n\,\#\mathsf{Soft}}{d(\Gamma,[a_-,b_-])^n}N^{k_0+n-1}\,e^{-N\eta}
\end{aligned}$$

Given the control provided by the recursion hypothesis, this inequality is correct provided we choose:

$$l(k_0+1;n_0) \geq \max\big[l(k_0;n_0-1)+2,\, l(k_0;n_0)+1,\, l(k_0;n_0+1)\big] \tag{118}$$

Accordingly, $R_{n_0}^{\{k_0\}} \to 0$ when $N \to \infty$. Eqn. 117 tells us that $E_{n_0}^{\{k_0\}} + R_{n_0}^{\{k_0\}} \in \mathrm{Im}\,\mathcal{K}$ for any N. Since $\mathrm{Im}\,\mathcal{K}$ is closed (Lemma 54), we know that $E_{n_0}^{\{k_0\}} \in \mathrm{Im}\,\mathcal{K}$, and also by difference $R_{n_0}^{\{k_0\}} \in \mathrm{Im}\,\mathcal{K}$ for any N. And, by continuity of $\mathcal{K}^{-1}$, we deduce:

$$\delta_{k_0} W_{n_0} = \frac{1}{N}\,W_{n_0}^{\{k_0+1\}} + \frac{1}{N}\,\delta_{k_0+1} W_{n_0} \tag{119}$$

where:

$$W_{n_0}^{\{k_0+1\}} = -\mathcal{K}^{-1}[E_{n_0}^{\{k_0\}}], \qquad \delta_{k_0+1} W_{n_0} = -\mathcal{K}^{-1}[R_{n_0}^{\{k_0\}}] \in o(1) \tag{120}$$

The previous inequality is more precise about the $o_{l(k_0+1;n_0)}(1)$: there exists a constant $\Delta_{n_0}^{\{k_0+1\}}$, such that

$$\varphi\delta_{k_0+1} W_{n_0} \varphi_{\Gamma_{l(k_0+1;n_0)}} \leq \Delta_{n_0}^{\{k_0+1\}} \max(N^{-1}\,;\,\varepsilon_N^{\{k_0\}}) \tag{121}$$

3.6.9 Remarks

The recursion hypothesis tells us that $W_{n_0}^{\{k_0\}} = 0$ whenever $n > k_0 + 2$ (we call $\star[k_0]$ this recursive assumption). Let us see what happens at order $k_0 + 1$ (here, k_0 is fixed, but n_0 is free), by looking at Eqn. 120.

- The term $W_{n_0+1}^{\{k_0\}}$ vanishes whenever $n_0 > k_0 + 1$.

- The term $W^{\{k\}}_{|J|+1}W^{\{k_0-k\}}_{n_0-|J|}$ may be non zero in case $k+1 \geq |J| \geq n_0-k_0-2+k$. This is impossible to fulfil as soon as $n_0 > k_0+3$.
- The term $\left(W^{\{k_0\}}_{n_0}\right)'$ vanishes whenever $n_0 > k_0+2$.
- The term involving $W^{\{k_0\}}_{n_0-1}$ vanishes whenever $n_0 > k_0+3$.

Accordingly, $W^{\{k_0+1\}}_{n_0} \equiv 0$ when $n_0 > (k_0+1)+2$, i.e. $\star[k_0+1]$ holds. This is just the manifestation of Lemma 55. Hence, we have propagated the full recursion hypothesis to order k_0+1. An easy recursion shows that $W^{\{k\}}_n$ are actually holomorphic functions on the domain $\mathbb{C}\setminus[\alpha_-,\alpha_+]$, i.e belongs to the subspace $\mathcal{H}^{(1)}_{n;[\alpha_-,\alpha_+]}$ of $\mathcal{H}^{(1)}_{n;[a_-,a_+]}$. Therefore, we can contract the contour to $\mathcal{C}([\alpha_-,\alpha_+])$ in the expression of $\mathcal{K}^{-1}$ (Eqn. 93) when computing $W^{\{k\}}_n$ with formula 120.

Since $l(k;n) = 4(n-1)$, the minimal solution of Eqn. 118 is $l(k;n) = 4(n+k)$. Indeed, in this proof, we need to have a more restrictive control on the error done at height $n+k$, in order to bound the error done at height $n+k+1$. Nevertheless, since $\Gamma_l \subseteq \mathrm{Int}(\Gamma_E)$ for all l, we can at the end make the weaker statement that, for any n and k:

$$\varphi\delta_k W_n \varphi_{\Gamma_E} \to 0 \tag{122}$$

when $N\to\infty$. However, we necessarily have $d(\Gamma_l,\Gamma_{l+1}) \to 0$ when $l\to\infty$, so that the constant ζ_l which allows us to bound the derivative of a function with the function itself (Eqn. 85), blows up. This means that Eqn. 122 cannot be uniform[3] in n and k, even when $\beta = 2$.

A posteriori, from Eqn. 122, we can deduce by choosing rather $l(k_0;n_0) = 8(n_0+k_0)$:

$$\varphi\delta_{k_0}W_{n_0}\varphi \leq N^{-1}\,\varphi W^{\{k_0+1\}}_{n_0}\varphi_{\Gamma_{l(k_0+1;n_0)}} + \Delta^{\{k_0\}}_{n_0}\max(N^{-1};\varepsilon^{\{k_0\}}_N) \tag{123}$$

Subsequently, upon redefinition of the constant $\Delta^{\{k_0\}}_{n_0}$, we may choose $\varepsilon^{\{k_0\}}_N = 1/N$. Finally, we can make the weaker statement that, for any n and k:

$$\varphi\delta_k W_n \varphi_{\Gamma_E} \in o(1/N) \tag{124}$$

without uniformity in n and k.

3.7 Proof of the main results

3.7.1 Expansion of the correlators

We wish to study the β ensembles on a given interval $[b_-,b_+]$, with the hypotheses 37 on the potential V. When both edges are hard, Hyp. 37 are equivalent to the five assumptions of Section 3.6, so the Proposition 38 is already proved, as we have shown recursively that Eqn. 113 holds for all k_0. Let us now assume that one of the edge is

[3] We thank Pavel Bleher for pointing out a mistake in a former version of the article, which we corrected by introducing this family of nested contours.

soft. The equilibrium measure $\mu_{\mathrm{eq}} := \mu_{\mathrm{eq}}^{V;[b_-,b_+]}$ with support $[\alpha_-,\alpha_+] \subset [b_-,b_+]$ also coincides with $\mu_{\mathrm{eq}}^{V;[a_-,a_+]}$, where a_- can be any point in $[b_-,\alpha_-[$ if b_- is a soft edge, and $a_- = b_-$ else (resp. a_+ can be any point in $]\alpha_+,b_+]$ if b_+ is a soft edge, and $a_+ = b_+$ else). When b_τ is a soft edge, "offcriticality" implies that $S(x)$ is positive in a neighborhood of α_τ in $[b_-,b_+]\setminus]\alpha_-,\alpha_+[$. So, one can choose an interval $[a_-,a_+] \subseteq U$, and such that the five assumptions of Section 3.6 are satisfied for $\mathrm{d}P_{N,\beta}^{V;[a_-,a_+]}$. Theorem 53 then can be applied: there exists an asymptotic expansion

$$W_n^{V;[a_-,a_+]}(x_1,\dots,x_n) = \sum_{k\geq n-2} N^{-k}\, W_n^{V;\{k\}}(x_1,\dots,x_n) \tag{125}$$

with respect to the norm $\varphi\cdot\varphi_{\Gamma_E}$ where $\Gamma_E \subseteq U$ can be any contour surrounding $[a_-,a_+]$ but not the zeroes of S. The "large deviation control" on $[b_-,b_+]$ allows to use Proposition 40: there exists $\eta > 0$ such that, for any contour $\Gamma'_E \subseteq \mathbb{C}$ surrounding $[b_-,b_+]$, there exists $T_{n,\Gamma} > 0$ such that:

$$\varphi W_n^{V;[b_-,b_+]} - W_n^{V;[a_-,a_+]}\varphi_{\Gamma'_E} \leq T_{n,\Gamma'_E}\, e^{-N\eta} \tag{126}$$

This implies that the right hand side of Eqn. 125 is an asymptotic series for $W_n^{V;[b_-,b_+]}(x_1,\dots,x_n)$, uniformly for $(x_1,\dots,x_n)$ in any compact of $(\mathbb{C}\setminus[b_-,b_+])^n$.

We give below a more transparent condition, which imply the "large deviation control" assumption on $[b_-,b_+]$:

Remark 62. If $S(x) > 0$ whenever $x \in [b_-,b_+]$, then $\mathcal{J}^{V;[b_-,b_+]}$ achieves its minimum value only on $[\alpha_-,\alpha_+]$,

Indeed, $\mathcal{J}^{V;[b_-,b_+]}(x)$ is differentiable when $x \in]b_-,b_+[\setminus[\alpha_-,\alpha_+]$, and we have:

$$\big(\mathcal{J}^{V;[b_-,b_+]}(x)\big)' = \frac{(V^{\{0\}})'(x)}{2} - W_1^{\{-1\}}(x) = y(x) = S(x)\sigma(x) \tag{127}$$

The sign of the square root $\sigma(x)$ is determined for example by the positivity conditions on $\mathcal{J}^{V;[b_-,b_+]}$. If we assume that S do not vanish on $[b_-,b_+]$, this implies that $\mathcal{J}^{V;[b_-,b_+]}$ is strictly decreasing in $[b_-,\alpha_-[$ and strictly increasing on $]\alpha_+,b_+]$, hence the remark.

3.7.2 Expansion of the free energy

So far, we only have determined the expansion of the correlators which are by definition derivatives of the free energy. To find the free energy itself, one would like to interpolate between our initial potential V, and a simpler situation, using that the difference depends on the correlators. For any fixed $\alpha_- < \alpha_+$, and fixed nature of the edges $\mathsf{X}_\pm \in \{\mathsf{hard},\mathsf{soft}\}$, we denote by $\mathcal{V}_{\alpha_-,\mathsf{X}_-}^{\alpha_+,\mathsf{X}_+}$ the set of potentials V:

- defined at least on some interval $[a_-,a_+] \supseteq [\alpha_-,\alpha_+]$, with $a_\tau \neq \alpha_\tau$ if $\mathsf{X}_\tau = \mathsf{soft}$, and $a_\tau = \alpha_\tau$ if $X_\tau = \mathsf{hard}$;
- which satisfies the five assumptions of Section 3.6.1 on $[a_-,a_+]$, in particular is offcritical on $[a_-,a_+]$;

- for which the equilibrium measure $\mu_{\mathrm{eq}}^{V;[a_-,a_+]}$ has $[\alpha_-,\alpha_+]$ as support,
- and such that a_τ is an edge of nature X_τ.

Lemma 63. *$\mathcal{V}_{\alpha_-,\mathsf{X}_-}^{\alpha_+,\mathsf{X}_+}$ is a convex set.*

Proof. Let $V_0,V_1 \in \mathcal{V}_{\alpha_-,\mathsf{X}_-}^{\alpha_+,\mathsf{X}_+}$, and set $V_s = (1-s)V_0 + sV_1$ for $s \in [0,1]$. V_0 and V_1 are at least defined on a common interval $[a_-,a_+] \supseteq [\alpha_-,\alpha_+]$. Let us call $\nu_s = \mathrm{d}\mu_{\mathrm{eq}}^{V_s;[a_-,a_+]}$ the equilibrium measure for the potential V_s on $[a_-,a_+]$. We observe that $(1-s)\mathrm{d}\nu_0 + s\mathrm{d}\nu_1$ is a probability measure which is solution of the characterization of $\mathrm{d}L_s$ by the two usual equations

$$V_c(x) - \int \log|x-y| dL_s$$

equals C on the support and is greater than C outside the support. Therefore, $\mathrm{d}L_s = (1-s)\mathrm{d}L_0 + s\mathrm{d}L_1$. Besides, we know that there exists a function S_s, regular in a neighborhood of $[\alpha_-,\alpha_+]$ in the complex plane, positive on $[a_-,a_+]$, such that:

$$\mathrm{d}\nu_s(\xi) = \frac{\mathrm{d}\xi}{\pi}\, S_s(\xi) \sqrt{\frac{\prod_{\tau/X_\tau=\mathsf{soft}} |\xi-\alpha_\tau|}{\prod_{\tau'/X_{\tau'}=\mathsf{hard}} |\xi-\alpha_{\tau'}|}}\, \mathbf{1}_{[\alpha_-,\alpha_+]}(\xi) \tag{128}$$

for $s=0$ or $s=1$. Since the edges are of the same nature in V_0 et V_1, we must have $S_s = (1-s)S_0 + sS_1$. Since S_0 and S_1 are positive on $[a_-,a_+]$, so is S_s. Hence $V_s \in \mathcal{V}_{\alpha_-,\mathsf{X}_-}^{\alpha_+,\mathsf{X}_+}$.

Corollary 64. *Let $V_0,V_1 \in \mathcal{V}_{\alpha_-,\mathsf{X}_-}^{\alpha_+,\mathsf{X}_+}$. When a_- and a_+ satisfy the condition above, the quantity:*

$$\ln Z_{N,\beta}^{V_1;[a_-,a_+]} - \ln Z_{N,\beta}^{V_0;[a_-,a_+]} = -\frac{N\beta}{2}\int_0^1 \mathrm{d}s \oint_{C([a_-,a_+])} \frac{\mathrm{d}\xi}{2i\pi}\big(V_1(\xi)-V_0(\xi)\big) W_1^{V_s;[a_-,a_+]}(\xi) \tag{129}$$

has a large N asymptotic expansion of the form:

$$\ln Z_{N,\beta}^{V_1} - \ln Z_{N,\beta}^{V_0} = \sum_{k\geq -2} N^{-k} F_\beta^{V_0\to V_1;[a_-,a_+];\{k\}} \tag{130}$$

where:

$$F_\beta^{V_0\to V_1;[a_-,a_+];\{k\}} = -\frac{\beta}{2}\int_0^1 \mathrm{d}s \oint_{C([a_-,a_+])} \frac{\mathrm{d}\xi}{2i\pi} \sum_{m=0}^{k+2} \big(V_1^{\{m\}}(\xi) - V_0^{\{m\}}(\xi)\big)\big(W_1^{V_s;[a_-,a_+]}\big)^{\{k+1-m\}}(\xi) \tag{131}$$

Proof. Since V_s satisfies the five assumptions of Section 3.6.1 for any $s \in [0,1]$, we can apply our main theorem to $W_1^{V_s}$. Moreover, since we do not reach a critical point when s is in the compact $[0,1]$, we know that the error $O(N^{-K})$ made if we replace $W_1^{V_s}$ by $\sum_{k=-1}^{K-1} N^{-k} W_1^{V_s;\{k\}}$ is uniformly bounded with respect to s on some contour surrounding $[a_-,a_+]$ and in the analyticity domain of V. Therefore, we can exchange the integral and the sum in the asymptotic expansion. $\diamond$

For instance, when V satisfies the five assumptions of Section 3.6.1 on some interval $[a_-,a_+]$, such that $a_\pm$ are soft edges, one can interpolate between V and a gaussian potential corresponding to an equilibrium measure with support $[\alpha_-,\alpha_+]$:

$$V_{\mathrm{G},\alpha_-,\alpha_+}(x) = \frac{8}{(\alpha_+-\alpha_-)^2}\Big(x-\frac{\alpha_-+\alpha_+}{2}\Big)^2 \tag{132}$$

Proposition 65. *Let V be a potential satisfying the five assumptions of Section 3.6.1 on some interval $[a_-,a_+]$, such that $a_\pm$ are soft edges. For all $s\in[0,1]$, $(1-s)V+sV_{\mathrm{G},\alpha_-,\alpha_+}$ belongs to $\mathcal{V}^{\alpha_+,\mathsf{soft}}_{\alpha_-,\mathsf{soft}}$ and we have the following asymptotic expansion when $N\to\infty$:*

$$Z^V_{N,\beta} = Z^{V_{\mathrm{G},\alpha_-,\alpha_+}}_{N,\beta}\,\exp\Big(\sum_{k\geq -2} N^{-k}\,F_\beta^{V\to V_{\mathrm{G},\alpha_-,\alpha_+};[a_-,a_+];\{k\}}\Big) \tag{133}$$

where the prefactor is a partition function of the gaussian β ensemble (see Eqn. 50):

$$Z^{V_{\mathrm{G},\alpha_-,\alpha_+}}_{N,\beta} = Z_{N,\mathrm{G}\beta\mathrm{E}}\,\Big(\frac{\alpha_+-\alpha_-}{4}\Big)^{N+\frac{\beta}{2}N(N-1)} \tag{134}$$

According to the discussion of Section 3.7.1, we can weaken the hypothesis of the proposition above to find Theorem 39.

3.7.3 Central limit theorem

Eventually, our results imply the central limit theorem proved by Johansson [73], but here integration is taken on a compact set $[a_-,a_+]$ instead of the real line (in fact as our derivation is quite similar to Johansson's, this is not surprising). For simplicity, we take here the hypotheses of Section 3.6, although we could refine to hypotheses 37 following Section 3.7.1.

Let $h:[a_-,a_+]\to\mathbb{R}$ be a function which can be extended as a holomorphic function defined on some neighborhood of $[a_-,a_+]$, let us take $V\equiv V^{\{0\}}$ independent of N, $V^{\{1\}}=\frac{2}{\beta}\,h$ and define $V_h = V^{\{0\}}+N^{-1}\,V^{\{1\}} = V-\frac{2}{N\beta}\,h$. Then:

$$P^{V;[a_-,a_+]}_{N,\beta}\left[\exp\Big(\sum_{i=1}^N h(\lambda_i)\Big)\right] = \frac{Z^{V_h;[a_-,a_+]}_{N,\beta}}{Z^{V;[a_-,a_+]}_{N,\beta}} \tag{135}$$

and we can use Corollary 64 to derive its large N asymptotics. Indeed, we have

$$\ln P^{V;[a_-,a_+]}_{N,\beta}\left[\exp\Big(\sum_{i=1}^N h(\lambda_i)\Big)\right] = \int_0^1 \mathrm{d}s \oint_{C([a_-,a_+])}\frac{\mathrm{d}\xi}{2i\pi}\,W_1^{V_{sh}}(\xi)\,h(\xi) \tag{136}$$

By Theorem 65, or simply at the point of Lemma 61, we have:

$$\begin{aligned}
W_1^{V_{sh};\{-1\}}(\xi) &= W_1^{V;\{-1\}}(\xi) = \int \frac{\mathrm{d}\mu_{\mathrm{eq}}(\eta)}{\xi-\eta} \\
W_1^{V_{sh};\{0\}}(\xi) &= \mathcal{K}^{-1}\Bigg\{-\Big(1-\frac{2}{\beta}\Big)\Bigg[\frac{\mathrm{d}}{\mathrm{d}x}\big(W_1^{V;\{-1\}}(x)\big)+\sum_{\tau\in\mathsf{Hard}}\frac{1}{a_\tau-a_{-\tau}}\,\frac{1}{x-a_\tau}\Bigg] \\
&\qquad -\frac{2s}{\beta}\mathcal{N}_{h'}(W_1^{V;\{-1\}})(x)\Bigg\} \\
W_1^{V_{sh}} &= N\,W_1^{V_{sh};\{-1\}}+W_1^{V_{sh};\{0\}}+o(1)
\end{aligned}$$

which shows the:

Proposition 66. *Central limit theorem.*

$$\ln P_{N,\beta}^{V;[a_-,a_+]}\left[\exp\Big(\sum_{i=1}^N h(\lambda_i)\Big)\right] = N\int \mathrm{d}\mu_{\mathrm{eq}}(\eta)\,h(\eta)+m[h]+\frac{1}{2}C[h]+o(1) \tag{137}$$

with $m[h]$ the linear in the function h, given by:

$$m[h] = -\Big(1-\frac{2}{\beta}\Big)\oint_{C([a_-,a_+])}\frac{\mathrm{d}\xi}{2i\pi}\,\mathcal{K}^{-1}\left\{\frac{\mathrm{d}}{\mathrm{d}x}\big(W_1^{V;\{-1\}}(x)\big)+\sum_{\tau\in\mathsf{Hard}}\frac{1}{a_\tau-a_{-\tau}}\,\frac{1}{x-a_\tau}\right\}h(\xi) \tag{138}$$

and $C[h]$ the quadratic function of h given by:

$$C[h] = -\frac{2}{\beta}\oint_{C([a_-,a_+])}\frac{\mathrm{d}\xi}{2i\pi}\,\mathcal{K}^{-1}\Big[\mathcal{N}_{h'}(W_1^{V;\{-1\}})\Big](\xi)\,h(\xi) \tag{139}$$

Therefore $\sum_{i=1}^N h(\lambda_i)-N\int \mathrm{d}\mu_{\mathrm{eq}}(\eta)\,h(\eta)$ converges towards a Gaussian variable with mean $m[h]$ and covariance $C[h]$.

Chapter 4

Several Matrices models

In this chapter, we study matrix models , that is the laws of interacting Hermitian matrices of the form

$$d\mu_V^{N,2}(\mathbf{A}_1,\cdots,\mathbf{A}_m) := \frac{1}{Z_V^N} e^{-N\mathrm{tr}(V(\mathbf{A}_1,\cdots,\mathbf{A}_m))} d\mu^{N,2}(\mathbf{A}_1)\cdots d\mu^{N,2}(\mathbf{A}_m)$$

where Z_V^N is the normalizing constant given by the matrix integral

$$Z_V^N = \int e^{-N\mathrm{tr}(V(\mathbf{A}_1,\cdots,\mathbf{A}_m))} d\mu^{N,2}(\mathbf{A}_1)\cdots d\mu^{N,2}(\mathbf{A}_m)$$

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

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

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, $d\mu^{N,2}(\mathbf{A})$ denotes the standard law of the **GUE**, i.e under $d\mu^{N,2}(\mathbf{A})$, $\mathbf{A}$ is a $N\times N$ Hermitian matrix such that

$$A(k,l) = \bar{A}(l,k) = \frac{g_{kl} + i\tilde{g}_{kl}}{\sqrt{2N}},\ k<l, \quad A(k,k) = \frac{g_{kk}}{\sqrt{N}}$$

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

$$d\mu^{N,2}(\mathbf{A}) = Z_N^{-1} 1_{\mathbf{A}\in\mathcal{H}_N^{(2)}} e^{-\frac{N}{2}\mathrm{tr}(\mathbf{A}^2)} \prod_{1\leq i\leq j\leq N} d\Re(A(i,j)) \prod_{1\leq i<j\leq N} d\Im(A(i,j)).$$

Since we restrict ourselves to Hermitian matrices, we shall drop the subscript $\beta = 2$ and denote $\mu^N = \mu^{N,2}$.

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

$$\hat{\mathbf{L}}^N(P) := \mathbf{L}_{\mathbf{A}_1,\cdots,\mathbf{A}_m}(P) = \frac{1}{N}\mathrm{tr}\left(P(\mathbf{A}_1,\cdots,\mathbf{A}_m)\right)$$

When V vanishes, we have seen in chapter 2.3 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 pair partitions of a set of points with m colors, or equivalently the number of planar maps with one star of type P. In this part, we shall generalize such a type of result to the case where V does not vanish but is 'small' and 'nice' in a sense to precise.

This part is motivated by a work of Brézin, Parisi, Itzykson and Zuber [32] and large developments which occured thereafter in theoretical physics [49]. They specialized an idea of 't Hooft [96] to show that if $V=\sum_{i=1}^n t_i q_i$ with fixed monomials q_i of m non-commutative variables, and if we see $Z_V^N = Z_{\mathbf{t}}^N$ as a function of $\mathbf{t}=(t_1,\cdots,t_n)$

$$\log Z_{\mathbf{t}}^N := \sum_{g\geq 0} N^{2-2g} F_g(\mathbf{t}), \tag{1}$$

where

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

is a generating function of integer numbers $\mathcal{M}_g((q_i,k_i)_{1\leq i\leq k})$ which count certain graphs called maps. A map is a connected oriented graph which 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 (which 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, which is a vertex with colored edges embedded into a surface (in particular an order on the colored edges is specified). More precisely, a star of type q, for some monomial $q=X_{\ell_1}\cdots X_{\ell_k}$, is a vertex with degree $\deg(q)$ 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, $1\leq i\leq n$.

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 (1)

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

The equalities (1) and (2) derived in [32] are only formal, i.e mean that all the derivatives on both sides of the equality coincide at $\mathbf{t}=0$. They can thus be deduced from Wick formula (which gives the expression of arbitrary moments of Gaussian variables) or equivalently by the use of Feynman diagrams.

Eventhough topological expansions such as (1) and (2) were first introduced by 't Hooft in the course of computing the integrals, the natural reverse question of computing the numbers $\mathcal{M}_g((q_i,k_i)_{1\le i\le k})$ by studying the associated integrals over matrices encountered a large success in theoretical physics (see e.g. the review papers [49, 57]). 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 is equal to $F_0(\mathbf{t})$. There is here the belief that one can interchange derivatives and limit, a claim that needs to be justified.

We shall indeed prove that the formal limit can be strenghtened into a large N expansion in the sense that

$$\mu_V^N[\hat{\mathbf{L}}_N(P)] = \sigma_0^V(P) + \frac{1}{N^2}\sigma_1^V(P) + o(N^{-2})$$

where $\sigma_g^V(q) = \sum_{k_1,\cdots,k_n\in\mathbb{N}^k}\prod_{i=1}^k \frac{(-t_i)^{k_i}}{k_i!}\mathcal{M}_g((q_i,k_i)_{1\le i\le k},(q,1))$ for monomial functions q and $g=0$ or 1 and $N^2\times o(N^{-2})$ goes to zero as N goes to infinity. This asymptotic expansion holds when V is small and satisfies some convexity hypothesis (which insures that the partition function Z_V^N is finite and the support of the limiting spectral measures of $\mathbf{A}_i$, $1\le i\le m$, under μ_V^N is connected, see [63]).

This part summarizes results from [59] and [62]. The full expansion (i.e higher order corrections) was obtained by E. Maurel Segala [79] in the multi-matrix setting. Such expansion in the one matrix case was already derived on a physical level of rigour in [5] and then made rigorous in [3, 51]. However, in the case of one matrix, orthogonal polynomial can be used to develop for instance Riemann Hilbert techniques as in [51]. In the multi-matrix case this approach fails in general (or at list has not yet been extended). [59, 62, 79] take a completly different route based on the free probability setting of limiting tracial states and of the so-called Master loop or Schwinger-Dyson equations. We start by introducing the formal expansion of Brezin, Itzykson, Parisi and Zuber. At the end of chapter **??**, we summarize some tricks to effectively compute $F_0(\mathbf{t})$, and therefore deduce from the asymptotic relations between matrix models and planar maps the actual enumeration of these graphs. The techniques we shall present here have the advantage to be robust. We use them here to study partition functions of Hermitian matrices, but they can be generalized to orthogonal or symplectic matrices or to matrices following the Haar measure on the unitary group [**?**]. The last extension is particularly interesting since then Gaussian calculus and Feynman diagrams techniques fail (since unitary matrices have no Gaussian entries) so that the diagrammatic representation of the limit is not straightforward even on a formal level (see [41] for a formal expansion with no diagrammatic interpretation).

4.1 Formal expansion of matrix integrals

The expansion obtained by 't Hooft is based on Feynman diagrams, or equivalently on Wick Formula which states as follows.

Lemma 67 (Wick's formula)**.** 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 $\{1, \cdots, 2n\}$.

Proof. Recall that if G is a standard gaussian variable, for all $n \in \mathbb{N}$,

$$\mathbb{E}[G^{2n}] = 2n!! := \frac{(2n)!}{2^n n!}$$

is the number of pair-partitions of the ordered set $\{1, 2, \cdots, 2n\}$. Thus, for any real numbers $(\alpha_1, \cdots, \alpha_{2n})$, since $\sum_{i=1}^{2n} \alpha_i G_i$ is a centered Gaussian variable with covariance $\sigma^2 = \sum_{i,j=1}^{2n} \alpha_i \alpha_j \mathbb{E}[G_i G_j]$, and so $\sum_{i=1}^{2n} \alpha_i G_i$ has the same law that σ times a standard Gaussian variable,

$$\mathbb{E}[\left(\sum_{i=1}^{2n} \alpha_i G_i\right)^{2n}] = (\sum_{i,j=1}^{2n} \alpha_i \alpha_j \mathbb{E}[G_i G_j])^n 2n!!.$$

Identifying on both sides the term corresponding to the coefficient $\alpha_1 \cdots \alpha_{2n}$, we obtain

$$(2n)! \mathbb{E}[G_1 \cdots G_{2n}] = 2n!! \sum_{\pi \in \Sigma} \prod_{(b,b') \in \pi} \mathbb{E}[G_b G_{b'}]$$

where Σ is the set of pairs of $2n$ elements. To compare this set with the collection of pairings of an ordered set, we have to order the elements of the pairs, and we have 2^n possible choices, and then order the pairs, which gives another $n!$ possible choices. Thus,

$$\sum_{\pi \in \Sigma} \prod_{(b,b') \in \pi} \mathbb{E}[G_b G_{b'}] = 2^n n! \sum_{\pi \in PP(2n)} \prod_{(i,j) \text{ block of } \pi} \mathbb{E}[G_i G_j]$$

completes the argument as $2^n n! 2n!! = (2n)!$. $\square$

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 68. Let $r, m \in \mathbb{N}$. Let $q_1, \cdots, q_r$ be r monomials in m non-commutative variables. A map of genus g with a star of type q_i for $i \in \{1, \cdots, r\}$ is a connected graph embedded into a surface of genus g with r vertices so that

1. for $1 \leq i \leq r$, 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 embedded into the surface in such a way that the orientation of the star

and the orientation of the surface agree). The half-edges inherit the orientation of their stars, i.e each side of each half-edge is endowed with an opposite orientation corresponding to the orientation of a path travelling around the star by following the orientation of the star.

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 orientation; thus edges have only one color and one orientation.

3. A path travelling along the edges of the map following their orientation will make a loop. The surface inside this loop is homeomorphic to a disk and called a face (see figure 4.1).

Note that each star has a distinguished half-edge and so each edge of a star is labelled. Moreover, all stars are labelled. Hence, the enumeration problem we shall soon consider can be thought as the problem of matching the labelled half-edges of the stars and so we will distinguish all the maps where the gluings are not done between exactly the same set of half-edges, regardless of symmetries. This is important to make clear since we shall shortly consider enumeration issues. The genus of a map is defined as in Definition **??**. Note that since at each vertex we imposed a cyclic orientation at the ends of the edges adjacent to this vertex, there is a unique way to embed the graph drawn with stars in a surface; we have to draw the stars so that their orientation agrees with the orientation of the surface.

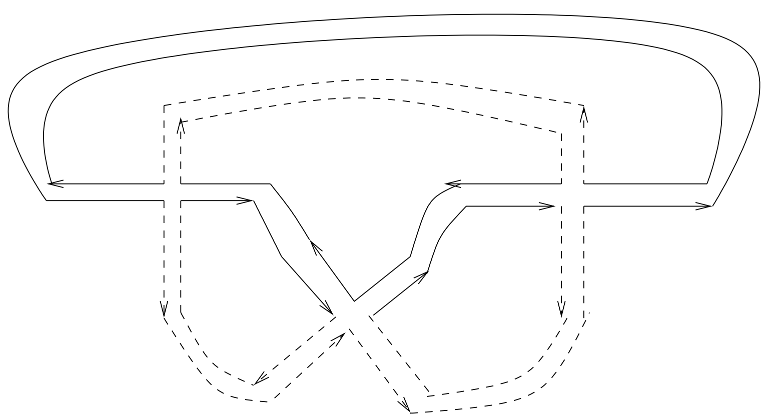


Figure 4.1: A planar bi-colored map with stars of type $q_1 = X_1X_2X_1X_2$, $q_2 = X_1^2X_2^2$, $q_3 = X_1X_2X_1X_2$

There is a dual way to consider maps in the spirit of figure (**??**); as in the figure in the center of figure (**??**), 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 covering of a surface (with the same genus as the map) by polygonals of type $q_1, \cdots, q_r$. The constraint on the colors becomes a constraint on the colors of the sides of the polygons of the covering.

Example 69. 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. $q(X) = X^4$) .

Exercise 70. Draw the quadrangulation corresponding to figure 4.1.

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

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

$$\text{and } k_i \text{ stars of type } q_i, 1 \le i \le n\}.$$

Note here that the stars are labelled in the counting. Hence, the problem amounts to count the possible matchings of the half-edges of the stars, all the half-edges being labelled.

In this section we shall first encounter eventually 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 which 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 labelled. Moreover, we shall count these graphs up to homeomorphism, that is up to continuous deformation of the surface on which the graphs are embedded. Thus, our problem is to enumerate the number of possible pairings of the half-edges (of a given color) of the stars in such a way that the resulting graph has a given genus.

We now argue that

Lemma 71. Let $q_1, \cdots, q_n$ be monomials. Then,

$$\int \prod_{i=1}^{n} (N\text{tr}(q_i(\mathbf{A}_1, \cdots, \mathbf{A}_m))) d\mu^N(\mathbf{A}_1) \cdots d\mu^N(\mathbf{A}_m)$$

$$= \sum_{g \in \mathbb{N}} \sum_{c \ge 1} \frac{1}{N^{2g-2c}} \sharp\{G_{g,c}((q_i, 1), 1 \le i \le n)\}$$

Here $\sharp\{G_{g,c}((q_i, 1), 1 \le i \le n)\}$ is the number of graphs of the set $G_{g,c}((q_i, 1), 1 \le i \le n)$.

As a warm up, let us show that

Lemma 72. Let q be a monomial. Then,we have the following expansion

$$\int N^{-1}\text{tr}(q(\mathbf{A}_1, \cdots, \mathbf{A}_m)) d\mu^N(\mathbf{A}_1) \cdots d\mu^N(\mathbf{A}_m) = \sum_{g \in \mathbb{N}} \frac{1}{N^{2g}} \sharp\{G_g((q, 1))\}$$

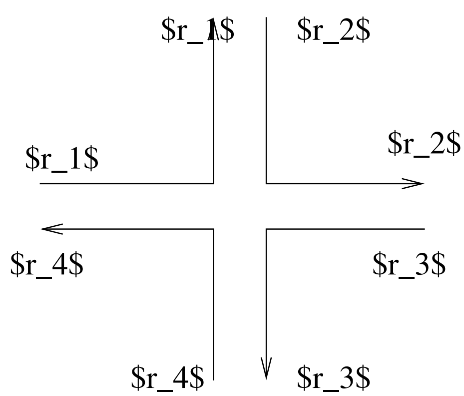
where $\sharp\{G_0((q, 1))\}$ equals $\sigma^m(q)$ as found by Voiculescu, Theorem 4.

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

$$\begin{aligned}
&\int \operatorname{tr}(q(\mathbf{A}_1, \cdots, \mathbf{A}_m)) d\mu^N(\mathbf{A}_1) \cdots d\mu^N(\mathbf{A}_m) \\
= &\sum_{1 \le r_1, \cdots, r_k \le N} \int A_{j_1}(r_1, r_2) \cdots A_{j_k}(r_k, r_1) d\mu^N(\mathbf{A}_1) \cdots d\mu^N(\mathbf{A}_m) \\
= &\sum_{r_1, \cdots, r_k} \sum_{\pi \in PP(k)} \prod_{\substack{(wv) \text{ block of } \pi \\ w<v}} \mathbb{E}[A_{j_w}(r_w, r_{w+1}) A_{j_v}(r_v, r_{v+1})]. \qquad (3)
\end{aligned}$$

Note that $\prod \mathbb{E}[A_{j_w}(r_w, r_{w+1}) A_{j_v}(r_v.r_{v+1})]$ 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 graphs 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 variables sit 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 4.1). 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 4.2: 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 correponding 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. (3) thus shows that

$$\begin{aligned}&\int \mathrm{tr}(q(\mathbf{A}_1,\cdots,\mathbf{A}_m))d\mu^N(\mathbf{A}_1)\cdots d\mu^N(\mathbf{A}_m)\\ = &\sum_{g\geq 0}\frac{N^F}{N^{\frac{k}{2}}}\sharp\{\text{maps with one star of type } q \text{ and } F \text{ faces}\}\end{aligned}$$

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

Remark 73. *In the above it is important to take μ^N to be the law of the GUE (and not GOE for instance) to insure that $E[(A_k)_{ij}(A_k)_{ji}]=1/N$ but $E[((A_k)_{ij})^2]=0$. 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 71. 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;

$$\begin{aligned}&\int\prod_{i=1}^n(N\mathrm{tr}(q_i(\mathbf{A}_1,\cdots,\mathbf{A}_m)))d\mu^N(\mathbf{A}_1)\cdots d\mu^N(\mathbf{A}_m)\\ &=N^n\sum_{\substack{i_1^k,\cdots,i_{d_k}^k\\ 1\leq k\leq n}}\mathbb{E}[\prod_{1\leq k\leq n}\mathbf{A}_{\ell_1^k}(i_1^k i_2^k)\cdots\mathbf{A}_{\ell_{d_k}^k}(i_{d_k}^k i_1^k)]\\ &=N^n\sum_{\substack{i_1^k,\cdots,i_{d_k}^k\\ 1\leq k\leq n}}\sum_{\pi\in PP(\sum d_i)}Z(\pi,\mathbf{i})\end{aligned}$$

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 variances over the corresponding blocks of the partition. A pictorial way to represent this sum over $PP(\sum d_i)$ is to represent $X_{\ell_1^k}(i_1^k i_2^k)\cdots X_{\ell_{d_k}^k}(i_{d_k}^k i_1^k)$ by its associated star of type q_k, for $1\leq k\leq n$. Note that in the counting this star will be labelled (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 variances, vanishes 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 $E[A_r(i,j)A_r(k,l)]$ 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, we find that

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

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 egdes of the graphs is $\sharp edges = \sum_{i=1}^n k_i/2$ and the total number of vertices is $\sharp vertices = n$, we see that if $g_i, 1 \leq i \leq c$, are the genera of each connected component of our graph, we must have

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

This completes the proof. □

We then claim that we find the topological expansion of Brézin, Itzykson, Parisi and Zuber [32] :

Lemma 74. Let $q_1, \cdots, q_n$ be monomials. Then, we have the folowing formal expansion

$$\begin{aligned}&\log\left(\int e^{\sum_{i=1}^n t_i N \mathrm{tr}(q_i(\mathbf{A}_1,\cdots,\mathbf{A}_m))} d\mu^N(\mathbf{A}_1)\cdots d\mu^N(\mathbf{A}_m)\right)\\ &= \sum_{g\geq 0} \frac{1}{N^{2g-2}} \sum_{k_1,\cdots,k_n\in\mathbb{N}} \prod_{i=1}^{n} \frac{(t_i)^{k_i}}{k_i!} \mathcal{M}_g((q_i,k_i), 1\leq i\leq n)\end{aligned}$$

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 convergent (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 chapters 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.

Proof of Lemma 74. The idea is to expand 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

$$\begin{aligned}
L \quad := \quad & \int e^{\sum_{i=1}^n t_i N \mathrm{tr}(q_i(\mathbf{A}_1,\cdots,\mathbf{A}_m))} d\mu^N(\mathbf{A}_1)\cdots d\mu^N(\mathbf{A}_m)\\
&= \int \prod_{i=1}^{n}\left(e^{t_i N \mathrm{tr}(q_i(\mathbf{A}_1,\cdots,\mathbf{A}_m))}\right) d\mu^N(\mathbf{A}_1)\cdots d\mu^N(\mathbf{A}_m)\\
&= \int \prod_{i=1}^{n}\left(\sum_{k_i\geq 0} \frac{(t_i)^{k_i}}{k_i!} (N\mathrm{tr}(q_i(\mathbf{A}_1,\cdots,\mathbf{A}_m)))^{k_i}\right) d\mu^N(\mathbf{A}_1)\cdots d\mu^N(\mathbf{A}_m)\\
&= \sum_{k_1,\cdots,k_n\in\mathbb{N}} \frac{(t_1)^{k_1}\cdots(t_n)^{k_n}}{k_1!\cdots k_n!} \int \prod_{i=1}^{n} (N\mathrm{tr}(q_i(\mathbf{A}_1,\cdots,\mathbf{A}_m)))^{k_i} d\mu^N(\mathbf{A}_1)\cdots d\mu^N(\mathbf{A}_m)\\
&= \sum_{k_1,\cdots,k_n\in\mathbb{N}} \frac{(t_1)^{k_1}\cdots(t_n)^{k_n}}{k_1!\cdots k_n!} \sum_{g\geq 0}\sum_{c\geq 0} \frac{1}{N^{2g-2c}} \sharp\{G_{g,c}((q_i,k_i), 1\leq i\leq n)\} \qquad (4)
\end{aligned}$$

where we finally used Lemma 71. 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 $\sharp\{G_{g,c}((q_i,k_i),1\le i\le n)\}$ with the number of maps. Since graphs in $G_{g,c}((q_i,k_i),1\le i\le n)$ can be decomposed into a union of disconnected maps, $\sharp\{G_{g,c}((q_i,k_i),1\le i\le n)\}$ is related with the ways to distribute the stars and the genus among the c maps, and the number of each of these maps. In other words, we have (since all stars are labelled)

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

Plugging this expression into (4) we get

$$\begin{aligned}
L\ :=\ &\sum_{k_1,\cdots,k_n\in\mathbb{N}}\frac{(t_1)^{k_1}\cdots(t_n)^{k_n}}{c!k_1!\cdots k_n!}\sum_{g\ge 0}\sum_{c\ge 0}\frac{1}{N^{2g-2c}}\sum_{\substack{\sum_{i=1}^c g_i=g\\ g_i\ge 0}}\frac{g!}{g_1!\cdots g_c!}\times\\
&\sum_{\substack{\sum_{j=1}^c l_i^j=k_i\\ 1\le j\le n}}\prod_{i=1}^n\frac{k_i!}{l_i^1!\cdots l_i^c!}\prod_{j=1}^c\mathcal{M}_g((q_i,l_i^j),1\le i\le n)\\
=\ &\sum_{c\ge 0}\frac{1}{c!}\sum_{g=\sum_{i=1}^c g_i}\frac{g!}{g_1!\cdots g_c!}\sum_{k_1,\cdots,k_n\in\mathbb{N}}\sum_{\substack{\sum_{j=1}^c l_i^j=k_i\\ 1\le j\le n}}\prod_{j=1}^c\left(\frac{1}{N^{2g_j-2}}\prod_{i=1}^n\frac{(t_i)^{l_i^j}}{l_i^j!}\mathcal{M}_g((q_i,l_i^j),1\le i\le n)\right)\\
=\ &\sum_{c\ge 0}\frac{1}{c!}\left(\sum_{g\ge 0}\frac{1}{N^{2g-2}}\sum_{l_1,\cdots l_n\ge 0}\prod_{i=1}^n\frac{(t_i)^{l_i}}{l_i!}\mathcal{M}_g((q_i,l_i),1\le i\le n)\right)^c\\
=\ &\exp\left(\sum_{g\ge 0}\frac{1}{N^{2g-2}}\sum_{l_1,\cdots l_n\ge 0}\prod_{i=1}^n\frac{(t_i)^{l_i}}{l_i!}\mathcal{M}_g((q_i,l_i),1\le i\le n)\right)
\end{aligned}$$

which completes the proof. □

The goal of the next chapters 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 migh also have been proven by use of differential calculus). The point here will be that we can design an asymptotic framework for differential calculus, which will then encode the combinatorics of the first order 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.

4.2 Combinatorics of maps and non-commutative polynomials

In this chapter, we introduce non-commutative polynomials and non-commutative laws such as the 'empirical distribution' of matrices $\mathbf{A}_1,\cdots,\mathbf{A}_m$ simply given as the complex valued linear functional on the set of polynomials which associate to a polynomial the normalized trace of the polynomial evaluated at $\mathbf{A}_1,\cdots,\mathbf{A}_m$. We will then describe precisely the combinatorial objects related with matrix integrals. Recalling the bijection between non-commutative monomials and graphical objects such as stars or ordered sets of colored point, we will show how 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 some surgery on maps, as introduced by Tutte [99] to prove induction relations for maps enumeration.

4.2.1 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

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

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 $\mathbf{A} = (\mathbf{A}_1,\cdots,\mathbf{A}_m)$ in the set $\mathcal{H}_N$ 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.

4.2.2 Convexity

We will say that V is convex if V is self-adjoint and for any $N\in\mathbb{N}$

$$\begin{array}{lccc} \varphi_V^N & :\mathcal{H}_N^m & \longrightarrow & \mathbb{R} \\ & (\mathbf{A}_1,\cdots,\mathbf{A}_m) & \longrightarrow & \mathrm{tr}(V(\mathbf{A}_1,\cdots,\mathbf{A}_m)) \end{array}$$

is a convex function of the entries of the Hermitian matrices $\mathbf{A}_1,\cdots,\mathbf{A}_m$.

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 by the notion of c-convexity.

Definition 75. 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

$$\begin{array}{lccc}\varphi_V^{N,c}: & \mathcal{E}_N^{(2)} & \longrightarrow & \mathbb{R} \\ & (\Re(A_k(i,j)),\Im(A_k(i,j)))^{1\le k\le m}_{1\le i\le j\le N} & \longrightarrow & \mathrm{tr}(V(\mathbf{A}_1,\cdots,\mathbf{A}_m)+\frac{1-c}{2}\sum_{k=1}^m \mathbf{A}_i^2)\end{array}$$

is non-negative. Here, for $k\in\{1,\cdots,m\}$, $\mathbf{A}_k$ is the Hermitian matrix with entries $\sqrt{2}^{-1}(A_k(p,q)+iA_k(q,p))$ above the diagonal and $A_k(i,i)$ on the diagonal.

An example of c-convex potential is

$$V=\sum_{i=1}^n P_i\left(\sum_{k=1}^m \alpha_k^i \mathbf{A}_k\right)+\sum_{k,l=1}^n \beta_{k,l}\mathbf{A}_k\mathbf{A}_l$$

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 22.

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 part will apply, in particular concentration inequalities and Brascamp-Lieb inequalities.

In the rest of this chapter, we assume that V is c-convex for some $c>0$ fixed. Arbitrary potentials could be considered as far as first order asymptotics are studied in [60], at the price of adding a cutoff. In fact, adding a cutoff and choosing the parameters t_i's small enough (depending eventually 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. Since $V=V_{\mathbf{t}}$ with $\mathbf{t}$ varying but fixed monomials, we will let $U_c=\{\mathbf{t}:V_{\mathbf{t}}\text{ is c-convex}\}\subset\mathbb{C}^n$. Moreover, B_η will denote the open ball in $\mathbb{C}^n$ centered at the origin and with radius $\eta>0$ (for instance for the metric $|\mathbf{t}|=\max_{1\le i\le n}|t_i|$).

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

$$\partial_i PQ=\partial_i P\times(1\otimes Q)+(P\otimes 1)\times\partial_i Q \tag{5}$$

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. 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 $2^{-1}[D_i^2 P\sharp(R,S) + D_i^2 P\sharp(S,R)]$ 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

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

Non-commutative laws

For $(\mathbf{A}_1,\cdots,\mathbf{A}_m) \in \mathcal{H}_N^m$, let us define the linear form $\mathbf{L}_{\mathbf{A}_1,\cdots,\mathbf{A}_m}$ from $\mathbb{C}\langle X_1,\cdots,X_m\rangle$ into $\mathbb{C}$ by

$$\mathbf{L}_{\mathbf{A}_1,\cdots,\mathbf{A}_m}(P) = \frac{1}{N}\mathrm{tr}\left(P(\mathbf{A}_1,\cdots,\mathbf{A}_m)\right)$$

where tr is the standard trace $\mathrm{tr}(A) = \sum_{i=1}^N A(i,i)$. $\mathbf{L}_{\mathbf{A}_1,\cdots,\mathbf{A}_m}$ will be called the empirical distribution of the matrices (note that in the case of one matrix, it is the empirical distribution of the eigenvalues of this matrix). When the matrices $\mathbf{A}_1,\cdots,\mathbf{A}_m$ are generic and distributed according to μ_V^N, we will drop the subscripts $\mathbf{A}_1,\cdots,\mathbf{A}_m$ and write in short $\hat{\mathbf{L}}^N = \mathbf{L}_{\mathbf{A}_1,\cdots,\mathbf{A}_m}$. We denote, when $V = V_{\mathbf{t}} = \sum_{i=1}^n t_i q_i$,

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

$\hat{\mathbf{L}}^N, \overline{\mathbf{L}}_t^N$ will be seen as elements of the algebraic dual $\mathbb{C}\langle X_1,\cdots,X_m\rangle^{\mathcal{D}}$ of $\mathbb{C}\langle X_1,\cdots,X_m\rangle$ equipped with the involution $*$. $\mathbb{C}\langle X_1,\cdots,X_m\rangle^{\mathcal{D}}$ is equipped with its weak topology.

Definition 76. A sequence $(\mu_n)_{n\in\mathbb{N}}$ in $\mathbb{C}\langle X_1,\cdots,X_m\rangle^{\mathcal{D}}$ converges weakly (or in moments) to $\mu \in \mathbb{C}\langle X_1,\cdots,X_m\rangle^{\mathcal{D}}$ iff for any $P \in \mathbb{C}\langle X_1,\cdots,X_m\rangle$,

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

Lemma 77. Let $C(\ell_1,\cdots,\ell_r), \ell_i \in \{1,\cdots,m\}, r\in\mathbb{N}$, be finite non-negative constants and

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

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

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 $(\varphi(n), n\in\mathbb{N})$, the terms $\mu_{\varphi(n)}(X_{\ell_1}\cdots X_{\ell_r}), \ell_i \in \{1,\cdots,m\}, r\in\mathbb{N}$ converge simultaneously. The limit defines an element of $\mathbb{C}\langle X_1,\cdots,X_m\rangle^{\mathcal{D}}$ by linearity. □

The following is a triviality, that we however recall since we will use it several times.

Corollary 78. *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 $K(C)$ which has a unique limit point. Then $(\mu_n)_{n\in\mathbb{N}}$ converges weakly (or in moments) to this limit point.*

Proof. Otherwise we could choose a subsequence which 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 $\mathbb{C}\langle X_1,\cdots,X_m\rangle^{\mathcal{D}}$. $\square$

Remark 79. *The laws $\hat{\mathbf{L}}^N, \overline{\mathbf{L}}_t^N$ are more than only linear forms on $\mathbb{C}\langle X_1,\cdots,X_m\rangle$; they satisfy also the properties*

$$\mu(PP^*) \geq 0, \quad \mu(PQ) = \mu(QP), \quad \mu(1) = 1 \tag{6}$$

for all polynomial functions P,Q. Since these conditions are closed for the weak topology, we see that any limit point of $\hat{\mathbf{L}}^N, \overline{\mathbf{L}}_t^N$ will as well satisfy these properties. A linear functional on $\mathbb{C}\langle X_1,\cdots,X_m\rangle$ which satisfies such conditions are called tracial states , or non-commutative laws. 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.

Maps and polynomials

In this section, we complete section 2.3.1 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 variables.

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

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

and for a monomial P

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

Maps and polynomials

Because there is a one to one mapping between stars and monomials, the operations on monomials such as involution or derivations have their graphical interpretation.

The involution comes to reverse the orientation and to shift the marked edge by one in the sense of the new orientation. This is equivalent to consider the star in a mirror. For derivations, the interpretation goes as follows.

Let q be a given monomial. The derivation ∂_i appears as a way to find out how to decompose a star of type q by pointing out an half-edge of color i: a star of type q can indeed be decomposed into one star of type q_1, one half-edge of color i and another star of type q_2, all sharing the same vertex, iff q can be written as $q = q_1 X_i q_2$. This is particularly useful to write induction relation on the number of maps. For instance, let us consider a planar map M and the event $A_M(X_i q)$ that, inside M, a star of type $X_i q$ is such that the first marked half-edge is glued with an half-edge of q. Then, if this happens, since the map is planar, it will be decomposed into two planar maps separated by the edge between these two X_i. Such a gluing can be done only with the edges X_i appearing in the decomposition of q as $q = q_1 X_i q_2$. Moreover, the two stars of type q_1 and q_2 will belong to two 'independent' planar maps. So, we can symbolically write

$$1_{A_M(X_i q)} = \sum_{q=q_1 X_i q_2} 1_{q_1 \in M_1} \otimes_{M=M_1 \otimes_i M_2} 1_{q_2 \in M_2} \tag{7}$$

where $M = M_1 \otimes_i M_2$ means that M decomposes into two planar maps M_1 and M_2, M_2 being surrounded by a cycle of color i which separates it from M_1 (see figure 4.2.2). Note here that we forgot in some sense that these 3 objects were sharing the same

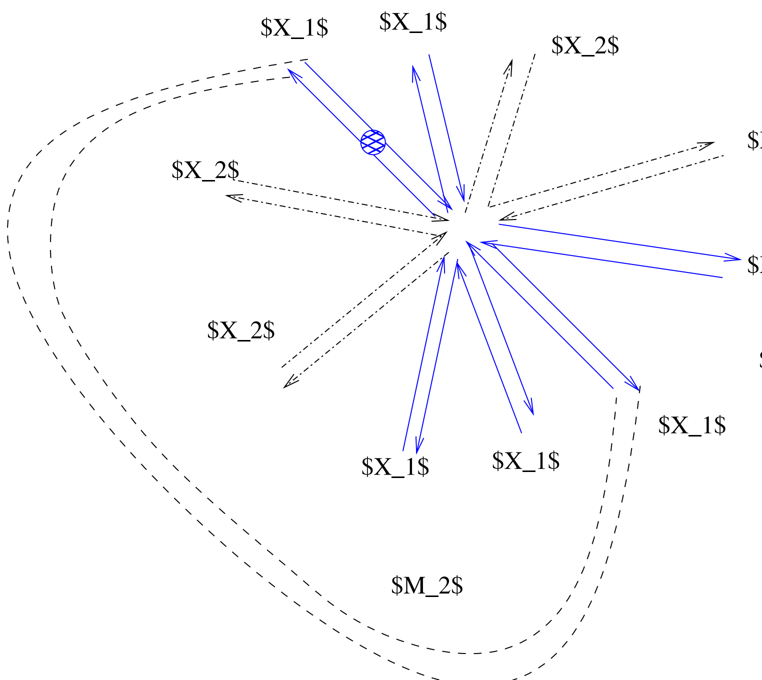


Figure 4.3: a star of type $q = X_1^2 X_2^2 X_1^4 X_2^2$ decomposed into $X_1(X_1 X_2^2 X_1) X_1 (X_1^2 X_2^2)$

vertex; this is somehow irrelevant here since a vertex is finally nothing but the point of junction of several edges; as long as we are concerned with the combinatorial problem of enumerating these maps, we can safely split the map M into these 3 objects. (4.2.2) is very close to the derivation operation ∂_i.

Similarly, let us consider again a planar map M containing given stars of type $X_i q$ and q' and the event $B_M(X_i q, q')$ that, inside M, the star of type $X_i q$ is such that the first marked half-edge is glued with an half-edge of the star of type q'. Once we know that this happens, we can write

$$1_{B_M(X_i q, q')} = \sum_{q'=q_1 X_i q_2} 1_{q_2 q_1 \bullet_i q \in M}. \tag{8}$$

$q_2 q_1 \bullet_i q$ is a new star made of a star of type q and one of type $q_2 q_1$ with an edge of color i from one to the other just before the marked half-edges. Again, once we know that this edge of color i exists, from a combinatorial point of view, we can simply shorten it till the two stars merge into a bigger star of type $q_2 q_1 q$. This is the merging operation; it corresponds to the cyclic derivative D_i (see figure 4.2.2).

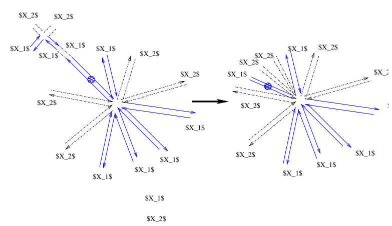


Figure 4.4: Merging of a star of type $q = X_1^2 X_2^2 X_1^4 X_2^2$ and a star of type $X_1^2 X_2^2$

4.3 First order expansion

At the end of this chapter (see Theorem 13) we will have proved that Lemma 74 holds as a first order limit, i.e.

$$\lim_{N\to\infty} \frac{1}{N^2} \log \int e^{\sum_{i=1}^n t_i N \mathrm{tr}(q_i(\mathbf{A}_1,\cdots,\mathbf{A}_m))} d\mu^N(\mathbf{A}_1)\cdots d\mu^N(\mathbf{A}_m)$$

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

provided the parameters $\mathbf{t} = (t_i)_{1\le i\le n}$ are sufficiently small and such that the polynomial $V = \sum t_i q_i$ is strictly convex (i.e belong to $U_c \cap B_\eta$ for some ¿0 and $\eta \le \eta(c)$ for some $\eta(c) > 0$). To prove this result we first show that, under the same assumptions, $\overline{\mathbf{L}}_t^N(q) = \mu^N_{\sum t_i q_i}(N^{-1}\mathrm{tr}(q))$ converges as N goes to infinity to a limit which is as well related with map enumeration (see Theorem 11).

The central tool in our asymptotic analysis will be the so-called Schwinger-Dyson (or loop) equations. In finite dimension, they are simple emanation of the integration by parts formula (or, somewhat equivalently, of the symmetry of the Laplacian in $L^2(dx)$). As dimension goes to infinity, concentration inequalities show that $\overline{\mathbf{L}}_t^N$ approximately satisfies a closed equation that we will simply refer to as the Schwinger-Dyson equation. The limit points of $\overline{\mathbf{L}}_t^N$ will therefore satisfy this equation. We will then show that this equation has a unique solution in some small range of the parameters. As a consequence, $\overline{\mathbf{L}}_t^N$ will converge to this unique solution. Showing that an appropriate generating function of maps also satisfies the same equation will allow us to determine the limit of $\overline{\mathbf{L}}_t^N$.

4.3.1 Finite dimensional Schwinger-Dyson's equations

Property 80. *For all $P \in \mathbb{C}\langle X_1, \cdots, X_m\rangle$, all $i \in \{1, \cdots, m\}$,*

$$\mu^N_{V_\mathbf{t}}\left(\hat{\mathbf{L}}^N \otimes \hat{\mathbf{L}}^N(\partial_i P)\right) = \mu^N_{V_\mathbf{t}}\left(\hat{\mathbf{L}}^N((X_i + D_i V_\mathbf{t})P)\right)$$

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

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

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 $r \in \{1, \cdots, m\}$,

$$N \int A_l(r,s) f(A_k(i,j), 1 \le i,j \le N, 1 \le k \le m) d\mu^N(\mathbf{A}_1) \cdots d\mu^N(\mathbf{A}_m) =$$

$$\int \partial_{A_l(s,r)} f(A_k(i,j), 1 \le i,j \le N, 1 \le k \le m) d\mu^N(\mathbf{A}_1) \cdots d\mu^N(\mathbf{A}_m).$$

Using repeatidly this equality, we arrive at

$$\begin{aligned} \int \frac{1}{N} \mathrm{tr}(\mathbf{A}_k P) d\mu^N_V(\mathbf{A}) &= \frac{1}{2N^2} \sum_{i,j=1}^N \int \partial_{A_k(j,i)} (P e^{-N\mathrm{tr}(V)})_{ji} \prod d\mu^N(\mathbf{A}_i) \\ &= \frac{1}{2N^2} \sum_{i,j=1}^N \int \Bigg(\sum_{P = QX_kR} 2 Q_{ii} R_{jj} \\ &\quad -N \sum_{l=1}^n t_l \sum_{q_l = QX_kR} \sum_{h=1}^N 2 P_{ji} Q_{hj} R_{ih} \Bigg) d\mu^N_V(\mathbf{A}) \\ &= \int \left(\frac{1}{N^2} (\mathrm{tr} \otimes \mathrm{tr})(\partial_k P) - \frac{1}{N} \mathrm{tr}(D_k V P) \right) d\mu^N_V(\mathbf{A}) \end{aligned}$$

which yields

$$\int \left(\hat{\mathbf{L}}^N((X_k + D_k V)P) - \hat{\mathbf{L}}^N \otimes \hat{\mathbf{L}}^N(\partial_k P)\right) d\mu_V^N(\mathbf{A}) = 0. \tag{9}$$

□

4.3.2 Tightness and limiting Schwinger-Dyson's equations

We say that $\tau \in \mathbb{C}\langle X_1, \cdots, X_m\rangle^{\mathcal{D}}$ satisfies the Schwinger-Dyson equation with potential V, denoted in short **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

Property 81. *Assume that $V_{\mathbf{t}}$ is c-convex. Then, $(\overline{\mathbf{L}}_t^N, N \in \mathbb{N})$ is tight. Its limit points satisfy* **SD[$V_{\mathbf{t}}$]** *and*

$$|\tau(X_{\ell_1} \cdots X_{\ell_r})| \leq M_0^r \tag{10}$$

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

By Lemma 28, we find that for all $\ell_1, \cdots, \ell_r$,

$$\begin{aligned} |\overline{\mathbf{L}}_t^N(X_{\ell_1} \cdots X_{\ell_r})| &\leq \mu_{V_{\mathbf{t}}}^N(|\lambda_{max}(\mathbf{A})|^r) \\ &= \int_0^\infty r x^{r-1} \mu_{V_{\mathbf{t}}}^N(|\lambda_{max}(\mathbf{A})| \geq x) dx \qquad (11) \\ &\leq M_0^r + \int_{M_0}^\infty r x^{r-1} e^{-\alpha N x} dx \\ &= M_0^r + r(\alpha N)^{-r} \int_0^\infty r x^{r-1} e^{-x} dx \qquad (12) \end{aligned}$$

Hence, if $K(C)$ denotes the compact set defined in Lemma 78, $\overline{\mathbf{L}}_t^N \in K(C)$ with $C(\ell_1, \cdots, \ell_r) = M_0^r + r\alpha^{-r} \int_0^\infty r x^{r-1} e^{-x} dx$. $(\overline{\mathbf{L}}_t^N, N \in \mathbb{N})$ is therefore tight. Let us consider now its limit points; let τ be such a limit point. By (12), we must have

$$|\tau(X_{\ell_1} \cdots X_{\ell_r})| \leq M_0^r. \tag{13}$$

Moreover, by concentration inequalities (see Lemma 31), 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 80 implies that

$$\limsup_{N\to\infty} \left| \overline{\mathbf{L}}_t^N((X_k + D_k V_{\mathbf{t}})P)) - \overline{\mathbf{L}}_t^N \otimes \overline{\mathbf{L}}_t^N(\partial_k P) \right| = 0. \tag{14}$$

Hence, (9) shows that

$$\tau((X_k + D_k V)P) = \tau \otimes \tau(\partial_k P). \tag{15}$$

□

Uniqueness of the solutions to Schwinger-Dyson's equations for small parameters

Let $R \in \mathbb{R}^+$ (we will always assume $R \geq 1$ in the sequel).

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

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

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 $1 \leq i \leq n$, $\deg(q_i) =: d_i \leq D$ and equality holds for some i.

The main result of this paragraph is

Theorem 10. *For all* $R \geq 1$*, there exists* $\varepsilon > 0$ *so that for* $|\mathbf{t}| = max_{1 \leq i \leq n} |t_i| < \varepsilon$*, there exists at most one solution* $\tau_{\mathbf{t}}$ *to* $\mathbf{SD}[V_{\mathbf{t}}]$ *which satisfies* **(CS(R))**.

Remark: Note that if $V = 0$, our equation becomes

$$\tau(X_i P) = \tau \otimes \tau(\partial_i P).$$

Because if P is a monomial, $\tau \otimes \tau(\partial_i P) = \sum_{P = P_1 X_i P_2} \tau(P_1)\tau(P_2)$ with P_1 and P_2 with degree smaller than P, we see that the equation **SD[0]** allows to define uniquely $\tau(P)$ for all P by induction. The solution can be seen to be exactly $\tau(P) = \sigma^m(P)$, σ^m the law of m free semi-circular found in Theorem 4. When V is not zero, such an argument does not hold a priori since the right hand side will also depend on $\tau(D_i q_j P)$, with $D_i q_j P$ of degree strictly larger than $X_i P$. However, our compactness assumption **(CS(R))** gives uniqueness because it forces the solution to be in a small neighborhood of the law $\tau_0 = \sigma^m$ of m free semi-circular variables, so that perturbation analysis applies. We shall see in Theorem 12 that this solution is actually the one which is related with the enumeration of maps.

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 $D - 1$,

$$\Delta_l(\tau, \tau') = \max_{P \text{ of degree } l-1} \max_{1 \leq i \leq m} \{|\tau(X_i P) - \tau'(X_i P)|\}$$

$$\leq 2\sum_{k=0}^{l-2}\Delta_k(\tau,\tau')R^{l-2-k} + C|t|\sum_{p=0}^{D-1}\Delta_{l+p-1}(\tau,\tau')$$

with a finite constant C (which depends on n only). For $\gamma > 0$, we set

$$d_\gamma(\tau,\tau') = \sum_{l\geq 0}\gamma^l\Delta_l(\tau,\tau').$$

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

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

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

$$\frac{1}{4R^2} + C|t|\sum_{p=0}^{D-1}(2R)^{p-1} < 1$$

so that when R is large, we see that we need $|t|$ to be at most of order $|R|^{-D+2}$. $\square$

4.3.3 Convergence of the empirical distribution

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

Theorem 11. *For all $c > 0$, there exists $\eta > 0$ and $M_0 \in \mathbb{R}^+$ (given in Lemma 28) so that for all $\mathbf{t} \in U_c \cap B_\eta$, $\hat{\mathbf{L}}^N$ (resp. $\overline{\mathbf{L}}_t^N$) converges almost surely (resp. everywhere) to the unique solution of* **SD[$V_\mathbf{t}$]** *such that*

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

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

Proof. By Property 81, the limit points of $\overline{\mathbf{L}}_t^N$ satisfy **CS(M_0)** and **SD[$V_\mathbf{t}$]**. Since M_0 does not depend on $\mathbf{t}$, we can apply Theorem 10 to see that if $\mathbf{t}$ is small enough, there is only one such limit point. Thus, by Corollary 78 we can conclude that $(\overline{\mathbf{L}}_t^N, N \in \mathbb{N})$ converges to this limit point. From Lemma 31, we have that

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

insuring by Borel-Cantelli's lemma that

$$\lim_{N\to\infty}(\hat{\mathbf{L}}^N - \overline{\mathbf{L}}_t^N)(P) = 0 \quad a.s$$

resulting with the almost sure convergence of $\hat{\mathbf{L}}^N$. $\square$

4.3.4 Combinatorial interpretation of the limit

In this part, we are going to identify the unique solution $\tau_{\mathbf{t}}$ of Theorem 10 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 $\mathbb{C}\langle X_1, \cdots, X_m\rangle$,

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

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

Theorem 12. *1. The family $\{\mathcal{M}_{\mathbf{k}}(P), \mathbf{k} \in \mathbb{N}^n, P \in \mathbb{C}\langle X_1, \cdots, X_m\rangle\}$ satisfies the induction relation: for all $i \in \{1, \cdots, m\}$, all $P \in \mathbb{C}\langle X_1, \cdots, X_m\rangle$, all $\mathbf{k} \in \mathbb{N}^n$,*

$$\mathcal{M}_{\mathbf{k}}(X_i P) = \sum_{\substack{0 \le p_j \le k_j \\ 1 \le j \le n}} \prod_{j=1}^n C_{k_j}^{p_j} \sum_{P = p_1 X_i p_2} \mathcal{M}_{\mathbf{p}}(P_1) \mathcal{M}_{\mathbf{k}-\mathbf{p}}(P_2) + \sum_{1 \le j \le n} k_j \mathcal{M}_{\mathbf{k}-1_j}([D_i q_j]P) \tag{16}$$

where $1_j(i) = 1_{i=j}$ and $\mathcal{M}_{\mathbf{k}}(1) = 1_{\bar{k}=0}$. (16) defines uniquely the family $\{\mathcal{M}_{\mathbf{k}}(P), \mathbf{k} \in \mathbb{N}^n, P \in \mathbb{C}\langle X_1, \cdots, X_m\rangle\}$.

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

$$|\mathcal{M}_{\mathbf{k}}(P)| \le \mathbf{k}! A^{\sum_{i=1}^n k_i} B^{deg(P)} \prod_{i=1}^n C_{k_i} C_{deg(P)} \tag{17}$$

with $\mathbf{k}! := \prod_{i=1}^n k_i!$ and C_p the Catalan numbers.

3. For $\mathbf{t}$ in $B_{(4A)^{-1}}$,

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

is absolutely convergent. For $\mathbf{t}$ small enough, $\mathcal{M}_{\mathbf{t}}$ is the unique solution of $\mathbf{SD}[V_{\mathbf{t}}]$ *which satisfies* **CS(4B)**.

By Theorem 10 and Theorem 11, we therefore readily obtain that

Corollary 82. *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 to*

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

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

$$\hat{\mathbf{L}}^N(PP^*) \ge 0 \quad \text{and} \quad \hat{\mathbf{L}}^N(PQ) = \hat{\mathbf{L}}^N(QP).$$

These conditions are closed for the weak topology and hence we find that

Corollary 83. *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 of any other way (and in particular any combinatorial way) to prove the positivity property $\mathcal{M}_\mathbf{t}(PP^*) \geq 0$ for all polynomial P, 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 $\mathcal{M}_\mathbf{t}$), see section **??**.

Proof of Theorem 12.

1. *Proof of the induction relation* (16).

   - We first check them for $\mathbf{k} = \mathbf{0} = (0, \cdots, 0)$. By convention, there is only one planar map with no vertex, so $\mathcal{M}_\mathbf{0}(1) = 1$. We now check that

     $$\mathcal{M}_\mathbf{0}(X_i P) = \mathcal{M}_\mathbf{0} \otimes \mathcal{M}_\mathbf{0}(\partial_i P) = \sum_{P = p_1 X_i p_2} \mathcal{M}_\mathbf{0}(p_1)\mathcal{M}_\mathbf{0}(p_2)$$

     But this is clear from (7) since for any planar map with only one star of type $X_i P$, the half-edge corresponding to X_i has to be glued with another half-edge of P, hence the event $A_M(X_i P)$ must hold, and if X_i is glued with the half-edge X_i coming from the decomposition $P = p_1 X_i p_2$, the map is split into two (independent) planar maps with stars p_1 and p_2 respectively (note here that p_1 and p_2 inherites the structure of stars since they inherite 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 the $\mathbf{k}$'s and the degree of P; we assume that (16) 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 $R = X_i P$ with P of degree less than L and the set of planar maps with a star of type $X_i Q$ 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 $P = p_1 X_i p_2$; we then can use (7) to see that this cuts the map M 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 labelled, 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 $X_i Q$ and k_i stars of type q_i, such that the marked half-edge of $X_i P$ is glued with an half-edge of P is

$$\sum_{P=p_1X_ip_2}\sum_{\substack{0\le r_i\le k_i\\1\le i\le n}}\prod_{i=1}^{n}C_{k_i}^{r_i}\mathcal{M}_{\mathbf{r}}(p_1)\mathcal{M}_{\mathbf{k}-\mathbf{r}}(p_2) \tag{18}$$

$\diamond$ 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 $q_j = q_1X_iq_2$. Then, we can use (8) to see that once we are given this gluing of the two edges, we can replace X_iP and q_j by q_2q_1P.

We have k_j ways to choose the star of type q_j and the total number of such maps is

$$\sum_{q_j=q_1X_iq_2}k_j\mathcal{M}_{\mathbf{k}-1_j}(q_2q_1P)$$

Summing over j, we obtain by linearity of $\mathcal{M}_{\mathbf{k}}$

$$\sum_{j=1}^{n}k_j\mathcal{M}_{\mathbf{k}-1_j}([D_iq_j]P) \tag{19}$$

(18) and (19) give (16). Moreover, it is clear that (16) defines uniquely $\mathcal{M}_{\mathbf{k}}(P)$ by induction.

2. *Proof of* (17). To prove the second point, we proceed also by induction over $\mathbf{k}$ and the degree of P. First, for $\mathbf{k}=\mathbf{0}$, $\mathcal{M}_{\mathbf{0}}(P)$ is the number of colored maps with one star of type P which is smaller than the number of planar maps with one star of type $x^{\deg \mathrm{P}}$ since colors only add constraints. Hence, we have, with C_k the Catalan numbers,

$$\mathcal{M}_{\mathbf{k}}(P)\le C_{[\frac{\deg(P)}{2}]}\le C_{\deg(P)}$$

showing that the induction relation is fine with $A=B=1$ at this step. Hence, let us assume that (17) is true for $\sum k_i\le M$ and all polynomials, and $\sum k_i = M+1$ for polynomials of degree less than L. Since $\mathcal{M}_{\mathbf{k}}(1)=0$ for $\sum k_i\ge 1$ we can start this induction. Moreover, using (16), we get that, if we denote $\mathbf{k}! = \prod_{i=1}^{n}k_i!$,

$$\begin{aligned}\frac{\mathcal{M}_{\mathbf{k}}(X_iP)}{\mathbf{k}!} &= \sum_{\substack{0\le p_i\le k_i\\1\le j\le n}}\sum_{P=P_1X_iP_2}\frac{\mathcal{M}_{\mathbf{p}}(P_1)}{\mathbf{p}!}\frac{\mathcal{M}_{\mathbf{k}-\mathbf{p}}(P_2)}{(\mathbf{k}-\mathbf{p})!}\\ &\quad+\sum_{\substack{1\le j\le n\\k_j\neq 0}}\frac{\mathcal{M}_{\mathbf{k}-\mathbf{1_j}}((D_iq_jP)}{(\mathbf{k}-1_j)!}\end{aligned}$$

Hence, taking P of degree less or equal to L and using our induction hypothesis, we find that

$$\left|\frac{\mathcal{M}_{\mathbf{k}}(X_iP)}{\mathbf{k}!}\right| \leq \sum_{\substack{0\leq p_j\leq k_j\\1\leq j\leq n}}\sum_{P=P_1X_iP_2} A^{\sum k_i}B^{degP-1}\prod_{i=1}^{n}C_{p_j}C_{k_j-p_j}C_{degP_1}C_{degP_2}$$
$$+2\sum_{1\leq l\leq n}A^{\sum k_j-1}\prod_j C_{k_j}B^{degP+deqq_l-1}C_{degP+deqq_l-1}$$
$$\leq A^{\sum k_i}B^{degP+1}\prod_i C_{k_i}C_{degP+1}\left(\frac{4^n}{B^2}+2\frac{\sum_{1\leq j\leq n}B^{degq_j-2}4^{degq_j-2}}{A}\right)$$

where we used Lemma **??** in the last line. It is now sufficient to choose A and B such that

$$\frac{4^n}{B^2}+2\frac{\sum_{1\leq j\leq n}B^{degq_j-2}4^{degq_j-2}}{A}\leq 1$$

(for instance $B=2^{n+1}$ and $A=4nB^{D-2}4^{D-2}$ if D is the maximal degree of the q_j) to verify the induction hypothesis works for polynomials of all degrees (all L's).

3. *Properties of* $\mathcal{M}_{\mathbf{t}}$. From the previous coniderations, we can of course define $\mathcal{M}_{\mathbf{t}}$ and the series is absolutely convergent for $|\mathbf{t}|\leq(4A)^{-1}$ since $C_k\leq 4^k$. Hence $\mathcal{M}_{\mathbf{t}}(P)$ depends analytically on $\mathbf{t}\in B_{(4A)^{-1}}$. Moreover, for all monomial P,

$$|\mathcal{M}_{\mathbf{t}}(P)|\leq\sum_{k\in\mathbb{N}^n}\prod_{i=1}^{n}(4t_iA)^{k_i}(4B)^{degP}\leq\prod_{i=1}^{n}(1-4At_i)^{-1}(4B)^{degP}.$$

so that for small t, $\mathcal{M}_{\mathbf{t}}$ satisfies **CS(4B)**.

4. $\mathcal{M}_{\mathbf{t}}$ *satisfies* **SD[$V_{\mathbf{t}}$]**. This is derived by summing (16) written for all $\mathbf{k}$ and multiplied by the factor $\prod(t_i)^{k_i}/k_i!$. From this point and the previous one (note that B is independent from $\mathbf{t}$), we deduce from Theorem 10 that for sufficiently small $\mathbf{t}$, $\mathcal{M}_{\mathbf{t}}$ is the unique solution of **SD[$V_{\mathbf{t}}$]** which satisfies **CS(4B)**. □

4.3.5 Convergence of the free energy

Theorem 13. *Let $c>0$. Then, for η small enough, for all $\mathbf{t}\in B_\eta\cap U_c$, the free energy converges towards a generating function of the numbers of certain planar maps;*

$$\lim_{N\to\infty}\frac{1}{N^2}\log\frac{Z_N^{V_{\mathbf{t}}}}{Z_N^0}=\sum_{\mathbf{k}\in\mathbb{N}^n\backslash(0,..,0)}\prod_{1\leq i\leq n}\frac{(-t_i)^{k_i}}{k_i!}\mathcal{M}_{\mathbf{k}}.$$

Moreover, the limit depends analytically on $\mathbf{t}$ *in a neighborhood of the origin.*

Proof. We may assume without loss of generality that $c\in(0,1]$. For $\alpha\in[0,1]$, $V_{\alpha\mathbf{t}}$ is c-convex since

$$V_{\alpha\mathbf{t}} + \frac{1}{2}\sum_{i=1}^{m} X_i^2 = \alpha(V_{\mathbf{t}}(X_1,\cdots,X_m) + \frac{1-c}{2}\sum_{i=1}^{m} X_i^2) + \frac{(1-\alpha)(1-c)+c}{2}\sum_{i=1}^{m} X_i^2$$

where all terms are convex (as we assumed $c \le 1$) whereas the last one is c-convex. Set

$$F_N(\alpha) = \frac{1}{N^2}\log Z_N^{V_{\alpha\mathbf{t}}}.$$

Then, $\frac{1}{N^2}\log\frac{Z_N^{V_{\mathbf{t}}}}{Z_N^0} = F_N(1) - F_N(0)$. Moreover

$$\partial_\alpha F^N_{\gamma(\alpha)} = -\overline{\mathbf{L}}^N_{\alpha t}(V_{\mathbf{t}}). \tag{20}$$

By Theorem 11, we know that for all $\alpha \in [0,1]$ (since $V_{\alpha\mathbf{t}}$ is c-convex),

$$\lim_{N\to\infty} \overline{\mathbf{L}}^N_{\alpha t}(V_{\mathbf{t}}) = \tau_{\alpha\mathbf{t}}(V_{\mathbf{t}})$$

whereas by (12), we know that $\overline{\mathbf{L}}^N_{\alpha t}(V_{\mathbf{t}})$ stays uniformly bounded. Therefore, a simple use of dominated convergence theorem shows that

$$\lim_{N\to\infty}\frac{1}{N^2}\log\frac{Z_N^{V_{\mathbf{t}}}}{Z_N^0} = -\int_0^1 \tau_{\alpha\mathbf{t}}(V_{\mathbf{t}})d\alpha = -\sum_{i=1}^{n} t_i \int_0^1 \tau_{\alpha\mathbf{t}}(q_i)d\alpha. \tag{21}$$

Now, observe that by Corollary 82,

$$\begin{aligned}
\tau_{\mathbf{t}}(q_i) &= \sum_{\mathbf{k}\in\mathbb{N}^n}\prod_{1\le j\le n}\frac{(-t_j)^{k_j}}{k_j!}\mathcal{M}_{\mathbf{k}+1_i} \\
&= -\partial_{t_i}\sum_{\mathbf{k}\in\mathbb{N}^n\setminus\{0,\cdots,0\}}\prod_{1\le j\le n}\frac{(-t_j)^{k_j}}{k_j!}\mathcal{M}_{\mathbf{k}}
\end{aligned}$$

so that (21) implies that

$$\begin{aligned}
\lim_{N\to\infty}\frac{1}{N^2}\log\frac{Z_N^{V_{\mathbf{t}}}}{Z_N^0} &= -\int_0^1 \partial_\alpha[\sum_{\mathbf{k}\in\mathbb{N}^n\setminus\{0,\cdots,0\}}\prod_{1\le j\le n}\frac{(-\alpha t_j)^{k_j}}{k_j!}\mathcal{M}_{\mathbf{k}}]d\alpha \\
&= -\sum_{\mathbf{k}\in\mathbb{N}^n\setminus\{0,\cdots,0\}}\prod_{1\le j\le n}\frac{(-t_j)^{k_j}}{k_j!}\mathcal{M}_{\mathbf{k}}.
\end{aligned}$$

$\square$

4.4 Second order expansion for the free energy

At the end of this chapter, we will have proved that Lemma 74 holds, up to the second order correction in the large N limit, i.e. that

$$\frac{1}{N^2}\log\left(\int e^{\sum_{i=1}^n t_i N \mathrm{tr}(q_i(X_1,\cdots,X_m))} d\mu^N(X_1)\cdots d\mu^N(X_m)\right)$$
$$=\sum_{g=0}^{1}\frac{1}{N^{2g-2}}\sum_{k_1,\cdots,k_n\in\mathbb{N}}\prod_{i=1}^{n}\frac{(t_i)^{k_i}}{k_i!}\mathcal{M}_g((q_i,k_i),1\le i\le n)+o(\frac{1}{N^2})$$

for some parameters t_i small enough and such that $\sum t_i q_i$ is c-convex. As for the first order, we shall prove first a similar large N expansion for $\overline{\mathbf{L}}_t^N$. We will first refine the arguments of the proof of Theorem 10 to estimate $\overline{\mathbf{L}}_t^N-\tau_{\mathbf{t}}$. This will already prove that $(\overline{\mathbf{L}}_t^N-\tau_{\mathbf{t}})(P)$ is at most of order N^{-2} for any polynomial P. To get the limit of $N^2(\overline{\mathbf{L}}_t^N-\tau_{\mathbf{t}})(P)$, we will first obtain a rough bound and then use the second order loop equation as we did for β-ensembles. The key argument in our approach, besides further uses of integration by parts-like arguments, will be the inversion of a differential operator acting on non-commutative polynomials which can be thought as a non-commutative analogue of a Laplacian operator with a drift.

We shall now estimate differences of $\hat{\mathbf{L}}^N$ and its limit . So, we set

$$\begin{aligned}
\hat{\delta}_{\mathbf{t}}^N &= N(\hat{\mathbf{L}}^N-\tau_{\mathbf{t}})\\
\bar{\delta}^N &= \int \hat{\delta}^N d\mu_V^N = N(\overline{\mathbf{L}}^N{}_{\mathbf{t}}-\tau_{\mathbf{t}})\\
\tilde{\delta}_{\mathbf{t}}^N &= N(\hat{\mathbf{L}}^N-\overline{\mathbf{L}}^N{}_{\mathbf{t}}) = \hat{\delta}_{\mathbf{t}}^N-\bar{\delta}^N.
\end{aligned}$$

In order to simplify the notations, we will make $\mathbf{t}$ implicit and drop the subscript $\mathbf{t}$ in the rest of this chapter so that we will denote $\overline{\mathbf{L}}^N,\tau,\hat{\delta}^N,\bar{\delta}_b^N$ and $\tilde{\delta}^N$ in place of $\overline{\mathbf{L}}^N{}_{\mathbf{t}},\tau_{\mathbf{t}},\hat{\delta}_{\mathbf{t}}^N,\bar{\delta}^N$ and $\tilde{\delta}_{\mathbf{t}}^N$, as well as V in place of $V_{\mathbf{t}}$.

4.4.1 Rough estimates on the size of the correction $\tilde{\delta}_{\mathbf{t}}^N$

In this section we improve on the perturbation analysis performed in section 4.3.2 in order to get the order of

$$\bar{\delta}_b^N(P)=N(\overline{\mathbf{L}}^N(P)-\tau)(P)$$

for all monomial P.

Proposition 84. *For all $c>0,\varepsilon\in]0,\frac{1}{2}[$, there exists $\eta>0,C<+\infty$, such that for all integer number N, all $\mathbf{t}\in B_\eta\cap U_c$, and all monomial function P of degree less than $N^{\frac{1}{2}-\varepsilon}$,*

$$|\bar{\delta}_b^N(P)|\le\frac{C^{deg\ (P)}}{N}.$$

Proof. The starting point is the finite dimensional Schwinger-Dyson equation of Property 80

$$\mu_V^N(\hat{\mathbf{L}}^N[(X_i+D_iV)P]) = \mu_V^N\left(\hat{\mathbf{L}}^N\otimes\hat{\mathbf{L}}^N(\partial_i P)\right) \tag{22}$$

Therefore, since τ satisfies the Schwinger-Dyson equation **SD[V]**, we get that for all polynomial P,

$$\overline{\delta}_b^N(X_iP) = -\overline{\delta}_b^N(D_iVP)+\overline{\delta}_b^N\otimes\overline{\mathbf{L}}^N(\partial_iP)+\tau\otimes\overline{\delta}_b^N(\partial_iP)+r(N,P) \tag{23}$$

with

$$r(N,P):=N^{-1}\mu_V^N\left(\tilde{\delta}^N\otimes\tilde{\delta}^N(\partial_iP)\right).$$

We take P a monomial of degree $d\le N^{\frac{1}{2}-\varepsilon}$ and see that

$$\begin{aligned}
|r(N,P)| &\le \frac{1}{N}\sum_{P=P_1X_iP_2}\mu_V^N\left(|\tilde{\delta}^N(P_1)|^2\right)^{\frac{1}{2}}\mu_V^N\left(|\tilde{\delta}^N(P_2)|^2\right)^{\frac{1}{2}}\\
&\le \frac{C}{N}\sum_{l=0}^{d-1}(Bl^2M^{2(l-1)}+C^lN^4e^{-\frac{\alpha MN}{2}})^{\frac{1}{2}}\times\\
&\qquad (B(d-l-1)^2M^{2(d-l-1)}+C^{(d-l-1)}N^4e^{-\frac{\alpha MN}{2}})^{\frac{1}{2}}\\
&\le \frac{C}{N}d(B(d-1)^2M^{2(d-2)}+C^{(d-1)}N^4e^{-\frac{\alpha MN}{2}}):=r(N,d,M)
\end{aligned}$$

where we used in the second line Lemma 31 and assumed $M\ge M_0$, and $d\le N^{\frac{1}{2}-\varepsilon}$. We set

$$\Delta_d^N:=\max_{P\text{ monomial of degree }d}|\overline{\delta}_b^N(P)|.$$

Observe that by (37), for any monomial of degree d less than $N^{\frac{1}{2}-\varepsilon}$,

$$|\overline{\mathbf{L}}_t^N(P)|\le C(\varepsilon)^d,\quad |\tau(P)|\le C_0^d\le C(\varepsilon)^d.$$

Thus, by (23), writing $D_iV=\sum t_jD_iq_j$, we get that for $d<N^{\frac{1}{2}-\varepsilon}$

$$\Delta_{d+1}^N\le\max_{1\le i\le m}\sum_{j=1}^n|t_j|\Delta^N_{d+\deg(D_iq_j)}+2\sum_{l=0}^{d-1}C(\varepsilon)^{d-l-1}\Delta_l^N+r(N,d,M).$$

We next define for $\kappa\le1$

$$\Delta^N(\kappa,\varepsilon)=\sum_{k=1}^{N^{\frac{1}{2}-\varepsilon}}\kappa^k\Delta_k^N.$$

We obtain, if D is the maximal degree of V,

$$\begin{aligned}
\Delta^N(\kappa,\varepsilon) &\le [C'|\mathbf{t}|+2(1-C(\varepsilon)\kappa)^{-1}\kappa^2]\Delta^N(\kappa,\varepsilon)\\
&\qquad +C|\mathbf{t}|\sum_{k=N^{\frac{1}{2}-\varepsilon}+1}^{N^{\frac{1}{2}-\varepsilon}+D}\kappa^{k-D}\Delta_k^N+\sum_{k=1}^{N^{\frac{1}{2}-\varepsilon}}\kappa^kr(N,k,M)
\end{aligned} \tag{24}$$

where we choose κ small enough so that $C(\varepsilon)\kappa < 1$. Moreover, since D is finite, bounding Δ_k^N by $2NC(\varepsilon)^k$, we get

$$\sum_{k=N^{\frac{1}{2}-\varepsilon}+1}^{N^{\frac{1}{2}-\varepsilon}+D} \kappa^{k-D}\Delta_k^N \leq 2DN(\kappa C(\varepsilon))^{N^{\frac{1}{2}-\varepsilon}}\kappa^{-D}.$$

When $\kappa C(\varepsilon) < 1$, as N goes to infinity, this term is neglectable with respect to N^{-1} for all $\varepsilon > 0$. The following estimate holds according to Lemma 31

$$\sum_{k=1}^{N^{\frac{1}{2}-\varepsilon}} \kappa^k r(N,k,M) \leq \frac{C}{N}\sum_{k=1}^{N^{\frac{1}{2}-\varepsilon}} k\kappa^k (B(k-1)^2 M^{2(k-2)} + C^{(k-1)}N^4 e^{-\frac{\alpha NM}{2}}) \leq \frac{C''}{N}$$

if κ is small enough so that $M^2\kappa < 1$ and $C\kappa < 1$. We observed here that $N^4 e^{-\frac{\alpha NM}{2}}$ is uniformly bounded independently of $N \in \mathbb{N}$. Now, if $|\mathbf{t}|$ is small, we can choose κ so that

$$\zeta := 1 - [C'|\mathbf{t}| + 2(1 - C(\varepsilon)\kappa)^{-1}\kappa^2] > 0.$$

Plugging these controls into (24) shows that for all $\varepsilon > 0$, and for $\kappa > 0$ small enough, there exists a finite constant $C(\kappa,\varepsilon)$ so that

$$\Delta^N(\kappa,\varepsilon) \leq C(\kappa,\varepsilon)N^{-1}$$

and so for all monomial P of degree $d \leq N^{\frac{1}{2}-\varepsilon}$,

$$|\overline{\delta}_b^N(P)| \leq C(\kappa,\varepsilon)\kappa^{-d}N^{-1}.$$

$\square$

To get the precise evaluation of $N\overline{\delta}_b^N(P)$, we shall first obtain a central limit theorem under μ_V^N which in turn will allow us to estimate the limit of $Nr(N,P)$.

4.4.2 Second order Schwinger-Dyson (or loop) equations

In this section we derive second order loop equation by making a small change in the potential $V \to V + N^{-1}\varepsilon W$ and identifying the first term in ε. @@ We denote by

$$\begin{aligned} W_2^V(P,Q) &:= \mathbb{E}[(\mathrm{tr}P - \mathbb{E}[\mathrm{tr}P])(\mathrm{tr}Q - \mathbb{E}[\mathrm{tr}Q])] \\ &:= \partial_t \mu^N_{V-N^{-1}tQ}[\mathrm{tr}P]|_{t=0} \\ W_3^V(P,Q,R) &= \partial_t W_2^{V-N^{-1}tR}(P,Q)|_{t=0} \end{aligned}$$

and forget the V in this notation when not needed. Differentiating (22) we deduce that

$$W_2((X_i + D_iV)P,Q) = (W_2 \otimes \tau_V + \tau_V \otimes W_2)(\partial_i P,Q) \tag{25}$$

$$+W_2(PDV,Q) + \tau_V(PDQ) \tag{26}$$

$$+N^{-1}\overline{\delta}_b^N(PDQ) + (W_2 \otimes \overline{\delta}_b^N + \overline{\delta}_b^N \otimes W_2)(\partial_i P,Q) + N^{-1}W_3(\partial_i P,Q)$$

By the previous rough estimates as well as the control on W_3 the last line is at most of order N^{-1} and therefore does not play any role in the limit of W_2. Observe therefore that at list when $V = 0$ the above equation allows to define the asymptotics of W_2 by induction over the degree of the first polynomial. We shall formalize this remark to include the ase where V is small.

To this end we shall take in (**??**) $P = D_i Q$ and sum the resulting equalities to find that the terms in W_2 sum up as $W_2(\Xi P, Q)$ with an operator Ξ that we shall invert. The resulting operator is a differential operator. As such, it may be difficult to find a normed space stable for this operator (since the operator will deteriorate the smoothness of the functions) in which it is continuous and invertible.

To avoid this issue, we will first divide each monomials of P by its degree (which more or less amounts to integrate and then divide by x the function in the one variable case).

Then, we define a linear map Σ on $\mathbb{C}\langle X_1, \cdots, X_m\rangle$ such that for all monomials q of degree greater or equal to 1

$$\Sigma q = \frac{q}{\deg q}.$$

Moreover, $\Sigma(q) = 0$ if $\deg q = 0$. For later use, we set $\mathbb{C}\langle X_1, \cdots, X_m\rangle(0)$ to be the subset of polynomials P of $\mathbb{C}\langle X_1, \cdots, X_m\rangle_{sa}$ such that $P(0, \cdots, 0) = 0$. We let Π be the projection from $\mathbb{C}\langle X_1, \cdots, X_m\rangle_{sa}$ onto $\mathbb{C}\langle X_1, \cdots, X_m\rangle(0)$ (i.e $\Pi(P) = P - P(0, \cdots, 0)$). We now define some operators on $\mathbb{C}\langle X_1, \cdots, X_m\rangle(0)$ i.e. from $\mathbb{C}\langle X_1, \cdots, X_m\rangle(0)$ into $\mathbb{C}\langle X_1, \cdots, X_m\rangle(0)$,

$$\Xi_1 : P \longrightarrow \Pi\left(\sum_{k=1}^{m} \partial_k \Sigma P \sharp D_k V\right)$$

$$\Xi_2 : P \longrightarrow \Pi\left(\sum_{k=1}^{m} (\mu \otimes I + I \otimes \mu)(\partial_k D_k \Sigma P)\right).$$

We denote $\Xi_0 = I - \Xi_2$ and $\Xi = \Xi_0 + \Xi_1$, where I is the identity on $\mathbb{C}\langle X_1, \cdots, X_m\rangle(0)$. Note that the images Ξ_i's and Ξ are indeed included in $\mathbb{C}\langle X_1, \cdots, X_m\rangle_{sa}$ since V is assumed self-adjoint. With these notations, Lemma 87, once applied to $P_i = D_i \Sigma P$, $1 \leq i \leq m$, reads

Proposition 85. *For all P in $\mathbb{C}\langle X_1, \cdots, X_m\rangle(0)$, $\hat{\delta}^N(\Xi P)$ converges in law to a centered Gaussian variable with covariance*

$$\mathcal{C}(P) := C(D_1 \Sigma P, \cdots, D_m \Sigma P).$$

Proof.

We have for all tracial state τ, $\tau(\partial_k P \sharp V) = \tau(D_k P V)$ and if P is in $\mathbb{C}\langle X_1, \cdots, X_m\rangle(0)$ (i.e $P(0, \cdots, 0) = 0$), we have the identity

$$P = \sum_k \partial_k \Sigma P \sharp X_k.$$

Then, as $\hat{\delta}^N$ is tracial (6) and vanishes on constant terms (so that the projection Π can be removed in the definition of Ξ), for all polynomial P,

$$\begin{aligned}
\hat{\delta}^N(\Xi P) &= \hat{\delta}^N(P+\sum_{k=1}^{m}\partial_k \Sigma P \sharp D_k V - \sum_{k=1}^{m}(\mu\otimes I + I\otimes\mu)(\partial_k D_k \Sigma P)) \\
&= \hat{\delta}^N(\sum_{k=1}^{m}(X_k + D_k V) D_k \Sigma P - \sum_{k=1}^{m}(\mu\otimes I + I\otimes\mu)(\partial_k D_k \Sigma P)) \\
&= Z_N(D_1\Sigma P, \cdots, D_m \Sigma P).
\end{aligned}$$

We then use Lemma 87 to conclude. □

To generalize the central limit theorem to all polynomial functions, we need to show that the image of Ξ is dense and to control approximations. If P is a polynomial and q a non-constant monomial we will denote $\lambda_q(P)$ the coefficient of q in the decomposition of P in monomials. We can then define a norm $\|.\|_A$ on $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)$ for $A>1$ by

$$\|P\|_A = \sum_{\deg q \neq 0} |\lambda_q(P)| A^{\deg q}.$$

In the formula above, the sum is taken over all non-constant monomials. We also define the operator norm given, for T from $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)$ to $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)$, by

$$|||T|||_A = \sup_{\|P\|_A=1} \|T(P)\|_A.$$

Finally, let $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)_A$ be the completion of $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)$ for $\|.\|_A$. We say that T is continuous on $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)_A$ if $|||T|||_A$ is finite. We shall prove that Ξ is continuous on $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)_A$ with continuous inverse when $\mathbf{t}$ is small.

Lemma 86. With the previous notations,

1. The operator Ξ_0 is invertible on $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)$.
2. There exists $A_0>0$ such that for all $A>A_0$, the operators Ξ_2, Ξ_0 and Ξ_0^{-1} are continuous on $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)_A$ and their norm are uniformly bounded for $\mathbf{t}$ in B_η.
3. For all $\varepsilon, A>0$, there exists $\eta_\varepsilon>0$ such for $|\mathbf{t}|<\eta_\varepsilon$, Ξ_1 is continuous on $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)_A$ and $|||\Xi_1|||_A \leq \varepsilon$.
4. For all $A>A_0$, there exists $\eta>0$ such that for $\mathbf{t}\in B_\eta$, Ξ is continuous, invertible with a continuous inverse on $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)_A$. Besides the norms of Ξ and Ξ^{-1} are uniformly bounded for $\mathbf{t}$ in B_η.
5. There exists $C>0$ such that for all $A>C$, $\mathcal{C}$ is continuous from $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)_A$ into $\mathbb{R}$.

Proof.

1. Observe that since Ξ_2 reduces the degree of a polynomial by at least 2,
$$P \to \sum_{n\geq 0} (\Xi_2)^n(P)$$
is well defined on $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)$ as the sum is finite for any polynomial P. This gives an inverse for $\Xi_0 = I - \Xi_2$.

2. First remark that a linear operator T has a norm less than C with respect to $\|.\|_A$ if and only if for all non-constant monomial q,
$$\|T(q)\|_A \leq CA^{\deg q}.$$
Recall that μ is uniformly compactly supported (see Lemma 27) and let $C_0 < +\infty$ be such that $|\mu(q)| \leq C_0^{\deg q}$ for all monomial q. Take a monomial $q = X_{i_1}\cdots X_{i_p}$, and assume that $A > 2C_0$,
$$\begin{aligned} &\|\Pi\left(\sum_k (I\otimes\mu)\partial_k D_k \Sigma q\right)\|_A \leq p^{-1} \sum_{\substack{k,q=q_1X_kq_2,\\ q_2q_1=r_1X_kr_2}} \|r_1\mu(r_2)\|_A \\ &\leq p^{-1} \sum_{\substack{k,q=q_1X_kq_2,\\ q_2q_1=r_1X_kr_2}} A^{\deg r_1} C_0^{\deg r_2} = \frac{1}{p}\sum_{n=0}^{p-1}\sum_{l=0}^{p-2} A^l C_0^{p-l-2} \\ &\leq A^{p-2}\sum_{l=0}^{p-2}\left(\frac{C_0}{A}\right)^{p-2-l} \leq 2A^{-2}\|q\|_A \end{aligned}$$
where in the second line, we observed that once $\deg(q_1)$ is fixed, q_2q_1 is uniquely determined and then r_1, r_2 are uniquely determined by the choice of l the degree of r_1. Thus, the factor $\frac{1}{p}$ is compensated by the number of possible decomposition of q i.e. the choice of n the degree of q_1. If $A > 2$, $P \to \Pi\left(\sum_k (I\otimes\mu)\partial_k D_k \Sigma P\right)$ is continuous of norm strictly less than $\frac{1}{2}$. And a similar calculus for $\Pi\left(\sum_k (\mu\otimes I)\partial_k D_k \Sigma\right)$ shows that Ξ_2 is continuous of norm strictly less than 1. It follows immediately that Ξ_0 is continuous. Since $\Xi_0^{-1} = \sum_{n\geq 0} \Xi_2^n$, Ξ_0^{-1} is continuous as soon as Ξ_2 is of norm strictly less than 1.

3. Let $q = X_{i_1}\cdots X_{i_p}$ be a monomial and let D be the degree of V and $B(\leq Dn)$ the sum of the maximum number of monomials in D_kV.
$$\begin{aligned} \|\Xi_1(q)\|_A &\leq \frac{1}{p}\sum_{k,q=q_1X_kq_2} \|q_1 D_k V q_2\|_A \leq \frac{1}{p}\sum_{k,q=q_1X_kq_2} |\mathbf{t}|BA^{p-1+D-1} \\ &= |\mathbf{t}|BA^{D-2}\|q\|_A. \end{aligned}$$
It is now sufficient to take $\eta_\varepsilon < (BA^{D-2})^{-1}\varepsilon$.

4. We choose $\eta < (BA^{D-2})^{-1}|||\Xi_0^{-1}|||_A^{-1}$ so that when $|\mathbf{t}| \leq \eta$,
$$|||\Xi_1|||_A|||\Xi_0^{-1}|||_A < 1.$$

By continuity, we can extend Ξ_0, Ξ_1, Ξ_2, Ξ and Ξ_0^{-1} on the space $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)_A$. The operator

$$P \to \sum_{n\geq 0}(-\Xi_0^{-1}\Xi_1)^n\Xi_0^{-1}$$

is well defined and continuous. This is an inverse of $\Xi = \Xi_0 + \Xi_1 = \Xi_0(I + \Xi_0^{-1}\Xi_1)$.

5. We finally prove that $\mathcal{C}$ is continuous from $\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)_A$ into $\mathbb{R}$ where we recall that we assumed $A > C_0$. Let us consider the first term

$$\mathcal{C}_1(P) := \sum_{k,l=1}^{m} \mu\otimes\mu(\partial_k D_l \Sigma P \times \partial_l D_k \Sigma P).$$

Then, we obtain as in the second point of this proof

$$\begin{aligned}
|\mathcal{C}_1(P)| &\leq 4\sum_{k,l=1}^{m}\sum_{q,q'}\frac{|\lambda_q(P)||\lambda_{q'}(P)|}{\deg q \deg q'}\sum_{\substack{q=q_1X_kq_2,q'=q'_1X_lq'_2\\ q_2q_1=r_1X_lr_2,q'_2q'_1=r'_1X_kr'_2}} C_0^{\deg q+\deg q'-4}\\
&\leq 4\sum_{q,q'}|\lambda_q(P)||\lambda_{q'}(P)|\deg q\deg q' C_0^{\deg q+\deg q'-4}\\
&\leq 4(\sup_{\ell\geq 0}\ell C_0^{\ell-2}A^{-\ell})^2\|P\|_A^2.
\end{aligned}$$

We next turn to show that

$$\mathcal{C}_2(P) := \sum_{k,l=1}^{m}\mu\left(\partial_k\circ\partial_l V\sharp(D_k\Sigma P, D_l\Sigma P)\right)$$

is also continuous for $\|.\|_A$. In fact, noting that we may assume $V\in\mathbb{C}\langle X_1,\cdots,X_m\rangle(0)$ without changing $\mathcal{C}_2$,

$$\begin{aligned}
|\mathcal{C}_2(P)| &\leq \sum_{p,q,q',k,l}|\lambda_p(V)|\sum_{\substack{q,q',p=p_1X_kp_2X_lp_3\\ q=q_1X_kq_2,q'=q'_1X_kq'_2}}\frac{|\lambda_q(P)||\lambda_{q'}(P)|C_0^{\deg p+\deg q+\deg q'-4}}{\deg q\deg q'}\\
&\leq n|\mathbf{t}|D^2\sum_{q,q'}|\lambda_q(P)||\lambda_{q'}(P)|C_0^{D+\deg q+\deg q'-4}\\
&\leq n|\mathbf{t}|D^2C_0^{D-4}\|P\|_A^2
\end{aligned}$$

The continuity of the last term $\mathcal{C}_3(P) = \sum_{i=1}^{m}\mu\left((D_j\Sigma P)^2\right)$ is obtained similarly.
□

Equation (**??**) gives

$$W_2(\Xi P,Q) = \sum_{i=1}^{p} \tau_V(\sum D_i \Sigma PQ) + \varepsilon_N(P)$$

It is not hard to verify that $\varepsilon_N(\Xi^1 P)$ goes to zero with N so that we conclude that

$$\lim_{N\to\infty} W_2(P,Q) = \sum_{i=1}^{p} \tau_V(\sum D_i \Sigma \Xi^{-1} PQ)$$

From the asymptotics of W_2 we can deduce those of $N\bar{\delta}_b^N(P)$. Indeed, coming back to (**??**) we find that

$$N\bar{\delta}_b^N(\Xi P) = W_2(\sum_{i=1}^{p} \partial_i D_i \Sigma P) + \varepsilon'_N(P)$$

We easily verify that $\varepsilon'_N(\Xi^{-1}P)$ goes to zero as N goes to infinity from which we readily deduce

$$\lim_{N\to\infty} N\bar{\delta}_b^N(P) = W_2(\sum_{i=1}^{p} \partial_i D_i \Sigma \Xi^{-1} P)$$

Chapter 5

Open questions or works in progress

5.1 One matrix model with several cuts

The topological expansions that appeared in these notes are only concerned with the case where the limiting spectral measure has a connected support. It is legitimate to wonder whether

(1) The large deviations for the probabilities that some eigenvalues lie outside this limiting support (should be easy)

(2) The topological expansion. The latter has more complicated expansions which contain θ functions. This is a work in progress with Gaetan Borot. A related central limit theorem was recently posted by M. Shcherbina on the arxiv.

5.2 Matrix model with complex potential or with domain of integration over complex curves

As the topological expansion are analytic it could be expected that they extend to the case where the potential is complex, at list when it is small. This is the subject of an article I have in preparation (unfortunately for quite a few years). In the one-matrix case, Bertola and Tovbis [19] could study the quartic case with a (non perturbative) complex weight. However the real case of interest is when one takes expectation on complex curves which are not perturbation of real curves. Indeed, it is now customary to be given a curve for the limit Stieltjes transform and a posteriori find and study the associated matrix model (cf some work by B. Eynard for instance). Such problems started to be studied by Bertola, cf [18].

5.3 Non perturbative several-matrix models

In non-perturbative cases it is not known in general whether the free energy converges. This is a famous open question related with the so-called free microstates entropy introduced by Voiculescu. With Dima Shlyktenkho, we could prove convergence and study the limit of such matrix model when the potential has some convexity property. In the one-matrix case this assumption implies that the support of the limiting measure is connected. An open question would be to obtain the next order asymptotics. For the sake of completness I copy below a summary of what we did with D. Shlyakhtenko to get the first order asymptotics, which is based on a nice use of dynamics (taken from my course at Park city).

Processes can be used to obtain non-perturbative results. In [33, 34, 21, 66], processes were the key to obtain large deviation estimates for Gaussian matrices. In [63], D. Shlyakhtenko and I used processes to show uniqueness of the solution to Schwinger-Dyson equation

$$\tau_V \otimes \tau_V(\partial_i P) = \tau_V((X_i + D_i V)P) \tag{1}$$

satisfying some bound such as $|\tau_V(X_{i_1} \cdots X_{i_p})| \leq R^p$ for all $i_j \in \{1, \cdots, m\}$. Indeed, this equation characterizes the invariant measures of Langevin dynamics driven by the free Brownian motion with drift given by the cyclic derivatives of the potential V. But when V is strictly convex (in the sense that its trace is a uniformly convex function of the entries of the matrices at which they are evaluated), these dynamics are shown to have only one equilibrium measure, and to converge exponentially fast to them. This insures the uniqueness of the solution to Schwinger Dyson equation and therefore the convergence of the empirical distribution as well as of the free energy.

5.4 More general laws on matrices

One can also wonder how to generalize the topological expansion to other settings. In [42], we considered unitary random matrices following the Haar measure on the unitary group and showed the convergence of the free energy in a perturbative regime. In [68] we could prove convergence of the free energy in the special case of the Harich-Chandra-Itzykson-Zuber integral

$$HCIZ(A,B) := \int e^{N\mathrm{Tr}(U^*AUB)} dU.$$

However the free energy is there given by a variational equation and is not related to a topological expansion. Moreover, it relies to the very special structure of this integral and in particular with its relation with the large deviation of the spectral measure of a Hermitian Brownian motion starting at B. We will not describe this result more precisely here but rather detail the results of [42] which generelize the strategy of the last chapter to the unitary matrices following the Haar measure.

The result is as follows. We consider matrix integrals given by

$$I_N(V, A_i^N) := \int e^{N\mathrm{Tr}(V(U_i, U_i^*, A_i^N, 1 \leq i \leq m))} dU_1 \cdots dU_m \tag{2}$$

where $(A_i^N, 1 \le i \le m)$ are $N \times N$ deterministic uniformly bounded matrices, dU denotes the Haar measure on the unitary group $\mathcal{U}(N)$ (normalized so that $\int_{\mathcal{U}(N)} dU = 1$) and V is a polynomial function in the non-commutative variables $(U_i, U_i^*, A_i^N, 1 \le i \le m)$. We assume that the joint distribution of the $(A_i^N, 1 \le i \le m)$ converges; namely for all polynomial function P in m non-commutative indeterminates

$$\lim_{N\to\infty} \frac{1}{N} \mathrm{Tr}(P(A_i^N, 1 \le i \le m)) = \tau(P) \tag{3}$$

for some linear functional τ on the set of polynomials. For technical reasons, we assume that the polynomial V satisfies
$\mathrm{Tr}(V(U_i, U_i^*, A_i^N, 1 \le i \le m)) \in \mathbb{R}$, for all $U_i \in \mathcal{U}(N)$, all Hermitian matrices A_i^N, for all $i \in \{1, \cdots, m\}$ and $N \in \mathbb{N}$.

Under those very general assumptions, the formal convergence of the integrals could already be deduced from [41]. The following Theorem is a precise decription of our results which gives an asymptotic convergence:

Theorem 87. *Under the above hypotheses and if we further assume that the spectral radius of the matrices $(A_i^N, 1 \le i \le m, N \in \mathbb{N})$ is uniformly bounded (by say M), there exists $\varepsilon = \varepsilon(M,V) > 0$ so that for $z \in [-\varepsilon, \varepsilon]$, the limit*

$$F_{V,\tau}(z) := \lim_{N\to\infty} \frac{1}{N^2} \log \int_{\mathcal{U}(N)^m} e^{zN\mathrm{Tr}(V(U_i,U_i^*,A_i^N,1\le i\le m))} dU_1 \cdots dU_m$$

exists. Moreover, $F_{V,\tau}(z)$ is an analytic function of $z \in \mathbb{C} \cap B(0,\varepsilon) = \{z \in \mathbb{C} : |z| \le \varepsilon\}$ and for all $k \in \mathbb{N}$,

$$\left.\frac{\partial^k}{\partial z^k} F_{V,\tau}(z)\right|_{z=0} = f_k(V,\tau).$$

Moreover, for all polynomial P there exists a limit

$$\tau_V(P) = \lim_{N\to\infty} I_N(V, A_i^N)^{-1} \int \frac{1}{N} \mathrm{Tr}(P(U_i, U_i^*, A_i^N, 1 \le i \le m)) e^{zN\mathrm{Tr}(V(U_i,U_i^*,A_i^N,1\le i\le m))} dU_1 \cdots dU_m$$

5.4.1 Idea of the proof

The strategy is again to find and study the Schwinger-Dyson (or loop) equationsunder the associated Gibbs measure

$$\mu_V^N(dU_1, \cdots, dU_m) = \frac{1}{Z^N} e^{zN\mathrm{Tr}(V(U_i,U_i^*,A_i^N,1\le i\le m))} dU_1 \cdots dU_m$$

To define this equation let us first define derivatives on polynomials in these matrices by the linear form such that

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

and satisfying the Leibnitz rule, namely, for monomials P, Q,

$$\partial_i(PQ) = \partial_i P \times (1 \otimes Q) + (P \otimes 1) \times \partial_i Q. \tag{4}$$

Here, $\times$ denotes the product $P_1 \otimes Q_1 \times P_2 \otimes Q_2 = P_1P_2 \otimes Q_1Q_2$. We also let D_i be the corresponding *cyclic* derivatives such that if $m(A \otimes B) = BA$, then $D_i = m \circ \partial_i$.

If q is a monomial, we more specifically have

$$\partial_i q = \sum_{q=q_1U_iq_2} q_1U_i \otimes q_2 - \sum_{q=q_1U_i^*q_2} q_1 \otimes U_i^* q_2 \tag{5}$$

$$D_i q = \sum_{q=q_1U_iq_2} q_2q_1U_i - \sum_{q=q_1U_i^*q_2} U_i^* q_2 q_1. \tag{6}$$

Then, using the invariance by multiplication of the Haar measure one can prove the asymptotic Schwinger-Dyson equation :

$$\mu_N^V(\frac{1}{N}\mathrm{Tr} \otimes \frac{1}{N}\mathrm{Tr}(\partial_i P)) + \mu_N^V(\frac{1}{N}\mathrm{Tr}(PD_iV)) = 0.$$

This is proved by noticing that if we set $U_p(t) = U_p e^{itB}$ and leave the other U_i's unchanged for a matrix $B = B^*$ then for all k, ℓ

$$\partial_t \int (P(U_i(t), U_i^*(t), A_i^N, 1 \le i \le m))_{k\ell} e^{zN\mathrm{Tr}(V(U_i(t), U_i^*(0), A_i^N, 1 \le i \le m))} dU_1 \cdots dU_m = 0$$

This reads

$$\int [(\partial_p P \sharp B)_{k\ell} + P_{k\ell} \mathrm{tr}(D_p V B)] d\mu_N^V = 0$$

Taking $B = 1_{k\ell} + 1_{\ell k}$ and $i(1_{k\ell} - 1_{\ell k})$ shows that we can by linearity choose $B = 1_{k\ell}$ even though this is not self-adjoint which yields the result after summation over k and ℓ.

We then claim that for all polynomial P,

$$\lim_{N\to\infty} \left\{ \frac{1}{N}\mathrm{Tr} \otimes \frac{1}{N}\mathrm{Tr}(\partial_i P) + \frac{1}{N}\mathrm{Tr}(D_iVP) \right\} = 0 \quad \mu_V^N \, a.s.$$

In particular, any limit point μ of $\hat{\mathbf{L}}^N$ under μ_V^N satisfies the Schwinger-Dyson equation

$$\mu \otimes \mu(\partial_i P) + \mu(D_iVP) = 0 \tag{7}$$

for all polynomial P and $\mu|_{(A_i)_{1\le i\le m}} = \tau$. Uniqueness of the solution to such an equation provides the convergence whereas the study of this solution shows that it expands as a generating series in the enumeration of some planar maps. In fact

5.4.2 discussion

The study of the second order (that is the fluctuations) around this limit is still an open questions. The strategy generalizes to the Haar measure on the orthogonal group (leading to the same limit except for a multiplicative constant) Finding loop equations for other classical ensembles could be a way to attack other integrals of interest in physics.

Chapter 6

Basics of matrices

6.1 Weyl's and Lidskii's inequalities

Theorem 14 (Weyl)**.** *Denote $\lambda_1(C) \leq \lambda_2(C) \leq \cdots \leq \lambda_N(C)$ the (real) eigenvalues of a $N \times N$ Hermitian matrix C. Let A, B be $N \times N$ Hermitian matrices. Then, for each $1 \leq j, k \leq N$ and $j + k \geq N + 1$,*

$$\lambda_{j+k-N}(A+B) \leq \lambda_j(A) + \lambda_k(B).$$

If $1 \leq j, k \leq N$ and $j + k \leq N + 1$,

$$\lambda_j(A) + \lambda_k(B) \leq \lambda_{j+k-1}(A+B).$$

Theorem 15 (Courant-Fischer)**.** *Let $A \in \mathcal{H}_N^{(2)}$ with ordered eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_N(A)$. For $k \in \{1, \cdots, N\}$,*

$$\lambda_k(A) := \min_{w_1, \cdots, w_{N-k} \in \mathbb{C}^N} max_{\substack{x \neq 0, x \in \mathbb{C}^N \\ x \perp w_1, \cdots, w_{N-k}}} \frac{x^* A x}{x^* x}.$$

Proof. We can without loss of generality assume that A is diagonal up to rotate the vectors $w_1, \cdots, w_{N-k}$ Then

$$\begin{aligned}
\max_{\substack{x \neq 0, x \in \mathbb{C}^N \\ x \perp w_1, \cdots, w_{N-k}}} \frac{x^* A x}{x^* x} &= \max_{\substack{\|x\|_2 = 1, x \in \mathbb{C}^N \\ x \perp w_1, \cdots, w_{N-k}}} \sum_{i=1}^{N} \lambda_i(A) |x_i|^2 \\
&\geq \max_{\substack{\|x\|_2 = 1, x \in \mathbb{C}^N, x_j = 0, j \leq k \\ x \perp w_1, \cdots, w_{N-k}}} \sum_{i=1}^{N} \lambda_i(A) |x_i|^2 \\
&\geq \lambda_k(A)
\end{aligned}$$

and equality holds when $w_i = u_{N-i+1}$ is the eigenvector corresponding to the eigenvalue $\lambda_{N-i+1}(A)$. Taking the minimum over the vectors w_i thus complete the proof. $\square$

We deduce that

Theorem 16 (Lidskii)**.** *Let $A \in \mathcal{H}_N^{(2)}$, $\eta \in \{+1,-1\}$ and $z \in \mathbb{C}^N$. We order the eigenvalues of $A+\eta zz^*$ in increasing order. Then*

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

Proof. Using Courant-Fischer theorem one gets for $k \ge 2$,

$$\begin{aligned}
\lambda_k(A+\eta zz^*) &:= \min_{w_1,\cdots,w_{N-k}\in\mathbb{C}^N} \max_{\substack{x\neq 0, x\in\mathbb{C}^N \\ x\perp w_1,\cdots,w_{N-k}}} \frac{x^*(A+\eta zz^*)x}{x^*x} \\
&\ge \min_{w_1,\cdots,w_{N-k}\in\mathbb{C}^N} \max_{\substack{x\neq 0, x\in\mathbb{C}^N \\ x\perp z, w_1,\cdots,w_{N-k}}} \frac{x^*Ax}{x^*x} \\
&\ge \min_{w_1,\cdots,w_{N-k+1}\in\mathbb{C}^N} \max_{\substack{x\neq 0, x\in\mathbb{C}^N \\ x\perp w_1,\cdots,w_{N-k+1}}} \frac{x^*Ax}{x^*x} \\
&= \lambda_{k-1}(A)
\end{aligned}$$

by a further use of Courant-Fischer theorem. Replacing $A' = A+\eta zz^*$, and η by $-\eta$ we also have proved $\lambda_k(A'-\eta zz^*) \ge \lambda_k(A')$, i.e $\lambda_k(A) \ge \lambda_{k-1}(A+\eta zz^*)$. ☐

Proof of Weyl's inequalities Theorem 20. We write $A_k = \sum_{i=k+1}^N \lambda_i(A)u_iu_i^*$ and $B_k = \sum_{i=k+1}^N \lambda_i(B)v_iv_i^*$ with $B_0 = B$, $A_0 = A$. Note that the fact that $\lambda_N(X+Y) \le \lambda_N(X) + \lambda_N(Y)$ is trivial (for instance from Courant-Fischer's theorem) and so for all j,k

$$\begin{aligned}
\lambda_N(A-A_j+B-B_k) &\le \lambda_N(A-A_j)+\lambda_N(B-B_k) \\
&= \lambda_j(A)+\lambda_k(B)
\end{aligned}$$

To bound from below $\lambda_N(A-A_j+B-B_k)$, note that the rank of A_j+B_k is at most $2N-j-k$. Applying Lidskii's theorem $2N-j-k$ times we thus get

$$\lambda_N(A-A_j+B-B_k) = \lambda_N((A+B)-(A_j+B_k)) \ge \lambda_{N-(2N-j-k)}(A+B) = \lambda_{j+k-N}(A+B)$$

which finishes the proof of the first inequality. The second is obtained by the change $A \to -A$ and $B \to -B$ which reverse the order of the eigenvalues too. The rank of A_j+B_k is at most $2N-j-k$. ☐

One also has

Theorem 17 (Lidskii)**.** *Let $A,E \in \mathcal{H}_N^{(2)}$. Then, there exists a doubly stochastic matrix B such that*

$$\lambda_k(A+E) - \lambda_k(A) = \sum_{m=1}^N B_{k,m}\lambda_m(E)\,. \tag{1}$$

In particular,

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

6.2 Non-commutative Hölder inequality

The following can be found in [85].

Theorem 18 (Nelson)**.** *For any* $P_1, \cdots, P_q \in \mathbb{C}\langle X_1, \cdots, X_m \rangle$*, any matrices* $\mathbf{A} = (A_1, \cdots, A_m)$ *in* $\mathcal{M}_N$ *any* $p_1, \cdots, p_q \in [0,1]^q$ *so that* $\sum p_i^{-1} = 1$*,*

$$|tr(P_1(\mathbf{A}) \cdots P_q(\mathbf{A}))| \leq \prod_{i=1}^{q} [tr(|P_i|^{q_i})]^{\frac{1}{q_i}}$$

This non-commutative Hölder inequality extends when tr is replaced by any Tracial states.

Chapter 7

Basics of Probability theory

7.1 Basic notions of large deviations

This appendix recalls basic definitions and main results of large deviations theory. We refer the reader to [48] and [47] for a full treatment.

In what follows, X will be assumed to be a Polish space (that is a complete separable metric space). We recall that a function $f : X \to \mathbb{R}$ is *lower semicontinuous* if the level sets $\{x : f(x) \leq C\}$ are closed for any constant C.

Definition 88. A sequence $(\mu_N)_{N\in\mathbb{N}}$ of probability measures on X satisfies a large deviation principle with speed a_N (going to infinity with N) and rate function I iff

$$I : X \to [0,\infty] \text{ is lower semicontinuous.} \tag{1}$$

$$\text{For any open set } O \subset X,\ \liminf_{N\to\infty} \frac{1}{a_N} \log \mu_N(O) \geq -\inf_O I. \tag{2}$$

$$\text{For any closed set } F \subset X,\ \limsup_{N\to\infty} \frac{1}{a_N} \log \mu_N(F) \leq -\inf_F I. \tag{3}$$

When it is clear from the context, we omit the reference to the speed or rate function and simply say that the sequence $\{\mu_N\}$ satisfies the LDP. Also, if x_N are X-valued random variables distributed according to μ_N, we say that the sequence $\{x_N\}$ satisfies the LDP if the sequence $\{\mu_N\}$ satisfies the LDP.

Definition 89. A sequence $(\mu_N)_{N\in\mathbb{N}}$ of probability measures on X satisfies a weak large deviation principle if (1) and (2) hold, and in addition (3) holds for all compact sets $F \subset X$.

The proof of a large deviation principle often proceeds first by the proof of a weak large deviation principle, in conjuction with the so-called exponential tightness property.

Definition 90. a. A sequence $(\mu_N)_{N\in\mathbb{N}}$ of probability measures on X is exponentially tight iff there exists a sequence $(K_L)_{L\in\mathbb{N}}$ of compact sets such that

$$\limsup_{L\to\infty}\limsup_{N\to\infty}\frac{1}{a_N}\log\mu_N(K_L^c) = -\infty.$$

b. A rate function I is good if the level sets $\{x\in X : I(x)\le M\}$ are compact for all $M\ge 0$.

The interest in these concepts lies in the

Theorem 19. *a. ([47, Lemma 1.2.18]) If $\{\mu_N\}$ satisfies the weak LDP and it is exponentially tight, then it satisfies the full LDP, and the rate function I is good.*
b. ([47, Exercise 4.1.10]) If $\{\mu_N\}$ satisfies the upper bound (3) with a good rate function I, then it is exponentially tight.

A weak large deviation principle is itself equivalent to the estimation of the probability of deviations towards small balls

Theorem 20. *[47, Theorem 4.1.11] Let $\mathcal{A}$ be a base of the topology of X. For every $A\in\mathcal{A}$, define*

$$\mathcal{L}_A = -\liminf_{N\to\infty}\frac{1}{a_N}\log\mu_N(A)$$

and

$$I(x) = \sup_{A\in\mathcal{A}:x\in A}\mathcal{L}_A.$$

Suppose that for all $x\in X$,

$$I(x) = \sup_{A\in\mathcal{A}:x\in A}\left\{-\limsup_{N\to\infty}\frac{1}{a_N}\log\mu_N(A)\right\}$$

Then, μ_N satisfies a weak large deviation principle with rate function I.

Let d be the metric in X, and set $B(x,\delta) = \{y\in X : d(y,x)<\delta\}$,

Corollary 91. *Assume that for all $x\in X$*

$$\begin{aligned} -I(x) &:= \limsup_{\delta\to 0}\limsup_{N\to\infty}\frac{1}{a_N}\log\mu_N(B(x,\delta)) \\ &= \liminf_{\delta\to 0}\liminf_{N\to\infty}\frac{1}{a_N}\log\mu_N(B(x,\delta)). \end{aligned}$$

Then, μ_N satisfies a weak large deviation principle with rate function I.

From a given large deviation principle one can deduce large deviation principle for other sequences of probability measures by using either the so-called contraction principle or Laplace's method.

Theorem 21 (Contraction Principle). *[47, Theorem 4.2.1] Assume that the sequence of probability measures $(\mu_N)_{N\in\mathbb{N}}$ on X satisfies a large deviation principle with good rate function I. Then, for any function $F : X \to Y$ with values in a Polish space Y which is continuous, the image $(F\sharp\mu_N)_{N\in\mathbb{N}} \in M_1(Y)^{\mathbb{N}}$ defined as $F\sharp\mu_N(A) = \mu \circ F^{-1}(A)$ also satisfies a large deviation principle with the same speed and rate function given for any $y \in Y$ by*

$$J(y) = \inf\{I(x) : F(x) = y\}.$$

Theorem 22 (Varadhan's Lemma). *[47, Theorem 4.3.1]): Assume that $(\mu_N)_{N\in\mathbb{N}}$ satisfies a large deviation principle with good rate function I. Let $F : X \to \mathbb{R}$ be a bounded continuous function. Then,*

$$\lim_{N\to\infty} \frac{1}{a_N} \log \int e^{a_N F(x)} d\mu_N(x) = \sup_{x\in X}\{F(x) - I(x)\}.$$

Moreover, the sequence

$$\nu_N(dx) = \frac{1}{\int e^{a_N F(y)} d\mu_N(y)} e^{a_N F(x)} d\mu_N(x) \in M_1(X)$$

satisfies a large deviation principle with good rate function

$$J(x) = I(x) - F(x) - \sup_{y\in X}\{F(y) - I(y)\}.$$

Large deviations principles are quite robust to exponential equivalence that we now define.

Definition 92. Let (X,d) be a metric space. Let $(\mu_N)_{N\in\mathbb{N}}$ and $(\tilde{\mu}_N)_{N\in\mathbb{N}}$ be two sequences of probability measures on X. $(\mu_N)_{N\in\mathbb{N}}$ and $(\tilde{\mu}_N)_{N\in\mathbb{N}}$ are said to be exponentially equivalent if there exists probability spaces $(\Omega, \mathcal{B}_N, P_N)$ and two families of random variables $Z_N, \tilde{Z}_N$ on Ω with values in X with joint distribution P_N and marginals μ_N and $\tilde{\mu}_N$ respectively so that for each $\delta > 0$

$$\limsup_{N\to\infty} P_N\left(d(Z_N, \tilde{Z}_N) > \delta\right) = -\infty.$$

We then have

Lemma 93. [47, Theorem 4.2.13] If a large deviation principle for μ_N holds with good rate function I and $\tilde{\mu}_N$ is exponentially equivalent to μ_N, then a $\tilde{\mu}_N$ satisfies a large deviation principle with the same rate function I.

$\mathcal{P}(\Sigma)$ possesses a useful criterion for compactness.

Theorem 23 (Prohorov). *Let Σ be Polish, and let $\Gamma \subset \mathcal{P}(\Sigma)$. Then $\overline{\Gamma}$ is compact iff Γ is tight.*

Since $\mathcal{P}(\Sigma)$ is Polish, convergence may be decided by sequences.

7.2 Basics of stochastic calculus

Definition 94 ([74], [90])**.** Let $(\Omega, \mathcal{F})$ be a measurable space.

- A filtration $\mathcal{F}_t, t \geq 0$ is a non-decreasing family of sub-σ-fields of $\mathcal{F}$.
- A random time T is a stopping time of the filtration $\mathcal{F}_t, t \geq 0$ if the event $\{T \leq t\}$ belongs to the σ-field $\mathcal{F}_t$ for all $t \geq 0$.
- A process $X_t, t \geq 0$ is adapted to the filtration $\mathcal{F}_t, t \geq 0$ if for all $t \geq 0$ X_t is an $\mathcal{F}_t$-measurable random variable.
- Let $\{X_t, \mathcal{F}_t, t \geq 0\}$ be an adapted process so that $\mathbb{E}[|X_t|] < \infty$ for all $t \geq 0$. The process $\{X_t, \mathcal{F}_t, t \geq 0\}$ is said to be a martingale if for every $0 \leq s < t < \infty$,
$$\mathbb{E}[X_t | \mathcal{F}_s] = X_s.$$
- Let $\{X_t, \mathcal{F}_t, t \geq 0\}$ be a martingale so that $E[X_t^2] < \infty$ for all $t \geq 0$. The martingale bracket (or the quadratic variation) $\langle X \rangle$ of X is the unique adapted increasing process so that $X^2 - \langle X \rangle$ is a martingale for the filtration $\mathcal{F}$

Let $\{X_t, \mathcal{F}_t, t \geq 0\}$ be a real-valued adapted process and let B be a Brownian motion. Assume that $E \int_0^T X_t^2 dt] < \infty$. Then,

$$\int_0^T X_t dB_t := \lim_{n \to \infty} \sum_{k=0}^{n-1} X_{\frac{Tk}{n}} (B_{\frac{T(k+1)}{n}} - B_{\frac{Tk}{n}})$$

exists, the convergence hold in L^2 and the limit does not depend on the above choice of the discretization of $[0, T]$ (see [74, section 3]). The limit is called a stochastic integral .

One can therefore consider the problem of finding solutions to the integral equation

$$X_t = X_0 + \int_0^t \sigma(X_s) dB_s + \int_0^t b(X_s) ds \tag{4}$$

with a given X_0, σ and b some functions on $\mathbb{R}^n$, and B a n-dimensional Brownian motion. This can be written under the differential form

$$dX_s = \sigma(X_s) dB_s + b(X_s) ds. \tag{5}$$

There are at least two notions of solutions; the strong solutions and the weak solutions.

Definition 95. [74, Definition 2.1] A strong solution of the stochastic differential equation (5) on the given probability space $(\Omega, \mathcal{F})$ and with respect to the fixed Brownian motion B and initial condition ξ is a process $\{X_t, t \geq 0\}$ with continuous sample paths so that

1. X is adapted to the filtration $\mathcal{F}$ given by

$\mathcal{G}_t = \sigma(B_s, s \leq t; X_0), \mathcal{N} = \{N \subset \Omega, \exists G \in \mathcal{G}_\infty \text{ with } N \subset G, P(G) = 0\}, \mathcal{F}_t = \sigma(\mathcal{G}_t \cup \mathcal{N}).$

2. $P(X_0 = \xi) = 1$.
3. $P(\int_0^t (|b_i(X_s)| + |\sigma_{ij}(X_s)|^2) ds < \infty) = 1$ for all $i, j \leq n$.
4. (4) holds almost surely.

Definition 96. [74, Definition 3.1] A weak solution of the stochastic differential equation (5) is a triple (X, B) and $(\Omega, \mathcal{F}, P)$ so that $(\Omega, \mathcal{F}, P)$ is a probability space equipped with a filtration $\mathcal{F}$, X is a continuous adapted process and B a n-dimensional Brownian motion. X satisfies (3) and (4) in Definition (116).

There are two also two notions of uniqueness;

Definition 97. [74, Definition 3.4]

- We say that strong uniqueness holds if two solutions with common probability space, common Brownian motion B and common initial condition are almost surely equal at all times.
- We say that weak uniqueness, or uniqueness in the sense of probability law, holds if any two weak solutions have the same law.

Theorem 24. *[74, Theorems 2.5 and 2.9]*
Suppose that b and σ satisfy

$$\begin{aligned} \|b(t,x) - b(t,y)\| + \|\sigma(t,x) - \sigma(t,y)\| &\leq K\|x-y\|, \\ \|b(t,x)\|^2 + \|\sigma(t,x)\|^2 \leq K^2(1 + \|x\|^2), \end{aligned}$$

for some finite constant K independent of t and $\|.\|$ the Euclidean norm on $\mathbb{R}^n$, then there exists a unique strong solution to (5). *Moreover, it satisfies*

$$\mathbb{E}[\int_0^T \|b(t, X_t)\|^2 dt] < \infty$$

for all $T \geq 0$.

Theorem 25. *[74, Proposition 3.10] Any two weak solutions $(X^i, B^i, \Omega^i, \mathcal{F}^i, P^i)_{i=1,2}$ of* (5) *so that*

$$\mathbb{E}[\int_0^T \|b(t, X_t^i)\|^2 dt] < \infty$$

for all $T < \infty$ and $i = 1, 2$ have the same law.

Theorem 26 (Itô(1944), Kunita-Watanabe (1967))**.** *[74, p. 149]*
Let $f : \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^2$ and let $X = \{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ be a continuous semi-martingale with decomposition

$$X_t = X_0 + M_t + A_t$$

where M is a local martingale and A the difference of continuous, adapted, non-decreasing processes. Then, almost-surely,

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)dM_s + \int_0^t f'(X_s)dA_s$$

$$+\frac{1}{2}\int_0^2 f''(X_s)d < M >_s, \quad 0 \le t < \infty.$$

We shall use the following well known results on martingales.

Theorem 27 (Burkholder-Davis-Gundy's inequality)**.** *[74, p. 166] Let $(M_t, t \ge 0)$ be a continuous local martingale with bracket $(A_t, t \ge 0)$. There exists universal constants λ_m, Λ_m so that for all $m \in \mathbb{N}$*

$$\lambda_m E(A_T^m) \le E(\sup_{t \le T} M_t^{2m}) \le \Lambda_m E(A_T^m).$$

Theorem 28 (Novikov(1972))**.** *[74, p. 199] Let $\{X_t, \mathcal{F}_t, t \ge 0\}$ be an adapted process with values in $\mathbb{R}^d$ such that*

$$E[e^{\frac{1}{2}\int_0^T \sum_{i=1}^d (X_t^i)^2 dt}] < \infty$$

for all $T \in \mathbb{R}^+$. Then, if $\{W_t, \mathcal{F}_t, t \ge 0\}$ is a d dimensional Brownian motion,

$$M_t = \exp\{\int_0^t X_u.dW_u - \frac{1}{2}\int_0^t \sum_{i=1}^d (X_u^i)^2 du\}$$

is a $\mathcal{F}_t$-martingale.

Theorem 29. *(Girsanov(1960)) [74, p. 191] Let $\{X_t, \mathcal{F}_t, t \ge 0\}$ be an adapted process with values in $\mathbb{R}^d$ such that*

$$E[e^{\frac{1}{2}\int_0^T \sum_{i=1}^d (X_t^i)^2 dt}] < \infty$$

Then, if $\{W_t, \mathcal{F}_t, P, 0 \le t \le T\}$ is a d dimensional Brownian motion,

$$\overline{W}_t^i = W_t^i - \int_0^t X_s^i ds, 0 \le t \le T$$

is a d dimensional Brownian under the probability measure

$$\overline{P} = \exp\{\int_0^T X_u.dW_u - \frac{1}{2}\int_0^T \sum_{i=1}^d (X_u^i)^2 du\}P.$$

Theorem 30. *[74, p. 14] Let $\{X_t, \mathcal{F}_t, 0 \le t < \infty\}$ be a submartingale whose every path is right-continuous. Then for any $\tau > 0$, for any $\lambda > 0$*

$$\lambda P(\sup_{0 \le t \le \tau} X_t \ge \lambda) \le E[X_\tau^+].$$

We shall use the following consequence

Corollary 98. *Let $\{X_t, \mathcal{F}_t, t \geq 0\}$ be an adapted process with values in $\mathbb{R}^d$ such that*

$$\int_0^T \|X_t\|^2 dt = \int_0^T \sum_{i=1}^d (X_t^i)^2 dt$$

is uniformly bounded by A_T. Let $\{W_t, \mathcal{F}_t, t \geq 0\}$ be a d dimensional Brownian motion. Then for any $L > 0$,

$$P(\sup_{0\leq t\leq T} |\int_0^t X_u.dW_u| \geq L) \leq 2e^{-\frac{L^2}{2A_T}}.$$

Proof. We denote in short $Y_t = \int_0^t X_u.dW_u$ and write for $\lambda > 0$,

$$\begin{aligned} P(\sup_{0\leq t\leq T} |Y_t| \geq A) &\leq P(\sup_{0\leq t\leq T} e^{\lambda Y_t} \geq e^{\lambda A}) + P(\sup_{0\leq t\leq T} e^{-\lambda Y_t} \geq e^{\lambda A}) \\ &\leq P\left(\sup_{0\leq t\leq T} e^{\lambda Y_t - \frac{\lambda^2}{2}\int_0^t \|X_u\|^2 du} \geq e^{\lambda A - \frac{\lambda^2 A_T}{2}}\right) \\ &\quad + P\left(\sup_{0\leq t\leq T} e^{-\lambda Y_t - \frac{\lambda^2}{2}\int_0^t \|X_u\|^2 du} \geq e^{\lambda A - \frac{\lambda^2 A_T}{2}}\right) \end{aligned}$$

By Theorem 34, $M_t = e^{-\lambda Y_t - \frac{\lambda^2}{2}\int_0^t \|X_u\|^2 du}$ is a non negative martingale. Thus, By Chebychev's inequality and Doob's inequality

$$\begin{aligned} P\left(\sup_{0\leq t\leq T} M_t \geq e^{\lambda A - \frac{\lambda^2 A_T}{2}}\right) &\leq e^{-\lambda A + \frac{\lambda^2 A_T}{2}} \mathbb{E}[M_T] \\ &= e^{-\lambda A + \frac{\lambda^2 A_T}{2}} \end{aligned}$$

Optimizing with respect to λ completes the proof.

□

Theorem 31 (Rebolledo's Theorem)**.** *Let $n \in \mathbb{N}$, and let M_N be a sequence of continuous centered martingales with values in $\mathbb{R}^n$ with bracket $< M_N >$ converging pointwise (i.e for all $t \geq 0$) in L^1 towards a continuous deterministic function $\varphi(t)$. Then, for any $T > 0$, $(M_N(t), t \in [0,T])$ converges in law as a continuous process from $[0,T]$ into $\mathbb{R}^n$ towards a Gaussian process G with covariance*

$$E[G(s)G(t)] = \varphi(t \wedge s).$$

7.3 Proof of (13)

Put

$$V(\mathbf{i}^1, \cdots, \mathbf{i}^l) = [[i_n^j]_{n=1}^k]_{j=1}^l, I = \bigcup_{j=1}^l \{j\} \times \{1, \ldots, k\}, A = [\{i_n^j, i_{n+1}^j\}]_{(i,n)\in I}.$$

We visualize A as a left-justified table of l rows. Let $G' = (V', E')$ be any spanning forest in $G(\mathbf{i}^1, \cdots, \mathbf{i}^l)$, with c connected components. Since every connected component of G' is a tree, we have

$$|V| = |V'| = c + |E'|. \tag{6}$$

Now let $X = \{X_{in}\}_{(i,n)\in I}$ be a table of the same "shape" as A, but with all entries equal either to 0 or 1. We call X an *edge-bounding table* under the following conditions:

- For all $(i,n) \in I$, if $X_{in} = 1$, then $A_{in} \in E'$.
- For each $e \in E'$ there exists distinct $(i_1, n_1), (i_2, n_2) \in I$ such that $X_{i_1 n_1} = X_{i_2 n_2} = 1$ and $A_{i_1 n_1} = A_{i_2 n_2} = e$.
- For each $e \in E'$ and index $i \in \{1, \ldots, j\}$, if e appears in the i^{th} row of A then there exists $(i,n) \in I$ such that $A_{in} = e$ and $X_{in} = 1$.

For any edge-bounding table X the corresponding quantity $\frac{1}{2}\sum_{(i,n)\in I} X_{in}$ bounds $|E'|$ by the second required property. At least one edge-bounding table exists, namely the table with a 1 in position (i,n) for each $(i,n) \in I$ such that $A_{in} \in E'$ and 0's elsewhere. Now let X be an edge-bounding table such that for some index i_0 all the entries of X in the i_0^{th} row are equal to 1. Then the graph $G(\mathbf{i}_0)$ is a tree (since all edges of $G(\mathbf{i}_0)$ could be kept in G'), and hence every entry in the i_0^{th} row of A appears there an even number of times and *a fortiori* at least twice. Now choose $(i_0, n_0) \in I$ such that $A_{i_0 n_0} \in E'$ appears in another row than i_0. Let Y be the table obtained by replacing the entry 1 of X in position (i_0, n_0) by the entry 0. Then Y is again an edge-bounding table. Proceeding in this way we can find an edge-bounding table with 0 appearing at least once in every row, and hence we have $|E'| \leq [\frac{|I|-l}{2}] = \frac{kl-l}{2}$. Together with (6) and the definition of I, this completes the proof. $\square$

Bibliography

[1] S. Aida and D. Stroock. Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. *Math. Res. Lett.*, 1(1):75–86, 1994.

[2] S. Albeverio, L. A. Pastur, and M. Shcherbina. On the $1/n$ expansion for some unitary invariant ensembles of random matrices. *Comm. Math. Physics*, 224:271–305, 2001.

[3] P. L. Albeverio S. and S. M. On the $1/n$ expansion for some unitary invariant ensembles of random matrices. *Comm. Math. Phys.*, 224(1):271–305, 2001. Dedicated to Joel L. Lebowitz.

[4] N. Alon, M. Krivelevich, and V. H. Vu. On the concentration of eigenvalues of random symmetric matrices. *Israel J. Math.*, 131:259–267, 2002.

[5] J. Ambjorn, Chekhov, L. Kristjansen, C.F., and Y. Makeenko. Matrix model calculations beyond the spherical limit. *Nuclear Physics B*, 404:127–172, 1993.

[6] J. Ambjørn, L. Chekhov, C. Kristjansen, and Y. Makeenko. Matrix model calculations beyond the spherical limit. *Nucl. Phys. B*, 449:681, 1995. [erratum] hep-th/9302014.

[7] J. Ambjørn, L. Chekhov, and Y. Makeenko. Higher genus correlators from the hermitian 1-matrix model. *Phys. Lett. B*, 282:341–348, 1992. hep-th/9203009.

[8] J. Ambjørn and Y. Makeenko. Properties of loop equations for the hermitian matrix model and for two-dimensional gravity. *Mod. Phys. Lett. A*, 5:1753–1763, 1990.

[9] G. Anderson, A. Guionnet, and O. Zeitouni. *Lectures notes on random matrices*.

[10] G. Anderson and O. Zetouni. A clt for a band matrix model. *Probab. Theory Related Fields*, To appear, 2005.

[11] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.

[12] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.

[13] O. M. B. Eynard, A-K Kashani-Poor. A matrix model for the topological string i: Deriving the matrix model. *arXiv:1003.1737*, 2010.

[14] C. Beenakker. Universality of Brézin and Zee's spectral correlator. *Nucl. Phys B*, 422:515–520, 1994. cond-mat/9310010.

[15] G. Ben Arous, A. Dembo, and A. Guionnet. Aging of spherical spin glasses. *Probab. Theory Related Fields*, 120(1):1–67, 2001.

[16] G. Ben Arous and A. Guionnet. Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Related Fields*, 108(4):517–542, 1997.

[17] G. Ben Arous and A. Guionnet. The spectrum of heavy tailed random matrices. *To appear in Comm. Math. Phys.*, 2007.

[18] M. Bertola. Boutroux curves with external field: equilibrium measures without a variational problem. *Anal. Math. Phys.*, 1:167–211., 2011.

[19] M. Bertola and A. Tovbis. Asymptotics of orthogonal polynomials with complex varying quartic weight:global structure, critical point behaviour and the first painlevé equation.

[20] P. Biane. Free probability for probabilists. In *Quantum probability communications, Vol. XI (Grenoble, 1998)*, QP-PQ, XI, pages 55–71. World Sci. Publishing, River Edge, NJ, 2003.

[21] P. Biane, M. Capitaine, and A. Guionnet. Large deviation bounds for matrix Brownian motion. *Invent. Math.*, 152(2):433–459, 2003.

[22] P. Biane and R. Speicher. Stochastic calculus with respect to free Brownian motion and analysis on Wigner space. *Probab. Theory Related Fields*, 112(3):373–409, 1998.

[23] P. Biane and R. Speicher. Free diffusions, free entropy and free Fisher information. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(5):581–606, 2001.

[24] P. Bleher and A. Its. Semiclassical asymptotics of orthogonal polynomials, Riemann-Hilbert problem, and universality in the matrix model. *Ann. Math.*, 150:185–266, 1999. math-ph/9907025.

[25] S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999.

[26] C. Bordenave, P. Caputo, and D. Chaifa. Spectrum of hermitian heavy tailed random matrices. *Comm. Math. Phys.*, To appear.

[27] G. Borot and B. Eynard. Enumeration of maps with self avoiding loops and the $O(\mathfrak{n})$ model on random lattices of all topologies. *J. Stat. Mech.*, (P01010), 2011. math-ph/0910.5896.

[28] G. Borot and A. Guionnet. Asymptotic expansion of beta matrix models in the one-cut regime. *arXiv:1107.1167*, 2011.

[29] A. Boutet de Monvel, A. Khorunzhy, and V. Vasilchuk. Limiting eigenvalue distribution of random matrices with correlated entries. *Markov Process. Related Fields*, 2(4):607–636, 1996.

[30] H. J. Brascamp and E. H. Lieb. 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):366–389, 1976.

[31] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.

[32] E. Brézin, C. Itzykson, G. Parisi, and J. B. Zuber. Planar diagrams. *Comm. Math. Phys.*, 59(1):35–51, 1978.

[33] T. Cabanal Duvillard and A. Guionnet. Large deviations upper bounds for the laws of matrix-valued processes and non-communicative entropies. *Ann. Probab.*, 29(3):1205–1261, 2001.

[34] T. Cabanal-Duvillard and A. Guionnet. Discussions around Voiculescu's free entropies. *Adv. Math.*, 174(2):167–226, 2003.

[35] L. A. Caffarelli. Boundary regularity of maps with convex potentials. *Comm. Pure Appl. Math.*, 45(9):1141–1151, 1992.

[36] L. A. Caffarelli. The regularity of mappings with a convex potential. *J. Amer. Math. Soc.*, 5(1):99–104, 1992.

[37] L. A. Caffarelli. Monotonicity properties of optimal transportation and the FKG and related inequalities. *Comm. Math. Phys.*, 214(3):547–563, 2000.

[38] S. Chatterjee. Concentration of Haar measures, with an application to random matrices. *J. Funct. Anal.*, 245(2):379–389, 2007.

[39] L. Chekhov and B. Eynard. Matrix eigenvalue model: Feynman graph technique for all genera. *JHEP*, (0612:026), 2006. math-ph/0604014.

[40] P. Cizeau and J.-P. Bouchaud. Theory of lévy matrices. *Physical Review E*, 50(3):1810–1822, 1994.

[41] B. Collins. Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability. *Int. Math. Res. Not.*, (17):953–982, 2003.

[42] B. Collins, A. Guionnet, and E. Maurel-Segala. Asymptotics of unitary and orthogonal matrix integrals. *Adv. Math.*, 222(1):172–215, 2009.

[43] A. B. de Monvel, L. Pastur, and M. Shcherbina. On the statistical mechanics approach in the random matrix theory. Integrated density of states. *J. Stat. Phys.*, 79(3-4):585–611, 1995.

[44] P. Deift, T. Kriecherbauer, K.-R. McLaughlin, S. Venakides, and X. Zhou. Strong asymptotics of orthogonal polynomials with respect to exponential weights via Riemann-Hilbert techniques. *Comm. Pure Appl. Math.*, 52(12):1491–1552, 1999.

[45] P. Deift, T. Kriecherbauer, K.-R. McLaughlin, S. Venakides, and X. Zhou. Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Comm. Pure Appl. Math.*, 52(11):1335–1425, 1999.

[46] P. Deift, T. Kriecherbauer, and K. T.-R. McLaughlin. New results on the equilibrium measure for logarithmic potentials in the presence of an external field. *J. Approx. Theory*, 95(3):388–475, 1998.

[47] A. Dembo and O. Zeitouni. Large deviations and applications. In *Handbook of stochastic analysis and applications*, volume 163 of *Statist. Textbooks Monogr.*, pages 361–416. Dekker, New York, 2002.

[48] J.-D. Deuschel and D. W. Stroock. *Large deviations*, volume 137 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1989.

[49] G. P. Di Francesco P. D. and Z.-J. J. 2d gravity and random matrices. *Phys. Rep.*, (254), 1995.

[50] F. J. Dyson. A Brownian-motion model for the eigenvalues of a random matrix. *J. Mathematical Phys.*, 3:1191–1198, 1962.

[51] N. M. Ercolani and K. D. T.-R. McLaughlin. Asymptotics of the partition function for random matrices via Riemann-Hilbert techniques and applications to graphical enumeration. *Int. Math. Res. Not.*, (14):755–820, 2003.

[52] B. Eynard. All genus correlation functions for the hermitian 1-matrix model. *JHEP*, (0411:031), 2004. hep-th/0407261.

[53] B. Eynard. Large N expansion of convergent matrix integrals, holomorphic anomalies, and background independence. *JHEP*, (0903:003), 2009. math-ph/0802.1788.

[54] B. Eynard. A matrix model for plane partitions. *arXiv:0905.0535*, 2009.

[55] B. Eynard and N. Orantin. Invariants of algebraic curves and topological expansion. 2007. math-ph/0702045.

[56] A. Fokas, A. Its, and A. Kitaev. The isomonodromy approach to matrix models in $2d$ quantum gravity. *Commun. Math. Phys.*, 147:395–430, 1992.

[57] P. T. Gross D. and W. S. Two dimensional quantum gravity and random surfaces. *Jerusalem winter school, World Scientific*, 1991.

[58] A. Guionnet. *Lectures on Random Matrices : Macroscopic asymptotics*. Lecture Notes for the 2006 Saint-Flour Probability Summer School. www.umpa.ens-lyon.fr/ aguionne, 2007.

[59] A. Guionnet and E. Maurel-Segala. Combinatorial aspects of matrix models. *arviv*, http://front.math.ucdavis.edu/math.PR/0503064, 2005.

[60] A. Guionnet and E. Maurel-Segala. Combinatorial aspects of matrix models. *ALEA Lat. Am. J. Probab. Math. Stat.*, 1:241–279 (electronic), 2006.

[61] A. Guionnet and E. Maurel-Segala. Second order asymptotics for matrix models. *Ann. Probab.*, 35:2160–2212, 2007. math.PR/0601040.

[62] A. Guionnet and E. Maurel-Segala. Second order asymptotics for matrix models. *Ann. Probab.*, 35(6):2160–2212, 2007.

[63] A. Guionnet and D. Shlyakhtenko. Free diffusions and matrix models with strictly convex interaction. *http://arxiv.org/abs/math/0701787*, 2007.

[64] A. Guionnet and B. Zegarlinski. Lectures on logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXVI*, volume 1801 of *Lecture Notes in Math.*, pages 1–134. Springer, Berlin, 2003.

[65] A. Guionnet and O. Zeitouni. Concentration of the spectral measure for large matrices. *Electron. Comm. Probab.*, 5:119–136 (electronic), 2000.

[66] A. Guionnet and O. Zeitouni. Large deviations asymptotics for spherical integrals. *J. Funct. Anal.*, 188(2):461–515, 2002.

[67] A. Guionnet and O. Zeitouni. Addendum to large deviations asymptotics for spherical integrals. *J. Funct. Anal.*, To appear, 2004.

[68] A. Guionnet and O. Zeitouni. Addendum to: "Large deviations asymptotics for spherical integrals" [J. Funct. Anal. **188** (2002), no. 2, 461–515; mr 1883414]. *J. Funct. Anal.*, 216(1):230–241, 2004.

[69] U. Haagerup, H. Schultz, and S. Thorbjørnsen. A random matrix approach to the lack of projections in $C^*_{\mathrm{red}}(\mathbb{F}_2)$. *Adv. Math.*, 204(1):1–83, 2006.

[70] U. Haagerup and S. Thorbjørnsen. A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group. *Ann. of Math. (2)*, 162(2):711–775, 2005.

[71] G. Hargé. A convex/log-concave correlation inequality for Gaussian measure and an application to abstract Wiener spaces. *Probab. Theory Related Fields*, 130(3):415–440, 2004.

[72] F. Hiai and D. Petz. A free analogue of the transportation cost inequality on the circle. *arxiv*, 2005.

[73] K. Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.*, 91:151–204, 1998.

[74] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[75] T. Kriecherbauer and M. Shcherbina. Fluctuations of eigenvalues of matrix models and their applications. 2010. math-ph/1003.6121.

[76] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[77] M. Ledoux. A remark on hypercontractivity and tail inequalities for the largest eigenvalues of random matrices. In *Séminaire de Probabilités XXXVII*, volume 1832 of *Lecture Notes in Math.*, pages 360–369. Springer, Berlin, 2003.

[78] E. Maurel-Segala. private communication.

[79] E. Maurel-Segala. High order expansion for matrix models. *http://front.math.ucdavis.edu/0608.5192*, 2006.

[80] R. J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80(2):309–323, 1995.

[81] M. W. Meckes. Concentration of norms and eigenvalues of random matrices. *J. Funct. Anal.*, 211(2):508–524, 2004.

[82] M. L. Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

[83] M. L. Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

[84] V. D. Milman and G. Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[85] E. Nelson. Notes on non-commutative integration. *J. Funct. Anal.*, 15:103–116, 1974.

[86] L. Pastur. On the spectrum of random matrices. *Teor. Mat. Fiz.*, 10(1):102–112, 1972.

[87] L. Pastur and M. Shcherbina. *Eigenvalue distribution of large random matrices*, volume 171 of *Mathematical Surveys and Monographs*. AMS, Providence, Rhode Island, 2011.

[88] G. K. Pedersen. *C^*-algebras and their automorphism groups*, volume 14 of *London Mathematical Society Monographs*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1979.

[89] M. Pimsner and D. Voiculescu. K-groups of reduced crossed products by free groups. *J. Operator Theory*, 8(1):131–156, 1982.

[90] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[91] W. Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991.

[92] D. Ruelle. *Statistical mechanics: rigorous results*. Benjamin, Amsterdam, 1969.

[93] A. Selberg. Remarks on a multiple integral. *Norsk. Mat. Tiddskr.*, 26:71–88, 1944.

[94] J. W. Silverstein and Z. D. Bai. On the empirical distribution of eigenvalues of a class of large-dimensional random matrices. *J. Multivariate Anal.*, 54(2):175–192, 1995.

[95] A. Soshnikov. Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.*, 207(3):697–733, 1999.

[96] G. 't Hooft. A planar diagram theory for strong interactions. *Nuclear Physics B*, 72(3):461–473, 1974.

[97] M. Talagrand. A new look at independence. *Ann. Probab.*, 24(1):1–34, 1996.

[98] F. G. Tricomi. *Integral equations*. Dover Publications, New York, NY, 1985. Reprint of the 1957 original.

[99] W. T. Tutte. On the enumeration of planar maps. *Bull. Amer. Math. Soc.*, 74:64–74, 1968.

[100] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[101] C. Villani. *Optimal transport, old and new.* Lecture Notes for the 2005 Saint-Flour Probability Summer School. www.umpa.ens-lyon.fr/ cvillani/StFlour/oldnew-9.ps, 2007.

[102] D. Voiculescu. Limit laws for random matrices and free products. *Invent. Math.*, 104(1):201–220, 1991.

[103] D. Voiculescu. Lectures on free probability theory. In *Lectures on probability theory and statistics (Saint-Flour, 1998)*, volume 1738 of *Lecture Notes in Math.*, pages 279–349. Springer, Berlin, 2000.

[104] E. P. Wigner. On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)*, 67:325–327, 1958.

[105] I. Zakharevich. A generalization of wigner's law. 2005.