Random matrices and free probability

Habilitation Thesis

(english version)

Florent Benaych-Georges

December 4, 2011

Introduction

This text is a presentation of my main researches in mathematics since the beginning of my PhD Thesis, which are all about random matrices and free probability theory, except one of them, devoted to the cycles of random permutations. I have divided them into seven chapters, to which are added a chapter where I present several research projects and an appendix giving a short introduction to free probability theory. Moreover, short reviews of various subjects are included in the form of inserts, so that the reader can easily refer to the ones which are useful to him and, just as easily, ignore those which are not.

The first chapter is about the construction of the rectangular free convolution $\boxplus_c$, and about the results I have obtained on this subject [A5, A8, A10, A16]. This work was initially part of a larger framework, the study of the *non-commutative joint distributions* of collections of rectangular random matrices. Indeed, I proved that *freeness with amalgamation* on a certain sub-algebra allows to model the asymptotic behavior of rectangular random matrices, once they have been embedded in larger square matrices. However, after my thesis, I have not developed this point of view much further, focusing on the rich enough and a bit more concrete study of rectangular free convolution. Besides, the asymptotic freeness results are quite hard to formulate, especially for random matrices, and can seem a bit arid to the reader unaware to their practical consequences. For these reasons, I chose to construct this chapter around the free rectangular convolution and to get into freeness with amalgamation only in the last paragraph, even though the existence of the free convolution follows from this freeness. For the same reasons, I do not tackle my work of [A8] on R-diagonal operators between two subspaces and on the associated entropy.

This convolution is the operation which allows to infer the empirical singular values distribution of the sum $A + B$ of two independent random matrices from the knowledge of the empirical singular values distributions of A and B. This operation is analogous to the classical convolution $*$ and to the "square type" free convolution $\boxplus$ of Voiculescu. It has recently been used by Gregoratti, Hachem and Mestre to analyse a telecommunications system in [59].

The second chapter is devoted to the so-called BBP phase transition, named after the authors Baik, Ben Arous and Péché who discovered it. It concerns the extreme eigenvalues of a random matrix X perturbed by a matrix whose rank stays bounded as the dimension goes to infinity. The general principle is that if the amplitude of the perturbation stays below a certain threshold, the largest eigenvalues do not move significantly, whereas above this threshold, they move significantly away from their initial positions. This phenomenon had first been proved for particular examples of random matrices X (namely Wigner and Wishart matrices). I have generalized it with my co-authors Raj Rao, Alice Guionnet and Mylène Maïda [A14, A15, A17], and in this text I propose an interpretation of this phenomenon via free probabilities. Moreover, with Alice Guionnet and Mylène Maïda, I have proposed a large deviations analysis of this model [A18].

The third chapter is devoted to the theory of free infinite divisibility and to its applications, which are about the definition of new matricial models, the regularization by convolution, and the repulsion of the singular values at the origin.

There exists a quite deep relation between infinitely divisible laws for the convolutions $*$, $\boxplus$ and $\boxplus_c$. These sets of laws are indeed indexed, via the Lévy-Kinchine Formulas, with the same objects. As a consequence, they are in bijective correspondance with each other. These bijections happen to preserve the limit theorems of the type Law of Large Numbers, Central Limit Theorem, etc... It has been proved for $\boxplus$ by Voiculescu, Bercovici and Pata in the papers [15, 16, 14] and in my paper [A4] for $\boxplus_c$. In this chapter, I present some new matrix ensembles I constructed in my papers [A2, A4], which generalize the GO(U)E and which give to the above-mentioned bijections a more concrete interpretation. Besides, in my paper [A7] with Serban Belinschi and Alice Guionnet, the particular properties of the infinitely divisible laws for $\boxplus_c$ and $\boxplus$ are used to prove some regularizing properties of their semigroups, and a *repulsion of the singular values at zero* phenomenon, that says that the singular values of the sum of two independent non Hermitian random matrices are likely to avoid a neighborhood of zero.

In the fourth chapter, I present a recent work, devoted to a universality result for the eigenvectors of Wigner matrices. It is proved that for $[u_{i,j}]_{i,j=1}^n$ the eigenvectors matrix of a Wigner matrix, the random bivalued process

$$\left(\sum_{1\le i\le ns,\,1\le j\le nt}(|u_{i,j}|^2-1/n)\right)_{(s,t)\in[0,1]^2}$$

converges in distribution (in a quite weak sense). The interesting fact is that when the entries of the Wigner matrix are centered with variance one, the law of the limit process depends only on their fourth moment (and not on their third one). In the case where this fourth moment coincides with the one of a Gaussian law, the limit process is a bivariate Brownian bridge and the convergence is proved in a stronger sense, the one of the Skorokhod topology.

In the fifth chapter, I present a joint work with Thierry Lévy, [A11], where I constructed a family of dependence structures in a non-commutative probability space, which interpolate between independence and freeness. It was known that if A and B are two large diagonal matrices whose eigenvalues are approximately distributed according to two probability distributions μ and ν, and if U is a permutation matrix chosen uniformly at random (resp. a unitary matrix chosen under the Haar measure), then the eigenvalues of $A + UBU^{-1}$ are distributed according to the measure $\mu * \nu$, the classical convolution of μ and ν (resp. according to $\mu \boxplus \nu$, the free convolution of μ and ν). Giving to U the distribution of a Brownian motion on the unitary group at time t whose distribution at time 0 is the uniform measure on the group of permutation matrices, we have defined an operation of convolution $*_t$ for all non-negative real numbers t, which for $t = 0$ is the classical convolution and, for t tending to infinity, tends to the free convolution. In fact, we defined the structure of dependence between two sub-algebras of a non-commutative probability space which underlies this convolution. Our initial hope was to identify cumulants associated with this t-free convolution, that is, universal multilinear forms, the cancellation of which would characterise the t-freeness of some of their arguments. We thought that they might interpolate between classical cumulants, which are intimately connected to the combinatorics of the partitions of a set, and free cumulants, related to the non-crossing partitions of a set endowed with a cyclic order. This hope has been turned down and we have in fact proved that there are no t-free cumulants.

The sixth chapter is devoted to a central limit theorem for the Brownian motion on the group of unitary $n \times n$ matrices, as n goes to infinity. More specifically, I consider linear combinations of the entries of such a process and I prove that as the dimension goes to infinity, there are three possible limit regimes, depending on whether we consider small, large or intermediate scales of time. In the first case, the limit process corresponds to a Brownian motion on the space of infinite skew-Hermitian matrices, in the second one, the limit process is a Brownian motion on the space of infinite complex matrices, and in the intermediate scale of time, one obtains an interpolation between both extremes. These three limit regimes can be explained by the way the unitary Brownian motion is constructed by *rolling* the unitary group on its Lie algebra along a Brownian motion on this algebra. A by-product of this work is a very short proof of the central limit theorem for the entries of a Haar-distributed unitary random matrix, a well known result already proved by Diaconis *et al.*

The seventh chapter is devoted to a work to which I devoted quite much energy and time during the two years following my thesis, that I still consider rather interesting and deep, though I have to confess it had very little echo. The reasons, totally justified in my opinion, why this work seemed to bring me to a kind of impasse, have taught me a lot to me about the way mathematics are built. Certain mistakes are instructive.

In this work, published in my paper [A12], we consider random permutations which can be written as free words in several independent random permutations: firstly, we fix a non trivial word w in letters $g_1, g_1^{-1}, \dots, g_k, g_k^{-1}$, secondly, for all n, we introduce a k-tuple $s_1(n), \dots, s_k(n)$ of independent random permutations of $\{1, \dots, n\}$, and the random permutation σ_n we are going to consider is the one obtained by replacing each letter g_i in w by $s_i(n)$. For example, for $w = g_1g_2g_3g_2^{-1}$, $\sigma_n = s_1(n) \circ s_2(n) \circ s_3(n) \circ s_2(n)^{-1}$. Moreover, we allow to restrict the set of possible lengths of the cycles of the $s_i(n)$'s: we fix sets $A_1, \dots, A_k$ of positive integers and suppose that for all i, $s_i(n)$ is uniformly distributed on the set of permutations of $\{1, \dots, n\}$ which have all their cycle lengths in A_i. For example, if $A_1 = \{1, 2\}$, $s_1(n)$ is a uniform random involution. We are interested in *small cycles* of σ_n, *i.e.* cycles whose length is fixed independently of n. Since the law of σ_n is invariant under conjugation, the positions of its cycles are uniform, and only their lengths contain some unknown randomness. So we introduce, for each positive integer ℓ, the number $N_\ell(\sigma_n)$ of cycles of length ℓ of σ_n. We are interested in the asymptotic behavior of the $N_\ell(\sigma_n)$'s as $n \longrightarrow \infty$. We first prove that the elements the word w represents in a certain quotient of the free group with generators $g_1, \dots, g_k$ determines the asymptotic order of the $N_\ell(\sigma_n)$'s and we prove that in many cases, the $N_\ell(\sigma_n)$'s are asymptotically independent, and distributed according to a Poisson law with parameter $1/\ell$. Beyond some asymptotic freeness issues, my interest for the question comes from another, seemingly quite hard problem, that Thierry Lévy and myself have tried to solve in vain: the characterization of the words w in the letters $g_1, g_1^{-1}, \dots, g_k, g_k^{-1}$ such that for any compact (or finite) group G, for any family $s_1, \dots, s_k$ of independent Haar distributed random variables on G, the law of the random variable obtained by replacing each letter g_i in w by s_i is the Haar measure on G.

At last, considering my papers [A1, A3, A6, A9, A20] as a bit isolated in my works, I will not present them here.

Contents

Chapter 1

Singular values of sums of random matrices

In this chapter, I am going to present my work on rectangular random matrices. As I explained in the introduction, this chapter is constructed around the rectangular free convolution, relegating its "ground", the freeness with amalgamation, to the second plan.

Most of the results presented in this text have been proved in both the real and complex cases. As a consequence, $\mathbb{K}$ denotes either $\mathbb{R}$ or $\mathbb{C}$ and $\beta = \dim_{\mathbb{R}} \mathbb{K}$. In the real case, the unitary matrices to consider are real, hence orthogonal.

1.1 Rectangular free convolution

The starting point of this study is the following question:

> *What can be said about the singular values of the sum $A + B$ of two matrices A and B, out of the singular values of A and B ?*

In full generality, it is of course very hard to answer this question, which is relevant to algebraic geometry (the definition and the geometric interpretation of the singular values of a matrix are recalled in Insert 1.1, at the end of this section). However, focusing on the large dimensional *generic* case, *i.e.* supposing A and B to be chosen at random, independently and according to isotropic distributions, and letting their dimensions go to infinity, one can give an answer.

More specifically, we shall consider random matrices $A, B \in \mathbb{K}^{n\times p}$ whose dimensions n, p will tend to infinity in such a way that $n/p \longrightarrow c \in [0, 1]$ (the dependance of A and B in n and p is implicit in the notation).

We make moreover the following hypotheses:

(a) A and B are independent,

(b) A or B is invariant, in law, under multiplication, on the right and on the left, by any unitary matrix,

(c) there exists μ_1, μ_2 laws on $\mathbb{R}_+$ such that, for the weak convergence in probability, we have

$$\frac{1}{n}\sum_{\sigma \text{ sing. val. of } A} \delta_\sigma \quad\longrightarrow\quad \mu_1 \qquad \text{and} \qquad \frac{1}{n}\sum_{\sigma \text{ sing. val. of } B} \delta_\sigma \quad\longrightarrow\quad \mu_2 \tag{1.1}$$

as $n, p \longrightarrow \infty$ with $n/p \longrightarrow c$.

Hypothesis (b) is an *isotropy* hypothesis (satisfied, for example, for a matrix with i.i.d. Gaussian entries) and Hypothesis (c) is the formalization of the idea that we know the singular values of A and B).

The answer to the question asked above is then given by the following theorem, which I proved during my thesis (see [A5, Th. 3.13][1]).

Theorem 1.1 *Under Hypotheses (a), (b) and (c) above, there is a non random law μ on $\mathbb{R}_+$, depending only on μ_1, μ_2 and c, such that for the weak convergence in probability, we have*

$$\frac{1}{n}\sum_{\sigma \text{ sing. val. of } A+B} \delta_\sigma \quad\longrightarrow\quad \mu \tag{1.2}$$

as $n, p \longrightarrow \infty$ with $n/p \longrightarrow c$.

The law μ, denoted by $\mu_1 \boxplus_c \mu_2$, is called the *rectangular free convolution with ratio c* of the laws μ_1 and μ_2.

Concretely, this theorem means that under Hypotheses (a) and (b) above, one can deduce the empirical singular values distribution of $A+B$ from the ones of A and B (and of n/p) : other types of informations on the singular values (extreme values, spacings,...) have no influence. Figure 1.1 gives an illustration of this phenomenon.

[1] The vocabulary used in certain articles, as [A5], is different: in order to have equlity, for $c = 1$, between $\boxplus_c$ and the convolution $\boxplus$ of Voiculescu, we therein consider the *symmetrizations* of the measures in (1.1) and (1.2). Since any law on $\mathbb{R}_+$ is entirely defined by its symmetrization, the results are equivalent.

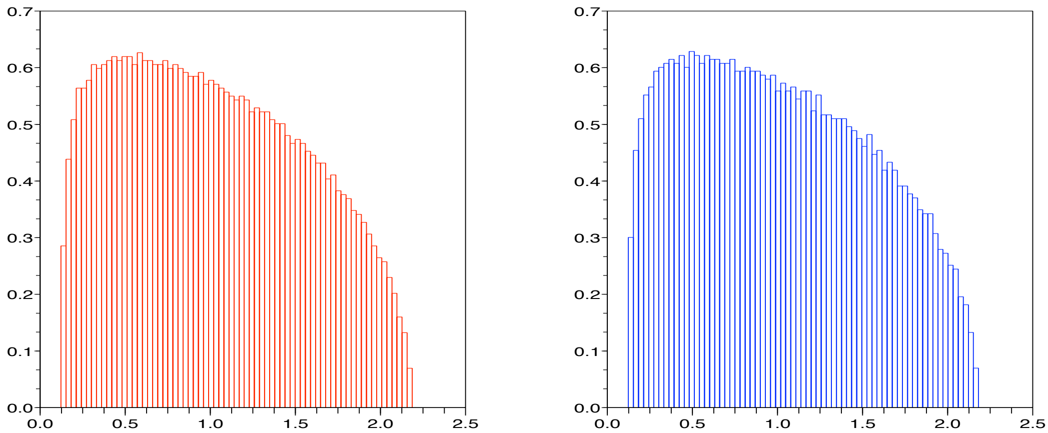


Figure 1.1: **Singular values of $A + B$ for different types of spacings :** Histograms of the singular values of $A + B$ for A a Gaussian matrix and B a matrix with singular values equal to $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}$ (left picture) or B a matrix with i.i.d. singular values uniformly distributed on $[0, 1]$ (right picture): the spacings of the singular values of B are much less regular in the second case than in the first one, and that does not impact the limit shape of the histogram (however, one can notice that the rate of convergence seems to be better in the left picture than in the right one, which corroborates the predictions given by the second order freeness theory). Here, the matrices have size $n \times p$ with $n = 4800$, $p = 6000$.

In the case of compactly supported probability measures (the general case being then deduced by approximation, with a control on the *rank* of the error), Theorem 1.1 is a consequence of Theorem 1.10, which characterizes the asymptotic behavior of the non-commutative distribution of large rectangular random matrices. This characterization allows to prove that the random variables

$$\frac{1}{n}\operatorname{Tr}[((A + B)(A + B)^*)^k] \qquad (k \geq 1)$$

(which are the *moments* of the empirical singular values distributions) concentrate around values which only depend on the numbers $\frac{1}{n}\operatorname{Tr}[(AA^*)^\ell]$ et $\frac{1}{n}\operatorname{Tr}[(BB^*)^\ell]$ (and of the ratio n/p). The dependance of the moments of the limit singular values distribution of $A + B$ on the numbers $\frac{1}{n}\operatorname{Tr}[(AA^*)^\ell]$ and $\frac{1}{n}\operatorname{Tr}[(BB^*)^\ell]$ is made explicit in the following section, via the *rectangular free cumulants* and the *rectangular R-transform*.

Remark 1.2 (Link with the convolution $\boxplus$ of Voiculescu) Theorem 1.1 is the analogue, for the singular values, of the result due to Voiculescu about the convolution $\boxplus$ and presented at Section 9.5 of the appendix. Relations between convolutions $\boxplus_c$ and $\boxplus$ will be presented below. We shall see that in the cases $c = 0$ and $c = 1$, $\boxplus_c$ can be expressed via $\boxplus$ and that for $0 < c < 1$, certain relations can be proved (*c.f.* Theorems 1.8 and 1.9).

Remark 1.3 (Singular values of AB) In the same way, for A, B isotropic matrices with respective sizes $n \times m$ and $m \times p$, as $n, m, p \longrightarrow \infty$ limit with $n/m \longrightarrow c$, $m/p \longrightarrow d$, the empirical singular values distribution of AB can be expressed out of the ones of A and B and of the limit ratios c, d. However, unlike for $A+B$, the "square type" multiplicative free convolution $\boxtimes$ of Voiculescu suffices to solve the problem. Indeed, up to some zeros and a square root, the singular values of AB are the eigenvalues of ABB^*A^*, i.e. of A^*ABB^*.

INSERT 1.1 – **Singular values of a matrix**

Let us recall the definition of the *singular values* of a matrix. Any matrix $A \in \mathbb{K}^{n\times p}$ can be written $A = UDV$, with $U \in \mathbb{K}^{n\times n}$, $V \in \mathbb{K}^{p\times p}$ both unitary and $D \in \mathbb{R}^{n\times p}$ null out of the diagonal and with non negative diagonal entries. The diagonal entries of D are then unique up to their order, and are called the *singular values* of A. The geometric interpretation of these values is the following: A maps the Euclidian unit ball to an ellipsoid, and the the singular values of A are exactly the half lengths of the $n \wedge p$ largest principal axes of this ellipsoid, the other axes having null length. Figure 1.2 gives an illustration in dimension two. The singular values of A are the eigenvalues of $\sqrt{AA^*}$ (resp. of $\sqrt{A^*A}$) if $n \leq p$ (resp. $n \geq p$).

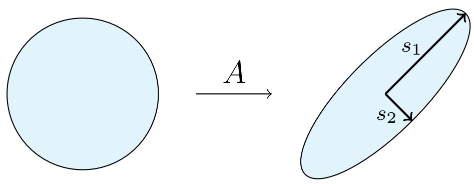


Figure 1.2: Singular values s_1 and s_2 of $A \in \mathbb{R}^{2\times 2}$.

1.2 An analytic tool to compute the convolution $\boxplus_c$: the rectangular R-transform

Theorem 1.1 states that the singular values of $A + B$ are distributed according to a law $\mu_1 \boxplus_c \mu_2$ which depends only on μ_1 and μ_2, but does not allow to compute it concretely out of μ_1 and μ_2. In this section, we make the dependence of $\mu_1 \boxplus_c \mu_2$ in μ_1, μ_2 and c explicit.

1.2.1 Definition

The rectangular free convolution $\boxplus_c$ can be computed thanks to an integral transform, like the classical convolution $*$ with the Fourier transform or the free convolution $\boxplus$ with the

R-transform. This transform, called the *rectangular R-transform with ratio c*, is defined as follows.

Let μ be a law on $\mathbb{R}_+$ and $c \in [0,1]$. One defines the functions[2]

$$M_\mu(z) := \int_{t\in\mathbb{R}_+} \frac{t^2 z}{1-t^2 z} \mathrm{d}\mu(t), \tag{1.3}$$

then

$$H_\mu^{(c)}(z) := z(cM_\mu(z)+1)(M_\mu(z)+1), \tag{1.4}$$

and at last

$$C_\mu^{(c)}(z) = T^{(c)^{-1}}\left(\frac{z}{H_\mu^{(c)^{-1}}(z)}\right) \quad \text{for } z \neq 0, \quad \text{and} \quad C_\mu^{(c)}(0) = 0, \tag{1.5}$$

where $T^{(c)}(z) = (cz+1)(z+1)$.

The function $C_\mu^{(c)}(\cdot)$ defined at (1.5) is called the *rectangular R-transform with ratio c* of μ. This transform allows to compute the rectangular free convolution of two laws, as asserted by the following theorem [A5, Th. 3.12].

Theorem 1.4 *(i) The analytic function* $C_\mu^{(c)}(\cdot)$ characterizes *the law μ.*

(ii) For μ_1, μ_2 laws on $\mathbb{R}_+$,

$$C^{(c)}_{\mu_1 \boxplus_c \mu_2}(z) = C^{(c)}_{\mu_1}(z) + C^{(c)}_{\mu_2}(z). \tag{1.6}$$

The proof of Formula (1.6) does not rely directly on the definition of the function $C_\mu^{(c)}$ via Formulas (1.3), (1.4) and (1.5), but on the study of the coefficients of the series expansion of $C_\mu^{(c)}$ around zero: setting, for μ compactly supported (the general case being then deduced by approximation)

$$C_\mu^{(c)}(z) = \sum_{n\geq 1} k_{2n}^{\boxplus_c}(\mu) z^n, \tag{1.7}$$

[2]In the case where μ is compactly supported, all these functions are analytic on a neighborhood of zero in $\mathbb{C}$, whose amplitude is controlled by the one of the support of μ. In the general case, the functions are still analytic, but they are defined on subsets of the type $\{\rho e^{i\theta}\,;\, 0 \leq \rho < c,\ |\theta - \pi| < c'\}$ ($c>0$, $0<c'<\pi$). The most adequate notion is then the one of *germ* of analytic functions on neighborhoods of zero which are non tangential to $\mathbb{R}_+$.

the coefficients $(k_{2n}^{\boxplus_c}(\mu))_{n\geq 1}$ are called the *rectangular free cumulants with ratio* c of μ. The function M_μ being the generating function of the moments of μ, it is not very hard to see that the relation between the moments $m_\ell(\mu) := \int_{\mathbb{R}} t^\ell \mathrm{d}\mu(t)$ and these cumulants is the following one:

$$m_{2n}(\mu) = \sum_{\pi} c^{e(\pi)} \prod_{v \text{ bloc de } \pi} k_{|v|}^{\boxplus_c}(\mu) \qquad (n \geq 1),$$

where the sum runs through the set of non-crossing partitions of the set $\{1, \ldots, 2n\}$ whose blocs all have an even cardinality and where $e(\pi)$ denotes the number of blocs of π with even minimal element. The study of the combinatoric structures underlying the *freeness with amalgamation* (*c.f.* Section 1.4) allows then to prove, via an analogue of Proposition 9.9, that

$$k_{2n}^{\boxplus_c}(\mu_1 \boxplus_c \mu_2) = k_{2n}^{\boxplus_c}(\mu_1) + k_{2n}^{\boxplus_c}(\mu_2) \qquad (n \geq 1).$$

Formula (1.6) follows immediately.

As showed by Formula (1.6), the rectangular R-transform with ratio c *linearizes* the convolution $\boxplus_c$. It will allow the concrete computation of $\boxplus_c$, but also to understand the links between the convolution $\boxplus_c$ and the convolution $\boxplus$ associated to Hermitian matrices (*c.f.* paragraph 1.3), to prove that $\boxplus_c$ is continuous for the weak topology, etc...

Let us now give an example of direct application of Theorem 1.4, corresponding to $\delta_1 \boxplus_c \delta_1$ (which is not a degenerate probability measure). Other examples can be found in [A5, Sect. 3.10].

Example 1.5 (Sum of isometries) *Consider $A, B \in \mathbb{K}^{n\times p}$ chosen at random, independently, one of them at least being invariant, in law, by left and right unitary actions, such that for all $\varepsilon > 0$, we have the convergence in probability*

$$\sharp\{\sigma \textit{ sing. val. of } A \textit{ or } B \textit{ such that } |\sigma - 1| > \varepsilon\} = o(n)$$

as $n, p \longrightarrow \infty$ with $n/p \longrightarrow c$. Then, we have the weak convergence in probability

$$\frac{1}{n} \sum_{\sigma \textit{ sing. val. of } A+B} \delta_\sigma \quad \longrightarrow \quad \frac{2\sqrt{\kappa^2 - (x^2-2)^2}}{c\pi x(4-x^2)} \mathbb{1}_{|x^2-2|\leq\kappa} \mathbb{1}_{x\geq 0} \mathrm{d}x, \tag{1.8}$$

with $\kappa = 2\sqrt{c(2-c)}$ (in the case where $c = 0$, the right term has to be interpreted as the Dirac mass at $\sqrt{2}$). Figure 1.3 illustrates this convergence.

1.2.2 Link with spherical integrals

Theorem 1.4 allows to assert that the rectangular R-transform plays an analogue role, for the convolution $\boxplus_c$, to the role played by the logarithm of the Laplace transform for the

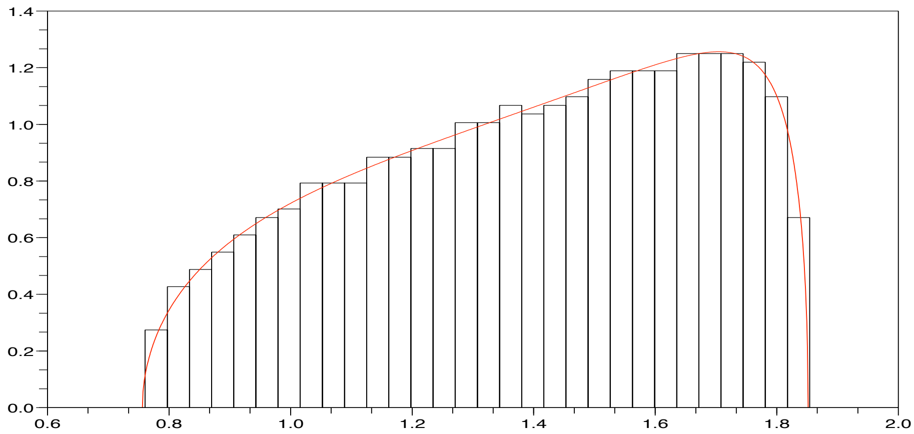


Figure 1.3: **Singular values of a sum of isometries:** Histogram of the singular values of $A+B$ and density of the limit law as predicted by Equation (1.8), under the hypotheses of Example 1.5. Here, $n = 10^3$, $c = 0.3$.

classical convolution $*$. The following theorem, from [A16], allows to claim that, beyond this analogy, the R-transform of a law μ is actually, up to an integration, the limit of the logarithm of a sequence of Laplace transforms.

We consider a matrix $M \in \mathbb{K}^{n\times p}$, depending implicitly on the parameters n and p. We suppose that M can be written $M = UDV$, with $U \in \mathbb{K}^{n\times n}$ and $V \in \mathbb{K}^{p\times p}$ independent and distributed according to the Haar measure on the unitary group, and D deterministic, uniformly bounded, such that

$$\frac{1}{n} \sum_{\sigma \text{ sing. val. of } D} \delta_\sigma \quad \longrightarrow \quad \mu,$$

where μ is a compactly supported law. We also introduce a matrix $E \in \mathbb{K}^{n\times p}$ whose entries are all zero, except one of them, equal to one. Then we have the following theorem.

Theorem 1.6 *As $n, p \longrightarrow \infty$ with $n/p \longrightarrow c \in [0, 1]$ (if $c = 0$, we also suppose that $p = o(n^2)$), we have, for θ a small enough real number,*

$$\frac{1}{n} \log \mathbb{E}\{\exp[\sqrt{np}\theta\Re(\mathrm{Tr}(EM))]\} \longrightarrow \beta \int_0^{\frac{\theta}{\beta}} \frac{C_\mu^{(c)}(t^2)}{t} \mathrm{d}t. \tag{1.9}$$

Let R_ν denote the "square type" R-transform of any probability measure ν (see Section 9.6.3 of the appendix). In the cases $c = 0$ and $c = 1$, the function $C^{(c)}$ being related respectively to the functions R_{μ^2} and $R_{\mathrm{s}(\mu)}$, where μ^2 denotes the law of X^2 for X distributed according to μ and $\mathrm{s}(\mu)$ denotes the symmetrization of μ (see Equation (1.11) below), one gets the following corollary.

Corollary 1.7 *In the case where $c = 0$ (resp. $c = 1$), we have*

$$\frac{1}{n}\log\mathbb{E}\{\exp[\sqrt{np}\theta\Re(\mathrm{Tr}(EM))]\} \quad\longrightarrow\quad \beta\int_0^{\frac{\theta}{2}} tR_{\mu^2}(t^2)\mathrm{d}t \qquad \textit{(resp. } \beta\int_0^{\frac{\theta}{2}} R_{\mathrm{s}(\mu)}(t)\mathrm{d}t\textit{)}.$$

The starting point of the study which led me to the previous results can be found in the works of Guionnet, Zeitouni, Collins, Zinn-Justin, Zuber, Maïda, Śniady, Mingo and Speicher, who proved, in the papers [63, 125, 39, 61, 41, 42, 40], that under various hypotheses on the $n \times n$ matrices A, E, for U distributed according to the Haar measure on the unitary group, for an adequate value of the exponent α, the asymptotic behavior of

$$\frac{1}{n^\alpha}\log\mathbb{E}\{\exp[n\theta\,\mathrm{Tr}(EUAU^*)]\}$$

is linked to free probability theory. For example, it has been proved [61, Th. 2] that if the empirical spectral law of an Hermitian matrix A tends to a law μ, then for $E = \mathrm{Diag}(1, 0, \ldots, 0)$ and θ small enough real number,

$$\frac{1}{n}\log\mathbb{E}\{\exp[n\theta\,\mathrm{Tr}(EUAU^*)]\} \underset{n\to\infty}{\longrightarrow} \frac{\beta}{2}\int_0^{\frac{2\theta}{\beta}} R_\mu(t)\mathrm{d}t. \tag{1.10}$$

Theorem (1.6) is proved by expressing the columns of U and V out of Gaussian vectors, which allows to express the expectation $\mathbb{E}\{\exp[\sqrt{np}\theta\Re(\mathrm{Tr}(EM))]\}$ as a Gaussian integral. An adequate change of variable makes then appear the rectangular R-transform of the empirical singular values distribution of M. To my knowledge, this technic is due to Guionnet and Maïda [61].

Before closing this paragraph, let us mention the fact that the expectations of exponentials of traces of random matrices (which are the partition functions of Gibbs measures) are usually called *spherical integrals*. The integral called *Harich-Chandra-Itzykson-Zuber Integral* is a well known example of those. Beyond the link with free probability mentioned above, the spherical integrals are of interest to the community of physicists and of Information Theory. The reader might find some references about spherical integrals involving square matrices in the texts [126, 63, 60]. The case of rectangular matrices has also been the object of investigations before the work presented here, *e.g.* in the papers [102, 57, 74].

1.3 Relations between square and rectangular free convolutions

The rectangular free convolution $\boxplus_c$ gives the distribution of the singular values of the sum of two non Hermitian matrices chosen at random, independently and in an isotropic way, whose dimensions $n, p \gg 1$ satisfy $n/p = c$. In the same way, the "square" free convolution $\boxplus$ gives the distribution of the eigenvalues of the sum of two large Hermitian matrices, chosen at random independently and in an isotropic way. The singular values of rectangular matrices and the eigenvalues of Hermitian matrices are related, since the singular values of an $n \times p$ matrix M are the eigenvalues of the matrix $\sqrt{MM^*}$. However, the operation $M \longmapsto \sqrt{MM^*}$ being non linear, it does not seem obvious that there exists a relation between the convolutions $\boxplus$ and $\boxplus_c$.

However, when $c = 1$ (*i.e.* when the matrices are all square ones), putting together the result of asymptotic freeness by Voiculescu [116] and the study of the sum of R-diagonal elements by Haagerup and Larsen [65, Prop. 3.5], one gets the following theorem, which can also be recovered very directly using the tools presented in this chapter, by noticing that for $c = 1$, the rectangular R-transform with ratio c is more or less the R-transform of Voiculescu. For μ a law on $\mathbb{R}_+$, we denote by $\mathrm{s}(\mu)$ its *symmetrization*, i.e. the law $\mathbb{R}$ defined by

$$\mathrm{s}(\mu)(A) = \frac{\mu(A) + \mu(-A)}{2} \tag{1.11}$$

for each Borel set A.

Theorem 1.8 *The rectangular free convolution with ratio* 1 *of two laws* μ_1, μ_2 *on* $\mathbb{R}_+$ *is the law on* $\mathbb{R}_+$ *whose symmetrization is the free convolution of the symmetrizations of* μ_1 *and* μ_2*. In other words, for* μ_1, μ_2 *laws on* $\mathbb{R}_+$*,* $\mu_1 \boxplus_1 \mu_2$ *is the law on* $\mathbb{R}_+$ *defined by the formula*

$$\mathrm{s}(\mu_1 \boxplus_1 \mu_2) = \mathrm{s}(\mu_1) \boxplus \mathrm{s}(\mu_2).$$

This result has a simple matricial interpretation. It means that for M, N independent large isotropic square matrices, the spectral law of

$$\begin{bmatrix} 0 & M \\ M^* & 0 \end{bmatrix} + \begin{bmatrix} 0 & N \\ N^* & 0 \end{bmatrix}$$

is close to the free convolution $\boxplus$ of the spectral laws of

$$\begin{bmatrix} 0 & M \\ M^* & 0 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 0 & N \\ N^* & 0 \end{bmatrix}, \tag{1.12}$$

though the matrices of (1.12), having too much structure in common, are in general not asymptotically free.

The following result, [A10, Th. 3], which involves the free multiplicative convolution $\boxtimes$, seems more surprising. For $c \in [0,1]$, let us introduce the *Marchenko-Pastur law*[3] *with parameter* c:

$$L_{\mathrm{MP},\,c} := \frac{\sqrt{(b-x)(x-a)}}{2\pi c x}\mathbb{1}_{x\in[a,b]}\mathrm{d}x, \tag{1.13}$$

for $a = (1-\sqrt{c})^2$, $b = (1+\sqrt{c})^2$ (when $c = 0$, this law has to be understood as the Dirac mass at 1). For μ a law on $\mathbb{R}_+$, we denote by $\sqrt{\mu}$ the law of $\sqrt{X}$ for X r.v. with law μ.

Theorem 1.9 *For μ_1, μ_2 laws on $\mathbb{R}_+$ and $c \in [0,1]$, we have*

$$\sqrt{\mu_1 \boxtimes L_{\mathrm{MP},\,c}} \boxplus_c \sqrt{\mu_2 \boxtimes L_{\mathrm{MP},\,c}} = \sqrt{(\mu_1 \boxplus \mu_2) \boxtimes L_{\mathrm{MP},\,c}}\ . \tag{1.14}$$

In particular, for $c = 0$,

$$\sqrt{\mu_1}\ \boxplus_0\ \sqrt{\mu_2} = \sqrt{\mu_1 \boxplus \mu_2}\ , \tag{1.15}$$

in other words, the rectangular free convolution with ratio 0 *of two laws is the unique law on $\mathbb{R}_+$ whose push-forward by the map $x \longmapsto x^2$ is the free additive convolution $\boxplus$ of the push-forwards of these laws by this map.*

This theorem has several unexpected consequences, concerning arithmetics of the convolutions $\boxplus$ and $\boxtimes$ of Voiculescu as well as the infinite divisibility. The reader can find them in [A10]. Let us now give interpretations, rather surprising in my point of view, of Formulas (1.14) and (1.15).

Let us begin with (1.15), which is more simple. From the random matrices point of view, this formula means that for A, B independent $n \times p$ random matrices, when $1 \ll n \ll p$, as far as the empirical spectral law is concerned, we have

$$(A+B)(A+B)^* \simeq AA^* + BB^*.$$

Let us now give a matricial interpretation of Formula (1.14). We consider the asymptotic regime $n, p \gg 1$ and $n/p \simeq c$. Let A, B be non Hermitian $n \times n$ random matrices, one at least being invariant by left and right unitary multiplications, such that

$$\frac{1}{n}\sum_{\lambda \text{ eig. of } AA^*} \delta_\lambda \longrightarrow \mu_1 \qquad \text{and} \qquad \frac{1}{n}\sum_{\lambda \text{ eig. of } BB^*} \delta_\lambda \longrightarrow \mu_2,$$

[3] The exists several conventions in the definition of the Marchenko-Pastur laws. This one corresponds to the limit spectral law of XX^*, for X and $n \times p$ matrix with i.i.d. centered entries with variance $1/p$, as $n, p \longrightarrow \infty$ with $n/p \longrightarrow c$.

and let X, X' be $n \times p$ matrices with i.i.d. Gaussian entries with variance $1/p$. Then $\sqrt{\mu_1 \boxtimes L_{\mathrm{MP},c}}$ and $\sqrt{\mu_2 \boxtimes L_{\mathrm{MP},c}}$ are the limit empirical singular values distributions of AX and BX', whereas $\sqrt{(\mu_1 \boxplus \mu_2) \boxtimes L_{\mathrm{MP},c}}$ is the limit empirical singular values distribution of $\sqrt{AA^* + BB^*}X$. Hence Equation (1.14) can be interpreted in the following way: as far as empirical singular values distributions are concerned

$$AX + BX' \simeq \sqrt{AA^* + BB^*}X. \tag{1.16}$$

Figure 1.4 illustrates this phenomenon.

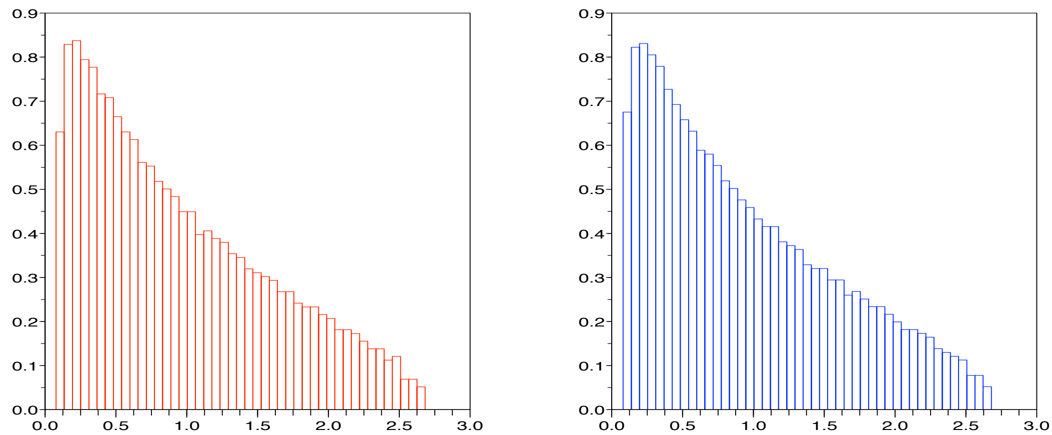


Figure 1.4: **Illustration of Identity** (1.16) **:** Histograms of the singular values of $AX+BY$ (left) and of $\sqrt{AA^* + BB^*}X$ (right) for A an $n \times n$ matrix whose eigenvalues are uniformly distributed on $[0,1]$ and B, X, Y matrices with respective sizes $n \times n$, $n \times p$, $n \times p$ having i.i.d. Gaussian entries with respective variances $1/n$, $1/p$, $1/p$. Here, $n = 2000$, $p = 2500$.

1.4 Asymptotic non-commutative distribution of rectangular random matrices

In this section, we explain how freeness with amalgamation over a certain sub-algebra allows to model the asymptotic behavior of rectangular random matrices, once they have been embedded in larger square matrices. These results are the foundations of what precedes in this chapter: indeed, Theorem 1.1, which allows to define the convolution $\boxplus_c$, follows from Theorem 1.10 below. The definitions relative to *$\mathcal{D}$-valued free probability theory* and to *freeness with amalgamation* can be found in Section 9.7 of the appendix.

Let us consider an integer $d \geq 1$ and, for each $n \geq 1$, some positive integers $q_1, \ldots, q_d$ depending on n (this dependence is let implicit, in order to lighten the notation) such that

$$q_1 + \cdots\cdots + q_d = n \tag{1.17}$$

and such that for all $i = 1, \ldots, d$, there is $\rho_i > 0$ such that $\frac{q_i}{n} \longrightarrow \rho_i$ as $n \longrightarrow \infty$.

Let us denote by $\mathcal{D}$ the algebra of $d \times d$ complex diagonal matrices. We shall endow the algebra $M_n(\mathbb{C})$ of $n \times n$ complex matrices with a structure of **$\mathcal{D}$-valued non-commutative probability space**. To do so, we consider the $n \times n$ matrices as $d \times d$ bloc matrices, following the subdivision of n given by formula (1.17). It allows, assimilating scalar matrices and scalars, to see $\mathcal{D}$ as a **sub-algebra** of $M_n(\mathbb{C})$. Moreover, we define an application $\varphi_{\mathcal{D},n} : M_n(\mathbb{C}) \to \mathcal{D}$ in the following way : for $M \in M_n(\mathbb{C})$, we write, via the bloc decomposition introduced above, $M = [M_{i,j}]_{i,j=1}^d$, and set

$$\varphi_{\mathcal{D},n}(M) = \begin{pmatrix} \operatorname{tr}(M_{1,1}) & & \\ & \ddots & \\ & & \operatorname{tr}(M_{d,d}) \end{pmatrix}, \tag{1.18}$$

tr denoting the normalized trace. This application $\varphi_{\mathcal{D},n} : M_n(\mathbb{C}) \to \mathcal{D}$ is a **conditional expectation**.

For each $n \geq 1$, we then consider two collections $(X_i)_{i \in I}$ and $(D_j)_{j \in J}$ of $n \times n$ matrices (the dependence in n of each X_i and each D_j is still implicit) satisfying the following hypotheses :

(a) $(D_j)_{j \in J}$ is a collection of deterministic matrices whose $\mathcal{D}$-distribution converges, as $n \longrightarrow \infty$, to the one of a collection $(d_j)_{j \in J}$ of elements of a $\mathcal{D}$-valued non-commutative probability space,

(b) $(X_i)_{i \in I}$ is a collection of independent random matrices such that for all $i \in I$, each of the d^2 coordinates of X_i in the bloc decomposition introduced above is invariant, in law, under the left and right unitary actions,

(c) for all $i \in I$, the non-commutative $\mathcal{D}$-distribution of X_i converges in probability, as $n \longrightarrow \infty$, to the one of an element x_i of a $\mathcal{D}$-valued non-commutative probability space.

The following theorem is the analogue of the one of Voiculescu about the asymptotic freeness of square matrices (Theorem 9.6 of the appendix).

Theorem 1.10 *Under the preceding hypotheses, the collection*

$$(X_i)_{i \in I} \cup (D_j)_{j \in J}$$

converges in $\mathcal{D}$-distribution, as $n \longrightarrow \infty$, to a collection $(\tilde{x}_i)_{i \in I} \cup (\tilde{d}_j)_{j \in J}$ whose $\mathcal{D}$-distribution is defined as follows :

- *we have the equalities in $\mathcal{D}$-distributions* $(\tilde{d}_j)_{j\in J} = (d_j)_{j\in J}$ *and* $\tilde{x}_i = x_i$ *for all* i*,*
- *the* $\tilde{x}_i$*'s are* free with amalgamation over $\mathcal{D}$ *together, and with* $\{\tilde{d}_j\, ;\, j \in J\}$*.*

One can find a slightly more sophisticated version of this theorem, allowing $\rho_i = 0$ and more latitude about the choice of the matrices X_i, in my paper [A5] (Theorems 1.6 and 1.7). I have also proved another version, for example for matrices with i.i.d. or band entries, in [A8]. At last, let us mention that a result relying on the same philosophy has been proved by Shlyakhtenko in [103].

Let us notice that this theorem allows to characterize the asymptotic non-commutative distributions of rectangular random matrices. Indeed, choosing the q_i's in an adequate way, one can always embed any rectangular matrix in a larger square matrix, by extending it with some zeros.

Before closing this chapter, let us say a few words about the proof of Theorem 1.10. It goes considering a matrix M which can be written

$$M = M_1 \cdots M_p,$$

where each M_k can be written

$$M_k = P(X_i, X_i^*) \qquad \text{or} \qquad M_k = P(D_j, j \in J),$$

with P a non-commutative polynomial with $\mathcal{D}$-valued coefficients such that as $n \longrightarrow \infty$,

$$\varphi_{\mathcal{D},n}(M_k) \longrightarrow 0 \tag{1.19}$$

and such that two following matrices M_k, M_{k+1} are always independent. One then has to prove that as $n \longrightarrow \infty$,

$$\varphi_{\mathcal{D},n}(M) \longrightarrow 0. \tag{1.20}$$

The proof of the convergence (1.20) is done by expanding the normalized traces appearing in the definition of $\varphi_{\mathcal{D},n}$ at Equation (1.18): one obtains a sum, rather heavy, where the exploitation of Hypothesis (1.19) is not obvious. But working patiently, using the adequate combinatorial structures to find out the order of each of the terms of the sum (see the appendix of [A5]), one at last manages to prove the convergence of (1.20)...

Chapter 2

The BBP phase transition

2.1 Presentation of the problem and case of Wigner and Wishart matrices

2.1.1 General context: finite rank perturbations of large random matrices

The question asked in this chapter is the following one :

> *What is the impact of a finite rank perturbation on the extreme eigenvalues of a large Hermitian matrix ?*

In the case where the matrix is not supposed to be Hermitian, we shall be interested in the same question, concerning the extreme singular values instead of the extreme eigenvalues. The eigenvectors and singular vectors will be also considered.

As it is asked, the question is quite vague. Let us make it more precise. One considers an Hermitian $n \times n$ matrix X, where n is an implicit parameter which shall tend to infinity.

We perturb X with an Hermitian matrix P whose rank r stays bounded as n tends to infinity. Hence we define

$$\widetilde{X} := X + P \qquad\qquad \text{(additive perturbation)},$$

or

$$\widetilde{X} := (I + P)X \qquad\qquad \text{(multiplicative perturbation)}$$

(in the case of a multiplicative perturbation, the matrix X is supposed to be non negative). We suppose that the operator norms of both X and P stay bounded as n goes to infinity. Up to an extraction, one can as a consequence make the following hypotheses.

Hypothesis 2.1 *The empirical spectral law of X converges, as $n \longrightarrow \infty$, to a compactly supported law μ.*

Hypothesis 2.2 *The rank r of P does not depend on n.*

Up to an extraction, again, one can suppose that the r non zero eigenvalues of P converge. In order to simplify the notation, we then make the following hypothesis[1]

Hypothesis 2.3 *The r non zero eigenvalues*

$$\theta_1 \geq \cdots\cdots \geq \theta_r$$

of P do not depend on n.

At last, one has to make a hypothesis on the relative positions of the eigenspaces of X and P, singular cases such as the one where X and P are codiagonalisable giving of course very particular results.

Hypothesis 2.4 *X and P are random and independent, in such a way that the eigenvectors of X and P are asymptotically in generic position[2] with each other.*

Let us then denote by

$$\lambda_1 \geq \cdots\cdots \geq \lambda_n \qquad \text{and} \qquad \widetilde{\lambda}_1 \geq \cdots\cdots \geq \widetilde{\lambda}_n$$

the respective eigenvalues of X and $\widetilde{X}$ (again, the dependence in the dimension n is implicit).

Let $r_0 \in \{0, \ldots, r\}$ be the number of positive eigenvalues of P (which has hence $r - r_0$ negative eigenvalues). By Weyl's interlacing inequalities, [2, Th. A.7], in the case of additive perturbations, for all $i = r_0 + 1, \ldots, n - (r - r_0)$, we have

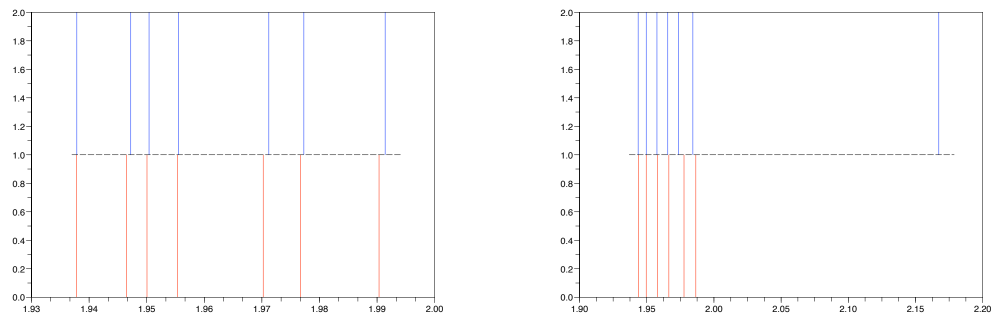
$$\lambda_{i-r_0} \geq \widetilde{\lambda}_i \geq \lambda_{i+(r-r_0)} \tag{2.1}$$

[1] In the results concerning the *convergence* and the *large deviations* of the extreme eigenvalues of $\widetilde{X}$, this hypothesis can, without any impact on the statements, be replaced by the weaker one asserting the *convergence* of the non zero eigenvalues of P to $\theta_1, \ldots, \theta_r$. However, when the *fluctuations* of the extreme eigenvalues of $\widetilde{X}$ are concerned, the rate of convergence of the non zero eigenvalues of P to $\theta_1, \ldots, \theta_r$ could change the conclusion.

[2] What being *asymptotically in generic position* exactly means is not specified here, but more precise hypotheses will be given later.

(for multiplicative perturbations, inequalities can also be proved). One deduces easily from (2.1) that the empirical spectral law of $\widetilde{X}$ converges to μ as $n \to \infty$, as the one of X. The perturbation did not modify the global repartition of the eigenvalues. The same cannot be said about the extreme ones : we shall see that large $|\theta_i|$'s give rise to important movings of the extreme eigenvalues, whereas for small ones, they shall stay quite close to their initial positions, so close that their fluctuations around their limits will keep the same orders and laws. This is the so-called *Baik-Ben Arous-Péché phase transition*, brought to light by these authors in the seminal paper [7].

Figure 2.1 illustrates this phenomenon and can give to the reader a preview of the results presented below.



(a) Case where $\theta = 0.5$ (b) Case where $\theta = 1.5$

Figure 2.1: **Comparison between the largest eigenvalues of a GUE matrix and those of the same matrix perturbed:** the abscises of the vertical segments correspond to the largest eigenvalues of X, a GUE matrix with size 2.10^3 (under the dotted line) or to those of $\widetilde{X} = X + \mathrm{diag}(\theta, 0, \ldots, 0)$ (above the dotted line). Left picture : $\theta = 0.5$. Right picture : $\theta = 1.5$. On the left picture, the eigenvalues of $\widetilde{X}$ are very close to the ones of X. On the right picture, one can observe the same mimicry between the eigenvalues of $\widetilde{X}$ and the ones of X, up to a shift : the larget eigenvalue of $\widetilde{X}$ equals ≈ 2.17.

2.1.2 Baik-Ben Arous-Péché phase transition: the two first examples

Deformed Wigner matrices

Let us consider a Wigner matrix[3] $\sqrt{n}X$. Then we know that, under certain hypotheses on the entries of X, we have the following convergences as $n \longrightarrow \infty$:

[3] A *Wigner matrix* is a random real symmetric or Hermitian $n \times n$ matrix with i.i.d. centered entries, having variance 1. More precisions about this definition and standard results (convergence to the semicircle law, extreme eigenvalues fluctuations,...) are recalled in Insert 2.3 below.

- the empirical spectral law of X tends to the semicircle law with support $[-2,2]$,

- for all fixed $k \geq 1$, the k^{th} extreme eigenvalues λ_k et λ_{n+1-k} of X tend to 2 and -2,

- for all fixed $k \geq 1$, the collection of the k largest eigenvalues of X has Tracy-Widom fluctuations with amplitude $n^{-2/3}$ around its limit 2 : $n^{2/3}(\lambda_i - 2)_{i=1,\ldots,k}$ converges in distribution to a Tracy-Widom law.

One then defines, as above, the additive perturbation $\widetilde{X}$ of X by the formula

$$\widetilde{X} = X + \mathrm{diag}(\underbrace{\theta_1, \ldots, \theta_{r_0}}_{>0}, \underbrace{\theta_{r_0+1}, \ldots, \theta_r}_{<0}, \underbrace{0, \ldots\ldots\ldots\ldots, 0}_{n-r \text{ null eigenvalues}}),$$

where r, r_0 and $\theta_1 \geq \cdots\cdots \geq \theta_r$ are independent of n.

One then has the following theorem, giving the asymptotic behavior of the largest eigenvalues $\widetilde{\lambda}_1 \geq \widetilde{\lambda}_2 \geq \cdots\cdots$ of $\widetilde{X}$. Of course, an identical result exists for the smallest ones. For all $\theta > 0$, we set

$$\rho_\theta := \begin{cases} \theta + \frac{1}{\theta} & \text{if } \theta > 1, \\ 2 & \text{if } 0 < \theta \leq 1. \end{cases}$$

Let us notice that the function $\theta \longmapsto \rho_\theta$ is increasing and that $\rho_\theta > 2$ when $\theta > 1$.

Theorem 2.5 *Under certain hypotheses*[4] *on the queues of the distributions of the entries of X, one has the following convergences.*

(a) For all fixed $i = 1, \ldots, r_0$,

$$\widetilde{\lambda}_i \underset{n\to\infty}{\longrightarrow} \rho_{\theta_i}$$

and for all fixed $i > r_0$,

$$\widetilde{\lambda}_i \underset{n\to\infty}{\longrightarrow} 2.$$

(b) If X is a GUE matrix, for all $i_0 = 1, \ldots, r_0$ such that $\theta_{i_0} > 1$, setting

$$J := \{i = 1, \ldots, r_0\,;\, \theta_i = \theta_{i_0}\},$$

[4]The hypotheses are of the type "sub-Gaussian tails" or "Poincaré inequality" (for the case "rank one" in (c), the laws of the entries need also being symmetric).

then up to a multiplicative constant, the collection

$$\{\sqrt{n}(\widetilde{\lambda}_i - \rho_{\theta_{i_0}})\,;\, i \in J\}$$

converges in distribution to the law of the collection of the ordered eigenvalues of a GUE matrix with size $\sharp J$.

(c) *Suppose firstly that $\theta_1 \in (0,1]$, and secondly that X is a GUE matrix or that $r = 1$. Then the fluctuations of $\widetilde{\lambda}_1$ around its limit 2 have amplitude $n^{-2/3}$ and are of Tracy-Widom type if $\theta_1 < 1$ or of generalized Tracy-Widom type if $\theta_1 = 1$.*

The (partial) definition of the Tracy-Widom distributions is recalled in Insert 2.4 below. Parts (a) and (b) of the theorem are due to Péché, Féral, Capitaine eand Donati-Martin [92, 52, 33, 34]. The proof of (c) is due to Péché (GUE case, [92, Th. 1.1]) and Péché and Féral ($r = 1$ case, [52, Th. 1.4]). In the case where X is not a GUE matrix and where $\sharp J = 1$, the convergence of (b) still holds, but the limit law is not Gaussian anymore [33, Th. 2.2].

Part (a) of the theorem signifies that to any eigenvalue $\theta_i > 0$ of the perturbing matrix $P := \operatorname{diag}(\theta_1, \ldots, \theta_r, 0, \ldots, 0)$, one can associate an eigenvalue $\widetilde{\lambda}_i$ of the perturbed matrix $\widetilde{X}$ which is out of the "bulk" of the spectrum of X if and only if $\theta_i > 1$: a *phase transition* with threshold equal to 1 occurs.

Parts (b) and (c) are manifestations of the *repulsion principle* for the eigenvalues.

Indeed, Part (b) can be interpreted as follows. To an eigenvalue $\theta_{i_0} > 1$ of P with multiplicity one, one can associate an eigenvalue $\widetilde{\lambda}_{i_0}$ of $\widetilde{X}$ which is isolated out of $[-2, 2]$, with rather large Gaussian fluctuations (with amplitude $1/\sqrt{n}$). We shall see later (*c.f.* Theorem 2.9) that the fluctuations of the eigenvalue $\widetilde{\lambda}_{i_1}$ associated to another eigenvalue $\theta_{i_1} > 1$ of P with multiplicity 1 are independent of the ones of $\widetilde{\lambda}_{i_0}$. However, if one lets θ_{i_0} and θ_{i_1} get closer and closer, until being equal, then the joint fluctuations are not anymore the ones of two independent Gaussian variables, but the ones of the eigenvalues of a Gaussian matrix with size 2, some variables in repulsive interaction.

Part (c) means that if the θ_{i_0} of the previous paragraph goes down below the threshold 1, then the associated eigenvalue $\widetilde{\lambda}_{i_0}$ enters the "bulk" of the spectrum, where it is not isolated anymore at all, and the amplitude of its fluctuations is sharply reduced from $1/\sqrt{n}$ to $1/n^{2/3}$. We shall even see later, at Theorem 2.9, that the *joint* fluctuations of such eigenvalues are of Tracy-Widom type, which supports this idea.

Deformed empirical covariance matrices

Let us consider a centered Gaussian vector $G \in \mathbb{K}^{n\times 1}$, with covariance matrix I. We consider p independent copies $G_1, \ldots, G_p$ of G and we define the *empirical covariance matrix* (also called a *Wishart matrix*)

$$X := \frac{1}{p}\sum_{k=1}^{p} G_k G_k^*.$$

One can then prove that if $n, p \longrightarrow \infty$ in such a way that $n/p \longrightarrow c \in (0, 1]$, we have :

- the empirical spectral distribution of X tends to the Marchenko-Pastur law $L_{\mathrm{MP},\,c}$ introduced at Formula (1.13),
- for all fixed $k \geq 1$, the k^{th} extreme eigenvalues λ_k and λ_{n+1-k} of X tend to the bounds $a = (1-\sqrt{c})^2$ and $b = (1+\sqrt{c})^2$ of the support of $L_{\mathrm{MP},\,c}$,
- for all fixed $k \geq 1$, the k^{th} largest eigenvalue λ_k of X has Tracy-Widom type fluctuations with amplitude $n^{-2/3}$ around its limit b : for a certain constant κ, $\kappa n^{2/3}(\lambda_k - b)$ converges in distribution to a Tracy-Widom law.

These results can be found respectively in [81, 56, 73].

We now consider the empirical covariance matrix

$$\widetilde{X} := \frac{1}{p}\sum_{k=1}^{p} \widetilde{G}_k \widetilde{G}_k^*$$

of a sample $\widetilde{G}_1, \ldots, \widetilde{G}_p$ of independent copies of a centered Gaussian vector $\widetilde{G} \in \mathbb{K}^{n\times 1}$, whose covariance matrix Σ has spectrum

$$\ell_1 \geq \cdots\cdots \geq \ell_r \; > \underbrace{1 \geq \cdots\cdots\cdots\cdots\cdots\cdots \geq 1}_{n-r \text{ eigenvalues equal to } 1},$$

where r and the ℓ_i's do not depend on n. The matrix $\widetilde{X}$ is a multiplicative perturbation of X because it can be realized via the formula $\widetilde{X} = \sqrt{I+P}X\sqrt{I+P}$, where $I + P = \Sigma$.

The following theorem has been proved by Baik, Ben Arous and Péché in the complex case [7] and by Paul in the real case [91]. Let us recall that b, the upper bound of the support of $L_{\mathrm{MP},\,c}$, is the limit of the largest eigenvalue λ_1 of X.

Theorem 2.6 *The largest eigenvalues of $\widetilde{X}$ satisfy the following phase transition.*

For all $i \in \{1, \ldots, r\}$, as $n, p \longrightarrow \infty$ with $n/p \longrightarrow c \in (0, 1]$,

$$\widetilde{\lambda}_i \quad \longrightarrow \quad \begin{cases} \ell_i(1 + \frac{c}{\ell_i - 1}) > b & \text{if } \ell_i > 1 + \sqrt{c}, \\ \\ b & \text{otherwise.} \end{cases} \tag{2.2}$$

*As for deformed Wigner matrices, (*c.f. *Theorem 2.5), the fluctuations of the eigenvalues with limit $> b$ have amplitude $1/\sqrt{n}$ and are distributed as the eigenvalues of a GO(U)E matrix with finite size, whereas the fluctuations of the eigenvalues tending to b have Tracy-Widom fluctuations with amplitude $n^{-2/3}$.*

Notice that this theorem can be considered as relevant to statistics (it seems by the way to be its origin [73]): it gives the base of the construction of a statistical *test* to detect a *signal* (variables with variance $\ell_i > 1 + \sqrt{c}$) among a *noise* (the variables with variance 1). It allows moreover to estimate such ℓ_i's.

Besides, this phase transition can be interpreted in terms of last passage percolation (see insert 2.2).

Insert 2.2 – **BBP transition for last passage percolation**

Fix $n \geq 1$, $\pi_1, \ldots, \pi_n > 0$ and, for all $p \geq 1$, $\hat{\pi}_p \in [0, +\infty)$.

We consider :

- a matrix (with n rows and infinitely many columns) $[A_{i,j}]_{1 \leq i \leq n,\, 1 \leq j}$ with independent complex entries such that for all i, j, the real and imaginary parts of $A_{i,j}$ are independent centered Gaussian variables with variance $\frac{1}{2(\pi_i + \hat{\pi}_j)}$,
- a collection $[W_{i,j}]_{1 \leq i \leq n,\, 1 \leq j}$ indexed by $\{1, \ldots, n\} \times \{1, 2, 3, \ldots\}$ of independent random variables, such that for all i, j, $W_{i,j}$ is an exponential random variable with parameter $\pi_i + \hat{\pi}_j$.

We denote by $\lambda_1(n, p)$ the largest eigenvalue of the matrix $A(n, p)A(n, p)^*$, where $A(n, p) = [A_{i,j}]_{1 \leq i \leq n,\, 1 \leq j \leq p}$ and we define the *last passage percolation time*

$$L(n, p) := \max_{\pi \in (1,1) \nearrow (n,p)} \sum_{(i,j) \in \pi} W_{i,j}, \tag{2.3}$$

where the $\max$ is over the up/right paths π on $\mathbb{Z}^2$ which join $(1, 1)$ to (n, p).

Then one can prove the following identity in laws.

Theorem 2.7 *The processes $(\lambda_1(n, p))_{p \geq 1}$ and $(L(n, p))_{p \geq 1}$ have the same law.*

The most accomplished version of this result, presented here, has been proved in [44], but preliminary versions had already appeared in [72, 7, 27]. One can then deduce from the behavior of the largest eigenvalue of Wishart matrices that in the case where all the $W_{i,j}$ have parameter 1, as $n, p \longrightarrow \infty$ with $n/p \longrightarrow c \in [0,1]$, one has

$$\frac{1}{p}L(n,p) \quad \longrightarrow \quad (1+\sqrt{c})^2.$$

Let us now multiply the $W_{i,j}$'s of the **first column** by $\ell \geq 1$: the $W_{i,j}$'s are then some independent exponential variables with parameter $1/\ell$ if $i = 1$ and 1 otherwise. Hence the $W_{i,j}$'s of the first column are likely to be larger than the ones of the other columns, and the maximizing paths, in (2.3), are likely to stay a bit on the first column before going to the right. The larger ℓ will be, the more important this new trend will be. More specifically, Theorem 2.6 says that for $n, p \gg 1$ such that $n/p \approx c \in [0,1]$,

$$\frac{1}{p}L(n,p) \approx \begin{cases} \ell(1+\frac{c}{\ell-1}) & \text{if } \ell > 1+\sqrt{c}, \\ \\ (1+\sqrt{c})^2 & \text{otherwise,} \end{cases}$$

which means that the "bonus" allocated to the first column modifies significantly $L(n,p)$ if and only if $\ell > 1+\sqrt{c}$. It follows that the maximizing paths shall stay long[5] on the first column if and only if $\ell > 1+\sqrt{c}$.

This result can be recovered heuristically with an elementary calculus by estimating the time passed on the first column [94, Sect. 3.2].

Insert 2.3 – Wigner matrices, GOE and GUE

A *Wigner matrix* is a random real symmetric or Hermitian matrix whose entries are independent, identically distributed on the diagonal, identically distributed above the diagonal and whose non diagonal entries are centered with variance one. A GOE matrix is a real symmetric matrix whose entries are Gaussian centered, with variance 2 on the diagonal. A GUE matrix is an Hermitian Wigner matrix whose entries are Gaussian centered, with variance 1 on the diagonal and covariance $\mathrm{diag}(1/2, 1/2)$ above the diagonal.

Let us consider an $n \times n$ Wigner matrix $\sqrt{n}X = [x_{i,j}]_{i,j=1}^n$, whose entries laws do not depend on n. The following results are standard ones (see respectively [6, Th. 2.5], [6, Th. 5.2] et [107, 98]).

- The empirical spectral law of X tends to the *semi circle law*, with support $[-2,2]$ and density $\frac{1}{2\pi}\sqrt{4-x^2}$.
- The extreme eigenvalues of X tend to -2 and 2 if and only if $\mathbb{E}[x_{1,1}^2]$ and $\mathbb{E}[|x_{1,2}|^4]$ are finite.

[5] Here, "long" means "a time proportional to p".

- If the laws of the $x_{i,j}$'s are symmetric or with sub-Gaussian tails[6], then for all fixed $k \geq 1$ denoting by
$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$$
the k largest eigenvalues of X, the law of the random vector
$$\{n^{2/3}(\lambda_1 - 2), \ldots\ldots, n^{2/3}(\lambda_k - 2)\}$$
tends to a Tracy-Widom law (either the real or complex one, depending on whether X is real or complex).

INSERT 2.4 – **Tracy-Widom laws**

The classical Fisher-Tippet-Gnedenko Theorem on statistics of extreme values says that for $X_1, \ldots, X_n$ i.i.d. with law $\mathcal{L}$, the maximum $\max\{X_1, \ldots, X_n\}$ can only converge (up to affine transformations), as $n \longrightarrow \infty$, to a Dirac mass or to one of the following laws:

- a Gumbel law (for example when $\mathcal{L}$ is exponential or Gaussian)
- a Fréchet law (when $\mathcal{L}$ has heavy tails),
- a Weibull law (when $\mathcal{L}$ has compact support).

The behavior of the cumulative distribution function of $\mathcal{L}$ on the right of its support determines the asymptotic fluctuations of $\max\{X_1, \ldots, X_n\}$.

Beyond the case where the X_i's are i.i.d., the presence of repulsion or attraction between the X_i's may completely change the asymptotic fluctuations of the maximum and give rise to new distributions at the limit. For example a class of laws called the *Tracy-Widom laws* appears in the following contexts:

- largest eigenvalue of a Wigner or Wishart matrix with sub-Gaussian tails,
- most right particle of a Coulomb gas,
- largest increasing subsequence of a random permutation,
- last passage percolation with geometric or exponential weights in dimension 2,
- ASEP and TASEP,
- polynuclear growth models.

[6]A law μ on $\mathbb{R}$ is said to have a *sub-Gaussian tail* if there exists $\alpha > 0$ such that for all t large enough, $\mu(\mathbb{R}\backslash[-t,t]) \leq e^{-t^\alpha}$.

This family of laws depends on two parameters : a positive integer k (the dimension of the considered vector) and a real number $\beta > 0$ (the inverse of a temperature). For example, the law of the fluctuations of the k largest eigenvalues of a GO(U)E matrix converges to the Tracy-Widom law with parameters k and $\beta = 1$ (resp. $\beta = 2$).

Let us now succinctly describe the definition of the cumulative distribution function F_β of the Tracy-Widom law with parameters $d = 1$ and $\beta = 1$ or 2 (more general descriptions can be found in [114, 115, 96]).

The function F_2 is defined by the Fredholm determinant

$$F_2(s) = \det(I - K_{\mathrm{Ai}})_{L^2([s,+\infty))},$$

where K_{Ai} is the *Airy kernel*, defined by

$$K_{\mathrm{Ai}}(x,y) = \frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}(y)}{x-y} = \int_0^{+\infty} \mathrm{Ai}(x+t)\,\mathrm{Ai}(y+t)\mathrm{d}t.$$

The function F_2 can also be defined by

$$F_2(s) = \exp\left(-\int_s^{+\infty}(x-s)q^2(x)\mathrm{d}x\right),$$

where q is the solution of the *Painlevé II* differential equation $q''(s) = sq(s) + 2q^3(s)$ such that $q(x) \sim \mathrm{Ai}(x)$ at $+\infty$.

The function F_1 can also be defined thanks to q via the formula

$$F_1(s) = \exp\left(-\frac{1}{2}\int_s^{+\infty} q(x) + (x-s)q^2(x)\mathrm{d}x\right).$$

Each of the models enumerated above allows exact computations, letting the Tracy-Widom law appear directly (except the non Gaussian matrices, that we simply compare to Gaussian ones). It seems however likely that these laws might be more universal than up to our current knowledge, appearing in various contexts where a maximum of repulsing variables appears.

2.2 Generalization

We expose here the works developed recently in collaboration with Raj Rao, Alice Guionnet and Mylène Maïda, generalizing the phase transition presented above for Wigner and Wishart matrices to more general matrix models, in the framework presented at Section 2.1.1. We also present the phase transition for eigenvectors and the perturbation of non Hermitian matrices. At last, we relate these results to free probability theory.

2.2.1 Main results

The models

We shall use here the framework outlined at Section 2.1.1 : X is a deterministic $n \times n$ Hermitian random matrix whose eigenvalues are denoted by

$$\lambda_1 \geq \cdots\cdots \geq \lambda_n.$$

We suppose that the empirical spectral law of X converges, as $n \longrightarrow \infty$, to a compactly supported probability measure μ. In order to locate the extreme eigenvalues of X, we also suppose that

$$\lambda_1 \quad \longrightarrow \quad b \qquad \text{and} \qquad \lambda_n \quad \longrightarrow \quad a,$$

where a and b are the lower and upper bounds of the support of μ (under a weaker version of this hypothesis, an important part of what follows stays true (*c.f.* [A14, Rem. 2.13])). The matrix X is supposed to be deterministic, but up to a conditioning, one can apply what follows to the case where X is chosen at random. We perturb X with an Hermitian matrix P whose rank stays bounded as the dimension n grows:

$$\widetilde{X} := X + P \qquad\qquad \text{(additive perturbation)}, \tag{2.4}$$

or

$$\widetilde{X} := (I + P)X \qquad\qquad \text{(multiplicative perturbation)}, \tag{2.5}$$

with

$$P := \sum_{j=1}^{r} \theta_j u_j u_j^*. \tag{2.6}$$

Here r and the real numbers

$$\underbrace{\theta_1 \geq \cdots\cdots \geq \theta_{r_0}}_{>0} \geq \underbrace{\theta_{r_0+1} \geq \cdots\cdots \geq \theta_r}_{<0}$$

are fixed independently of n and the column vectors $u_1, \ldots, u_r \in \mathbb{K}^{n\times 1}$ are :

- either the column vectors of a matrix $\frac{1}{\sqrt{n}} \begin{bmatrix} g_{1,1} & \cdots & g_{1,r} \\ \vdots & & \vdots \\ g_{n,1} & \cdots & g_{n,r} \end{bmatrix}$, where the $g_{i,j}$'s are i.i.d. centered variables with variance 1, satisfying a log-Sobolev inequality[7],

- either obtained from the column vectors of the matrix above by the Schmidt orthonormalization process.

[7] *Log-Sobolev inequalities* allow concentration in large dimensions (see [2]).

Both of these ways to define P are formalization of Hypothesis 2.4 of Section 2.1.1. They define two models, called respectively the *i.i.d. perturbations model* and the *orthonormalized perturbations model*, which are in fact asymptotically very close, the u_j's of the i.i.d. model being almost orthornormalized for $n \gg 1$. Notice that in the orthonormalized model, if the $g_{i,j}$'s are Gaussian, the the collection $\{u_1, \ldots, u_r\}$ is uniformly distributed on the manifold of r-uplets of orthonormal vectors.

Let us add a last hypothesis: in the case of a multiplicative perturbation, X is supposed to be positive and $\mu \neq \delta_0$.

Additive perturbations : $\widetilde{X} = X + P$

Let us introduce the Cauchy transform

$$G_\mu(z) := \int \frac{\mathrm{d}\mu(x)}{z - x}$$

of the limit empirical spectral distribution μ of X, defined for z out of the support of μ, for example if $z < a$ or $z > b$. We shall see that the isolated eigenvalues of $\widetilde{X}$, when they exist, are close to the solutions of the equations $G_\mu(z) = \theta_i^{-1}$.

Indeed, setting, with the convention $\frac{1}{\pm\infty} = 0$,

$$\overline{\theta} := \frac{1}{\lim_{z \downarrow b} G_\mu(z)} \geq 0, \qquad \underline{\theta} := \frac{1}{\lim_{z \uparrow a} G_\mu(z)} \leq 0 \tag{2.7}$$

and, for $\theta \in \mathbb{R} \backslash \{0\}$,

$$\rho_\theta := \begin{cases} G_\mu^{-1}(1/\theta) & \text{if } \theta \in (-\infty, \underline{\theta}) \cup (\overline{\theta}, +\infty), \\ a & \text{if } \theta \in [\underline{\theta}, 0), \\ b & \text{if } \theta \in (0, \overline{\theta}], \end{cases} \tag{2.8}$$

we have the following theorem [A14, Th. 2.1], [A17, Th. 1.3].

Theorem 2.8 *For all $i \in \{1, \ldots, r_0\}$, we have*

$$\widetilde{\lambda}_i \quad \longrightarrow \quad \rho_{\theta_i} \tag{2.9}$$

and for all $i \in \{r_0 + 1, \ldots, r\}$,

$$\widetilde{\lambda}_{n-r+i} \quad \longrightarrow \quad \rho_{\theta_i}. \tag{2.10}$$

Moreover, for all fixed $i > r_0$ (resp. for all fixed $i \geq r - r_0$),

$$\widetilde{\lambda}_i \quad \longrightarrow \quad b \qquad \textit{(resp. } \widetilde{\lambda}_{n-i} \quad \longrightarrow \quad a\textit{)}. \tag{2.11}$$

Moreover, the fluctuations of the extreme eigenvalues of $\widetilde{X}$ are given by the following theorem, which compiles theorems 3.2, 3.4, 4.3, 4.4 et 4.5 of [A17]. Part (1) is about the isolated eigenvalues, and Part (2) is about the eigenvalues with limit a or b.

Theorem 2.9 *Under a few supplementary technical hypotheses, we have the following results.*

(1) Let $\underbrace{\alpha_1 > \cdots > \alpha_{q_0}}_{>0} > \underbrace{\alpha_{q_0+1} > \cdots > \alpha_q}_{<0}$ *be the pairwise distinct values of the* θ_i*'s such that* $\rho_{\theta_i} \notin \{a, b\}$ *and for all* $j = 1, \ldots, q$*, let*

$$I_j := \{i = 1, \ldots, r\,;\, \theta_i = \alpha_j\}.$$

Then the law of the random vector

$$\left\{\sqrt{n}(\widetilde{\lambda}_i - \rho_{\alpha_j})\,;\, i \in I_j\right\}_{1 \le j \le q_0} \cup \left\{\sqrt{n}(\widetilde{\lambda}_{n-r+i} - \rho_{\alpha_j})\,;\, i \in I_j\right\}_{q_0+1 \le j \le q}$$

converges to the one of the vector

$$\{\textit{ordered eigenvalues of } c_{\alpha_j} M_j\}_{1 \le j \le q}, \tag{2.12}$$

where the matrices M_j *are independent GO(U)E* $k_j \times k_j$ *matrices and the* c_{α_j} *are constants, depending only on* μ *and on* α_j *(see [A17, Eq. (6)]).*

(2)(2.1) If no θ_i *is critical (i.e. equal to* $\underline{\theta}$ *or* $\overline{\theta}$*), then there exists* $\varepsilon > 0$ *small such that with large probability, the eigenvalues of* $\widetilde{X}$ *with limit* a *or* b *are within* $n^{-1+\varepsilon}$ *of the spectrum of* X*.*

(2.2) If, moreover, we consider the i.i.d. perturbations model or if $r = 1$*, then, denoting by* p_+ *and* p_- *the number of eigenvalues of* $\widetilde{X}$ *with limit respectively* $> b$ *and* $< a$*, we have, for all fixed* $i \ge 1$*,*

$$n^{1-\varepsilon}(\widetilde{\lambda}_{p_+ + i} - \lambda_i) \quad \longrightarrow \quad 0 \tag{2.13}$$

and

$$n^{1-\varepsilon}(\widetilde{\lambda}_{n+1-(p_- + i)} - \lambda_{n+1-i}) \quad \longrightarrow \quad 0. \tag{2.14}$$

Let us make precise the technical hypotheses needed here. The supplementary hypotheses needed to the proof of (1) are about the rate of convergence of the empirical spectral law of X to μ, (which has to be $\ll n^{-1/2}$) and about the fourth cumulant of the $g_{i,j}$'s. The supplementary hypotheses needed to the proof of (2) are about the spacings of the extreme eigenvalues of X. All of these hypotheses are satisfied almost surely when X is random and chosen according to a standard model (Wigner, Wishart, Coulomb gas...). Moreover, in these models, the ε of (2) is small enough so that the $\widetilde{\lambda}_i$'s with limit a or b inherit the Tracy-Widom fluctuations of the λ_i's (see the Section "Examples" in [A17]).

Instead of detailing these examples, let us make both previous theorems more concrete by showing how they allow to explain Figure 2.1 presented at the beginning of this chapter.

In this figure, we perturb a GUE matrix with a rank-one matrix P whose sole non zero eigenvalue is denoted by θ. The law of X is invariant under unitary conjugation, so this fits with the model with orthonormalized perturbations, up to a conditioning by X. By the formula of the Cauchy transform of the semi-circle law, we know that the threshold $\overline{\theta}$ of the phase transition equals 1 and that $\rho_\theta = \theta + \frac{1}{\theta}$ when $\theta > b = 2$.

- In the left picture of Figure 2.1, $\theta = 0.5 < \overline{\theta}$ and as predicted by (2.9), $\widetilde{\lambda}_1 \approx b = 2$, whereas in the right one, $\theta = 1.5 > \overline{\theta}$, which indeed implies $\widetilde{\lambda}_1 \approx \rho_\theta = 2.17$ and $\widetilde{\lambda}_2 \approx b$, in accordance with (2.9) and (2.11).

- Moreover, in the left picture, we have, for all i, $\widetilde{\lambda}_i \approx \lambda_i$, with some deviations

  $$|\widetilde{\lambda}_i - \lambda_i| \ll \text{deviation of } \lambda_i \text{ from its limit 2.}$$

  In the same way, in the right picture, for all i, $\widetilde{\lambda}_{i+1} \approx \lambda_i$, with some deviations

  $$|\widetilde{\lambda}_{i+1} - \lambda_i| \ll \text{deviation of } \lambda_i \text{ from its limit 2.}$$

  Both of these observations are conform to (2.13).

- At last, here, $n = 2.10^3$, and Equation (6) of [A17] gives the formula of the c_α in (2.12) : $c_\alpha^2 = 1 - \alpha^{-2}$. In the right picture, we have $\widetilde{\lambda}_1 \approx 2.167$, which gives $\frac{\sqrt{n}(\widetilde{\lambda}_1 - \rho_\theta)}{c_\theta} \approx 0.040$, a reasonable value for a standard Gaussian variable.

Let us now give a sketch of the proofs of these theorems. For all z out of the spectrum of X, we have

$$\det(z - \widetilde{X}) = \det(z - X - P) = \det(z - X)\det(1 - (z - X)^{-1}P).$$

From there, using the formula of P given in (2.6) and the identity $\det(1+AB) = \det(1+BA)$, true even when A, B are not square matrices, one characterizes the eigenvalues of $\widetilde{X}$ which are out of the spectrum of X in the following way: these are the z's such that

$$\det\left(\begin{bmatrix}\theta_1^{-1} & & \\ & \ddots & \\ & & \theta_r^{-1}\end{bmatrix} - \begin{bmatrix}u_1^*(z-X)^{-1}u_1 & \cdots & u_1^*(z-X)^{-1}u_r \\ \vdots & & \vdots \\ u_r^*(z-X)^{-1}u_1 & \cdots & u_r^*(z-X)^{-1}u_r\end{bmatrix}\right) = 0. \tag{2.15}$$

This is a characterization of the eigenvalues of $\widetilde{X}$ via an $r\times r$ (and not $n\times n$) determinant.

This characterization is enough to prove Theorem 2.8. Indeed, it is enough to control Equation (2.15) for the z's whose distance to $[a,b]$ has order 1, which can easily be done thanks to concentration inequalities like Hanson-Wright Theorem [67]. They imply that for $n \gg 1$,

$$\begin{bmatrix}u_1^*(z-X)^{-1}u_1 & \cdots & u_1^*(z-X)^{-1}u_r \\ \vdots & & \vdots \\ u_r^*(z-X)^{-1}u_1 & \cdots & u_r^*(z-X)^{-1}u_r\end{bmatrix} \approx \begin{bmatrix}G_\mu(z) & & \\ & \ddots & \\ & & G_\mu(z)\end{bmatrix}, \tag{2.16}$$

which makes the characterization of the limits of the isolated eigenvalues of $\widetilde{X}$ as the solutions of the equations $G_\mu(z) = \theta_i^{-1}$ quite obvious.

Let us now explain how GO(U)E matrices appear in Part (1) of Theorem 2.9. We suppose for example that

$$\theta_1 = \theta_2 > \theta_3 \geq \theta_4 \geq \cdots \geq \theta_r$$

and that $\theta_1 > \overline{\theta}$, so that $G_\mu(\rho_{\alpha_1}) = \frac{1}{\theta_1} = \frac{1}{\theta_2}$. In order to determine the fluctuations of the vector $(\widetilde{\lambda}_1, \widetilde{\lambda}_2)$ around its limit $(\rho_{\alpha_1}, \rho_{\alpha_1})$, we shall Taylor-expand the left hand term of (2.15) for $z = \rho_{\alpha_1} + \frac{x}{\sqrt{n}}$. The Central Limit Theorem and a fine analysis of the orthonromalization process allow to push the approximation of (2.16) further and give, for $z = \rho_{\alpha_1} + \frac{x}{\sqrt{n}}$,

$$\begin{aligned}&\begin{bmatrix}\theta_1^{-1} & & \\ & \ddots & \\ & & \theta_r^{-1}\end{bmatrix} - \begin{bmatrix}u_1^*(z-X)^{-1}u_1 & \cdots & u_1^*(z-X)^{-1}u_r \\ \vdots & & \vdots \\ u_r^*(z-X)^{-1}u_1 & \cdots & u_r^*(z-X)^{-1}u_r\end{bmatrix} \\ &\approx \begin{bmatrix}\frac{1}{\sqrt{n}}(x-N) & -\frac{1}{\sqrt{n}}N'' & & & \\ -\frac{1}{\sqrt{n}}N'' & \frac{1}{\sqrt{n}}(x-N') & & & \\ & & \theta_3^{-1}-\theta_1^{-1} & & \\ & & & \ddots & \\ & & & & \theta_r^{-1}-\theta_1^{-1}\end{bmatrix},\end{aligned}$$

where the matrix $\begin{bmatrix}N & N'' \\ N'' & N'\end{bmatrix}$ is a GO(U)E one. Considering the determinants of these matrices, we deduce, up to an approximation, that $z = \rho_{\alpha_1} + \frac{x}{\sqrt{n}}$ is a solution of the equation (2.15) if and only if x is a solution of Equation

$$\det\left(x - \begin{bmatrix} N & N'' \\ N'' & N' \end{bmatrix}\right) = 0.$$

Part (2) of Theorem 2.9, is proved by showing that when no θ_i is critical, for z closed enough to a or b, Equation (2.15) can only be verified if z is very close to one of the λ_i's. To simplify, let us suppose for example that $r = 1$, consider the model with i.i.d. perturbations and restrict ourselves to the vicinity of b. We want to prove that according to whether $\theta_1 < \overline{\theta}$ or $\theta_1 > \overline{\theta}$, we have

$$\widetilde{\lambda}_1 \approx \lambda_1\,,\ \widetilde{\lambda}_2 \approx \lambda_2\,,\ \widetilde{\lambda}_3 \approx \lambda_3\,,\ \ldots\ldots$$

or

$$\widetilde{\lambda}_2 \approx \lambda_1\,,\ \widetilde{\lambda}_3 \approx \lambda_2\,,\ \widetilde{\lambda}_4 \approx \lambda_3\,,\ \ldots\ldots,$$

the symbol $\approx$ meaning here *equal up to an $n^{-2/3-\varepsilon}$ error*, with $\varepsilon > 0$ small.

In the particular case where $r = 1$, Equation (2.15) can be written

$$\frac{1}{n}\sum_{j=1}^{n} \frac{g_j^2}{z - \lambda_j} = \frac{1}{\theta_1}. \tag{2.17}$$

The left hand term of this equation being a decreasing function of z, with limit $+\infty$ at the right of each λ_j and $-\infty$ at the left of each λ_j, the equation admits exactly one solution in each interval $(\lambda_{j+1}, \lambda_j)$, what Weyl's interlacing inequalities already said to us, up to the equality case. If $\theta_1 > 0$ (resp. $\theta_1 < 0$), as the left hand term of (2.17) vanishes at $\pm\infty$, we know that the equation admits a solution $> \lambda_1$ (resp. $< \lambda_n$). Thus, we have

$$\widetilde{\lambda}_1 > \lambda_1 > \widetilde{\lambda}_2 > \lambda_2 > \cdots\cdots \ (\text{resp. } \lambda_1 > \widetilde{\lambda}_1 > \lambda_2 > \widetilde{\lambda}_2 > \cdots\cdots).$$

Let us focus for example on the interval $[\lambda_2, \lambda_1]$ and prove that, according to whether we are in one of the two following alternative situations

$$0 < \theta_1 < \overline{\theta} \qquad \text{or} \qquad \theta_1 \in \mathbb{R}\backslash[0, \overline{\theta}], \tag{2.18}$$

the solution of Equation (2.17) is this interval is situated respectively to the right or the left of the sub-interval $[\lambda_2 + \frac{1}{n^{2/3+\varepsilon}}, \lambda_1 - \frac{1}{n^{2/3+\varepsilon}}]$. Let us first notice that it suffices to prove that on the interval $[\lambda_2 + \frac{1}{n^{2/3+\varepsilon}}, \lambda_1 - \frac{1}{n^{2/3+\varepsilon}}]$,

$$\frac{1}{n}\sum_{j=1}^{n} \frac{g_j^2}{z - \lambda_j} \approx \frac{1}{\overline{\overline{\theta}}}. \tag{2.19}$$

Indeed, according to the one of the two alternatives of (2.18) which is verified, we have respectively

$$\frac{1}{\theta_1} > \frac{1}{\overline{\overline{\theta}}} \qquad \text{or} \qquad \frac{1}{\theta_1} < \frac{1}{\overline{\overline{\theta}}}$$

and the decreasingness of the function $z \longmapsto \frac{1}{n}\sum_{j=1}^n \frac{g_j^2}{z-\lambda_j}$ allows then to locate the solution of Equation (2.17).

Let us now explain quickly how one obtains Approximation (2.19). Let z be such that

$$\lambda_2 + n^{-2/3-\varepsilon} \le z \le \lambda_1 - n^{-2/3-\varepsilon} \tag{2.20}$$

with $\varepsilon > 0$ small. We decompose the left term of (2.19) in the following way :

$$\frac{1}{n}\sum_{j=1}^{n} \frac{g_j^2}{z-\lambda_j} = \frac{1}{n}\sum_{j=1}^{m} \frac{g_j^2}{z-\lambda_j} + \frac{1}{n}\sum_{j=m+1}^{n} \frac{g_j^2}{z-\lambda_j}. \tag{2.21}$$

First, when $m \gg 1$ and $n-m \gg 1$, a concentration result of the type Hanson-Wright [67] allows us to replace the g_j^2's by 1 in both sums above :

$$\frac{1}{n}\sum_{j=1}^{m} \frac{g_j^2}{z-\lambda_j} \approx \frac{1}{n}\sum_{j=1}^{m} \frac{1}{z-\lambda_j} \qquad \text{et} \qquad \frac{1}{n}\sum_{j=m+1}^{n} \frac{g_j^2}{z-\lambda_j} \approx \frac{1}{n}\sum_{j=m+1}^{n} \frac{1}{z-\lambda_j}. \tag{2.22}$$

For z satisfying (2.20), choosing $m = n^{1/3-\varepsilon'}$, with $\varepsilon' > \varepsilon$ small, we can neglect the proximity effects due to the first λ_j's, and obtain

$$\frac{1}{n}\sum_{j=m+1}^{n} \frac{1}{z-\lambda_j} \approx \lim_{z\downarrow b} G_\mu(z) = \frac{1}{\overline{\theta}}. \tag{2.23}$$

Besides, Equation (2.20) gives

$$\left|\frac{1}{n}\sum_{j=1}^{m} \frac{1}{z-\lambda_j}\right| \le \frac{1}{n}\sum_{j=1}^{m} \frac{1}{n^{-2/3-\varepsilon}} = n^{\varepsilon-\varepsilon'} \quad \underset{n\to\infty}{\longrightarrow} \quad 0. \tag{2.24}$$

Joining (2.21), (2.22), (2.23) and (2.24), we see that for $z \in [\lambda_2 + \frac{1}{n^{2/3+\varepsilon}}, \lambda_1 - \frac{1}{n^{2/3+\varepsilon}}]$,

$$\frac{1}{n}\sum_{j=1}^{n} \frac{g_j^2}{z-\lambda_j} \approx \frac{1}{\overline{\theta}}.$$

Eigenvectors

Let us now present our results about the eigenvectors. For $i \in \{1,\ldots,r_0\}$, denoting by $\widetilde{u}_i$ a unitary eigenvector of $\widetilde{X}$ associated to $\widetilde{\lambda}_i$, the scalar product

$$\langle \widetilde{u}_i, u_i \rangle$$

contains a hint of the "trace" (in the common sense of the term) that the deformation P has let on X. We are going to see that this scalar product satisfies the same phase transition, *i.e.* that to each θ_i beyond the threshold, one can associate an eigenvalue with an eigenvector that has a positive component according to u_i, which does not happen when θ_i is below the threshold.

The following theorem, [A14, Th. 2.2 et 2.3], formalizes these ideas. It has only been proved for the model with orthonormalized perturbations, but can be easily extended to the one with i.i.d. perturbations. For F a subspace of a Hilbert space and x a vector of this space, we denote by $|\langle x, F\rangle|$ the norm of the orthogonal projection of x onto F.

Theorem 2.10 *(1) Let i_0 be such that $\theta_{i_0} > \overline{\theta}$ and $\widetilde{u}_{i_0}$ a unit eigenvector of $\widetilde{X}$ associated to $\widetilde{\lambda}_{i_0}$. Then as $n \longrightarrow \infty$,*

$$|\langle \widetilde{u}_{i_0}, \mathrm{Span}\{u_i\,;\, \theta_i = \theta_{i_0}\}\rangle|^2 \quad \longrightarrow \quad \frac{-1}{\theta_{i_0}^2 G_\mu'(\rho_{\theta_{i_0}})} > 0 \tag{2.25}$$

and

$$|\langle \widetilde{u}_{i_0}, \mathrm{Span}\{u_i\,;\, \theta_i \neq \theta_{i_0}\}\rangle| \quad \longrightarrow \quad 0.$$

(2) Suppose that $r = 1$, that the sole non zero eigenvalue θ of P is below the threshold $\overline{\theta}$ and that $\lim_{z\downarrow b} G_\mu'(z) = -\infty$, as it is the case in most of the classical models. Let $\widetilde{u}$ be a unit eigenvector of $\widetilde{X}$ associated to $\widetilde{\lambda}_1$. Then

$$|\langle \widetilde{u}, \ker(\theta - P)\rangle| \quad \longrightarrow \quad 0.$$

Multiplicatives perturbations: $\widetilde{X} = (I + P)X$

In the case of multiplicative perturbations, everything which has been seen for additive perturbations stays true (for both the largest and the smallest eigenvalues), up to a replacement of the Cauchy transform by the *T-transform*

$$T_\mu(z) := \int_{x\in\mathbb{R}} \frac{x}{z-x} \mathrm{d}\mu(x)$$

(only Formula (2.25) is changed, see [A14, Th. 2.8]). The empirical covariance matrices are of course the standard application. However, the hypothesis needed to apply these results to their smallest eigenvalues have not been verified (they have been verified for their largest eigenvalues).

The key of the proofs is an analogue of Equation (2.15): a real number z out of the spectrum of X is an eigenvalue of $\widetilde{X}$ if and only if

$$\det\left(\begin{bmatrix}\theta_1^{-1} & & \\ & \ddots & \\ & & \theta_r^{-1}\end{bmatrix} - \begin{bmatrix} u_1^* X(z-X)^{-1}u_1 & \cdots & u_1^* X(z-X)^{-1}u_r \\ \vdots & & \vdots \\ u_r^* X(z-X)^{-1}u_1 & \cdots & u_r^* X(z-X)^{-1}u_r\end{bmatrix}\right) = 0. \tag{2.26}$$

2.2.2 Singular values of finite rank perturbations of non Hermitian matrices

One of the benefits of the method relying on Equation (2.15) is that it can be adapted to other models, as we saw above with multiplicative perturbations. In the paper [A15], co-written with Raj Rao, we prove that the extreme singular values of non Hermitian (possibly rectangular) matrices satisfy the same kind of phase transition when submitted to finite rank perturbations.

The model is analoguous to the one presented in Section 2.2.1, except that X is an $n \times p$ matrix, with $n, p \longrightarrow \infty$ in such a way that $n/p \longrightarrow c \in [0, 1]$. The hypothesis relative to the convergence of the empirical spectral distribution of X to μ is replaced by the convergence of the *empirical singular values distribution* of X to a compactly supported law μ and the hypothesis relative to the convergence of the extreme eigenvalues is replaced by the convergence of the *extreme singular values* of X to the bounds a and b of the support of μ. At last, P is defined by the formula

$$P = \sum_{j=1}^{r} \theta_j u_j v_j^*,$$

where the u_j's and the v_j's are respectively $n\times 1$ and $p\times 1$ column vectors and the collections $\{u_j\,;\, j = 1, \ldots, r\}$ and $\{v_j\,;\, j = 1, \ldots, r\}$ are independent, each of them being defined according to one of the processes presented above (*i.i.d.* and *orthonormalized* perturbations).

We introduce the integral transform

$$D_\mu(z) = \left[\int_{x\in\mathbb{R}} \frac{z}{z^2 - t^2} \mathrm{d}\mu(z)\right] \times \left[\int_{x\in\mathbb{R}} \frac{z}{z^2 - t^2} \mathrm{d}\widetilde{\mu}_X(z)\right], \tag{2.27}$$

where $\widetilde{\mu}_X = c\mu_X + (1-c)\delta_0$. The results of the paper [A15] show that the extreme singular values of $\widetilde{X}$ satisfy the same kind of phase transition as the previous ones. However, when the limit ratio c is < 1, certain technical difficulties restrict the results to the largest singular values. Besides, the fluctuations of the singular values with limit a or b have not been studied.

For example, if X is an $n \times p$ matrix with i.i.d. Gaussian entries with variance $1/p$, the empirical singular values distribution of X tends to the law

$$\frac{\sqrt{(b^2 - x^2)(x^2 - a^2)}}{\pi c x} \mathbb{1}_{x\in[a,b]} \mathrm{d}x,$$

where $a = 1 - \sqrt{c}$, $b = 1 + \sqrt{c}$. In this case, each $\theta_i > c^{1/4}$ gives rise to a singular value of

$\widetilde{X}$ with limit out of the support of μ, equal to

$$\frac{\sqrt{(1+\theta_i^2)(c+\theta_i^2)}}{|\theta_i|},$$

whereas the θ_i's which are $\le c^{1/4}$'s do not cause any isolated singular value.

Another example: the one where X is an $n \times n$ unitary matrix with Haar distribution. Then all singular values of X are equal to 1 and each θ_i gives rise to an isolated singular value of $\widetilde{X}$ on the right of 1 with limit

$$\frac{\theta_i + \sqrt{\theta_i^2+4}}{2}$$

and to an isolated singular value of $\widetilde{X}$ on the left of 1 with limit

$$\frac{-\theta_i + \sqrt{\theta_i^2+4}}{2}.$$

More details on both of these examples are given to Section 3 of [A15].

At last, let us mention that the results of both papers [A14, A15] have been the base of the construction of a parameters estimation algorithm by Hachem, Loubaton, Mestre, Najim and Vallet [66].

2.2.3 Link with free probability theory

It has to be noticed that the phase transitions brought to light here (the one relative to additive perturbations of Hermitian matrices, the one relative to multiplicative perturbations of Hermitian matrices and the one relative to additive perturbations of non Hermitian, possibly rectangular, matrices) are all governed by functions playing a role in the respective convolutions $\boxplus$, $\boxtimes$ and $\boxplus_c$. Indeed,

- the function G_μ^{-1} is linked to the R-transform R_μ of μ by formula

$$G_\mu^{-1}(z) = R_\mu(z) + \frac{1}{z},$$

- the function T_μ^{-1} is linked to the S-transform S_μ of μ by the formula

$$T_\mu^{-1}(z) = \frac{z+1}{zS_\mu(z)},$$

- the function D_μ^{-1} is linked to the rectangular R-transform with ratio c of μ by the formula
$$\left(D_\mu^{-1}(z)\right)^2 = \frac{(cC_\mu^{(c)}(z)+1)(C_\mu^{(c)}(z)+1)}{z}.$$

Hence it seems that free probability theory is at the heart of this phase transition. That can be appear as surprising, because until now, the link between free probability and random matrices only concerned *macroscopic* issues, *i.e.* the whole spectrum, and did not allow to locate any isolated eigenvalue.

One can explain these coincidences. The spectrum of a large matrix X deformed by a finite rank matrix P with spectrum
$$\theta_1,\ldots,\theta_r,\underbrace{0,\ldots\ldots\ldots\ldots,0}_{n-r\text{ null eigenvalues}}$$
can be understood as the limit, as $\varepsilon \longrightarrow 0$, of the free convolution of the spectral law of X with the law
$$(1-\varepsilon)\delta_0 + \varepsilon\frac{\delta_{\theta_1}+\cdots+\delta_{\theta_r}}{r}.$$

Let us consider for example the case of an additive perturbation of an Hermitian matrix X by a rank one matrix with sole non zero eigenvalue θ. As the dimension n gets large, the empirical spectral distribution of $\widetilde{X}$ gets close to
$$\mu_\varepsilon := \mu \boxplus ((1-\varepsilon)\delta_0 + \varepsilon\delta_\theta) \tag{2.28}$$
with $\varepsilon = \frac{1}{n}$.

Lemma 2.11 *As $\varepsilon \longrightarrow 0$, the law μ_ε admits a Taylor expansion*
$$\mu_\varepsilon = \mu + \varepsilon\,\mathfrak{m} + o(\varepsilon),$$
where $\mathfrak{m}$ is a null mass signed measure which satisfies the following phase transition (with the notations introduced at (2.7) and (2.8)) :

- *if $\underline{\theta} \le \theta \le \overline{\theta}$, then the support of $\mathfrak{m}$ is contained in the one of μ,*
- *if $\theta < \underline{\theta}$ or $\theta > \overline{\theta}$, $\mathfrak{m}$ can be written as the sum of a measure with support contained in the one of μ and of a Dirac mass at ρ_θ, with weight 1.*

Before proving the lemma, notice that it gives an explanation to the BBP phase transition as far as *positions* of extreme eigenvalues are concerned (the *fluctuations* issue is more delicate, but we shall see later that one can however say something about it). Indeed, up to the approximation of the empirical spectral distribution of $\widetilde{X}$ by μ_ε with $\varepsilon = 1/n$, it allows to understand the apparition, in this spectral law, of a Dirac mass with weight $1/n$ (*i.e.* of an eigenvalue of $\widetilde{X}$) at ρ_θ when $\theta \notin [\underline{\theta}, \overline{\theta}]$.

Proof of Lemma 2.11. By [A7, Lem. 2.11] and [8, Lem. 2.17], proving this lemma amounts to prove that the Cauchy transform of μ_ε admits the Taylor expansion

$$G_{\mu_\varepsilon}(z) = G_\mu(z) + \varepsilon\, \mathfrak{g}(z) + o(\varepsilon),$$

where $\mathfrak{g}$ is a function satisfying the following properties :

- if $\underline{\theta} \le \theta \le \overline{\theta}$, then $\mathfrak{g}$ is analytic on the complementary of the support of μ with real values on the real line,
- if $\theta < \underline{\theta}$ or $\theta > \overline{\theta}$, then $\mathfrak{g}$ is a meromorphic function on the complementary of the support of μ with with real values on the real line and a unique pole, located at ρ_θ and with residue 1.

The R-transform of μ_ε is $R_\mu + R_{(1-\varepsilon)\delta_0 + \varepsilon\delta_\theta}$. The function $R_{(1-\varepsilon)\delta_0 + \varepsilon\delta_\theta}$ can be computed and admits the Taylor expansion

$$R_{(1-\varepsilon)\delta_0 + \varepsilon\delta_\theta}(z) = \frac{\varepsilon\theta}{1 - \theta z} + o(\varepsilon),$$

hence

$$R_{\mu_\varepsilon} = R_\mu(z) + \frac{\varepsilon\theta}{1 - \theta z} + o(\varepsilon).$$

By formula $G_{\mu_\varepsilon}^{-1}(z) = R_{\mu_\varepsilon}(z) + \frac{1}{z}$, we deduce

$$G_{\mu_\varepsilon}(z) = G_\mu(z) - \varepsilon \frac{\theta G_\mu'(z)}{1 - \theta G_\mu(z)} + o(\varepsilon),$$

which allows to conclude immediately. □

The theory of *Second Order Freeness*, developed by Mingo, Nica, Speicher, Śniady and Collins [83, 84, 85, 40] allows to understand the asymptotic fluctuations of the spectral law of sums or products of random matrices. This theory explains, at least at the heuristic level, why the law of the eigenvalues of GO(U)E matrices is the limit law of the joint fluctuations of collections of eigenvalues of $\widetilde{X}$ with common limit located out of the support of μ. We do not detail it here, but the approach is the one already used above.

Besides, in the recent prepublications [35, 30], Capitaine, Donati-Martin, Féral and Février have proved that in the case where X is the sum of a Wigner or Wishart matrix and a deterministic matrix, the phase transitions presented in Theorems 2.8 et 2.10 can be understood via the *subordination* (presented in this text at Section 3.3.1).

Up to my knowledge, Part (2) of Theorem 2.9 (about the eigenvalues of $\widetilde{X}$ which "stick" to the bulk) cannot be related to free probability theory.

2.3 Large deviations for deformed matrix models

Large deviations principles for random matrices are quite seldom. The first ones are those of Ben Arous and Guionnet [11] for the GO(U)E matrices (which can be generalized to Coulomb matrices [51, 2]) and those of Ben Arous and Zeitouni for Gaussian non symmetric matrices [13]. At the same time, Hiai et Petz have also studied large deviations for Wishart and uniform unitary matrices [68]. Cabanal-Duvillard, Guionnet, Zeitouni, Capitaine and Biane, in [29, 63, 23], have established LDP for matricial processes, linked to spherical integrals and to the free entropy of Voiculescu. Besides, a LDP for the largest eigenvalues of a GO(U)E matrix is proved in [2], based on the works of Ben Arous, Dembo and Guionnet on the spin glasses [10]. As I an writing this memoir, Chaterjee et Varadhan make their preprint [38] public, devoted to large deviations of random matrices in a quite different context. At last, deviations of the largest eigenvalue of a GO(U)E matrix perturbed by the addition of a rank one matrix, have been studied by Maïda in [80] and large deviations of the largest eigenvalues of a Wishart matrix of the type MM^*, with M an $n \times p$ matrix such that $n/p \longrightarrow 0$, have been studied by Fey, van der Hofstad and Klok in [53].

Most of these LDPs (more specifically: all but the three last ones) are about some matrices with density proportional to $e^{-n \operatorname{tr} V(M)}$, whose eigenvalues distribution can be exactly computed. For more general matrices, as non Gaussian Wigner matrices, the rate function of a LDP might depend on the laws of the entries, and we do not even have any hint on the definition of such a function.

In the paper [A18], co-written with Alice Guionnet and Mylène Maïda, we study the large deviations of the extremes eigenvalues in the model $\widetilde{X} = X + P$ of Section 2.2.1, and also in a quite different model, where we do not suppose anymore that the extreme eigenvalues of X tend to a and b. In a first time we focus on the case where X is deterministic and diagonal, having exactly p_+ eigenvalues on the right of the support of its limit empirical spectral distribution μ. We denote by r_0 the number of j's such that $\theta_j > 0$, as previously.

Theorem 2.12 *The $r_0 + p_+$ largest eigenvalues of $\widetilde{X}$ satisfy a LDP with scale n and with a good rate function.*

The rate function is not explicit generally, however, one can say enough of it to deduce a new proof of Theorem 2.8 (see [A18, Rem. 6.5]).

From this theorem, one can deduce some results about the case where X is chosen at random if we control the deviations of X well enough. For example, (*c.f.* [A18, Th. 2.13]), it allows to prove a LDP with scale n and good rate function for the the k largest eigenvalues (k being any integer) of $\widetilde{X}$ when X is chosen at random in a so-called classical ensemble, *i.e.* with a distribution of the type

$$\frac{1}{Z_n^\beta} e^{-n \operatorname{tr} V(X)} \mathrm{d}^\beta X,$$

where $\mathrm{d}^\beta X$ is the standard Lebesgue measure in the space of symmetric real matrices ($\beta = 1$) or Hermitian matrices ($\beta = 2$) and $V : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is a continuous function $\geq \beta \log(x)$. Deformed Gaussian Wishart matrices and GO(U)E matrices are examples of applications of this work.

The proof of Theorem 2.12 relies again on the representation of the eigenvalues of $\widetilde{X}$ as the solutions of Equation (2.15): any real number z out of the spectrum of X is an eigenvalue of $\widetilde{X}$ if and only if

$$\det \left(\begin{bmatrix} \theta_1^{-1} & & \\ & \ddots & \\ & & \theta_r^{-1} \end{bmatrix} - \begin{bmatrix} u_1^*(z-X)^{-1}u_1 & \cdots & u_1^*(z-X)^{-1}u_r \\ \vdots & & \vdots \\ u_r^*(z-X)^{-1}u_1 & \cdots & u_r^*(z-X)^{-1}u_r \end{bmatrix} \right) = 0.$$

We first establish a LDP for a function which, up to a random but well controlled factor, is the left term of the previous equation, then deduce a LDP for the zeros of this function. This last step, a kind of application of the contraction principle, is quite delicate because the application which, to a function, associates its zeros is not continuous in the space we are working in.

Chapter 3

Infinite divisibility and limit theorems for free convolutions, applications

In this chapter, we recall some basic results on infinitely divisible laws for the classical convolution $*$, that one can find for example in [58, 100], and then we expose the generalization of this theory to the free additive convolution, due to Voiculescu, Bercovici et Pata, bringing to light a deep relation between the convolutions $*$ and $\boxplus$. We shall then expose the matricial interpretation I gave in [A2] and the rectangular version of infinite divisibility, that I developped during my PhD Thesis.

At last, we give applications of this theory, first to the regularization of measures, and secondly bringing to light a phenomenon of repulsion of the singular values faraway from zero.

3.1 The Bercovici-Pata bijection between $*$- and $\boxplus$-infinitely divisible law, matricial interpretation

One can define infinitely divisible laws as the limit laws of sums of i.i.d. variables : a law μ on $\mathbb{R}$ is said to be $*$-*infinitely divisible* if there exists a sequence (k_n) of integers tending to infinity and a sequence (ν_n) of laws such that as $n \longrightarrow \infty$,

$$\underbrace{\nu_n * \cdots * \nu_n}_{k_n \text{ times}} \longrightarrow \mu. \tag{3.1}$$

In this case, there exists a unique family $(\mu^{*t})_{t\in[0,\infty)}$ of laws, starting at δ_0 such that $\mu^{*1} = \mu$ and which is a semigroup for the convolution $*$.

These laws are characterized and classified thanks to their Fourier transforms: μ is

$*$-infinitely divisible if and only if there exists $\gamma \in \mathbb{R}$ and a positive measure σ on $\mathbb{R}$ such that the Fourier transform of μ can be written $\int_{t\in\mathbb{R}} e^{it\xi} \mathrm{d}\mu(t) = e^{\Psi_\mu(\xi)}$, with

$$\Psi_\mu(\xi) = i\gamma\xi + \int_{\mathbb{R}} \underbrace{(e^{it\xi} - 1 - \frac{it\xi}{t^2+1})\frac{t^2+1}{t^2}}_{:=-\frac{\xi^2}{2} \text{ for } t=0} \mathrm{d}\sigma(t). \tag{3.2}$$

Moreover, in this case, the pair (γ, σ), unique, is called the *Lévy pair*[1] of μ and the law μ is denoted by $\nu_*^{\gamma,\sigma}$.

Voiculescu and Bercovici have proved in [15, 16] that everything that precedes stays true if one replaces the classical convolution $*$ by the free additive convolution $\boxplus$, except that the characterization via the Fourier transform in Formula 3.2 has to be replaced by the following one, via the R-transform: μ is $\boxplus$-infinitely divisible if and only if there exists $\gamma \in \mathbb{R}$ and a positive finite measure σ on $\mathbb{R}$ such that R-transform of μ can be written

$$R_\mu(z) = \gamma + \int_{\mathbb{R}} \frac{z+t}{1-tz} \mathrm{d}\sigma(t). \tag{3.3}$$

The law μ is then denoted by $\nu_{\boxplus}^{\gamma,\sigma}$.

Once the characterizations (3.2) and (3.3), called *Lévy-Kinchine Formulas*, etablished, one can define a bijection Λ from the set of $*$-infinitely divisible laws to the set of $\boxplus$-infinitely divisible laws, which associates $\nu_{\boxplus}^{\gamma,\sigma}$ to the law $\nu_*^{\gamma,\sigma}$. This bijection, called *the Bercovici-Pata bijection*, is clearly a morphism for the operations $*$ and $\boxplus$. Moreover, one can prove that it is a homeomorphism for the weak topology and that it commutes with the push-forwards by affine maps. It can then be fully characterized by the fact that, when restricted to the set of laws with moments, it transforms classical cumulants into free ones (the definitions of cumulants are recalled at Section 9.6 of the appendix). The following facts can be deduced of this characterization :

- the image, by the bijection Λ, of the Gaussian law with mean m and variance σ^2 is the semi-circle law with center m and radius 2σ,

- the bijection Λ admits the Dirac masses and the Cauchy measures for fixed points,

[1]The real number γ is sometimes called the *drift* of the semigroup associated to μ, whereas the finite measure σ can be interpreted as follows: $\sigma(\{0\})$ represents the *Brownian component* of this semigroup and the measure $\mathbb{1}_{x\neq 0}\frac{1+x^2}{x^2}\mathrm{d}\sigma(x)$, when it is finite, represents its *compound Poisson process* component (in the case where this measure is not finite, the semigroup can be understood by a subtile limit on such decompsitions).

- for all $\lambda > 0$, Λ maps the Poisson law with parameter λ to the law[2]

$$\begin{cases} \frac{\sqrt{(b-x)(x-a)}}{2\pi x} \mathbb{1}_{x\in[a,b]}\mathrm{d}x & \text{if } \lambda \geq 1, \\ \\ (1-\lambda)\delta_0 + \frac{\sqrt{(b-x)(x-a)}}{2\pi x} \mathbb{1}_{x\in[a,b]}\mathrm{d}x & \text{if } \lambda \leq 1, \end{cases} \qquad \text{with } a = (1-\sqrt{\lambda})^2,\ b = (1+\sqrt{\lambda})^2.$$

The correspondence between classical and free infinitely divisible laws given by the Bercovici-Pata bijection would be quite formal, wouldn't have the following theorem been proved by Bercovici and Pata in [14].

Theorem 3.1 *The Bercovici-Pata bijection preserves the limit theorems for sums of i.i.d. random variables. More specifically, for any sequence of integers (k_n) tending to $+\infty$, for any sequence (ν_n) of laws, for any $*$-infinitely divisible law μ, we have*

$$\underbrace{\nu_n * \cdots\cdots * \nu_n}_{k_n \text{ times}} \underset{n\to\infty}{\longrightarrow} \mu \qquad \Longleftrightarrow \qquad \underbrace{\nu_n \boxplus \cdots\cdots \boxplus \nu_n}_{k_n \text{ times}} \underset{n\to\infty}{\longrightarrow} \Lambda(\mu). \tag{3.4}$$

Notice that this theorem is not a consequence of the fact that Λ is a morphism for $*$ and $\boxplus$ and an homeomorphism, because the laws ν_n are not supposed to be infinitely divisible and we do not take their images by Λ. It means that to an infinitesimal level, $*$ and $\boxplus$ behave in parallel ways. The proof of the theorem, relying on integral transforms and complex analysis, do not give a very concrete aproach of the phenomena. My first PhD Thesis work, that I present below, was motivated by the aim of giving an interpretation of the previous theorem in terms of random matrices, making the Bercovici-Pata bijection ad the equivalence (3.4) more intuitive.

Let us fix a $*$-infinitely divisible law μ, and consider a sequence (k_n) of integers tending to $+\infty$ and a sequence (ν_n) of laws such that $\nu_n^{*k_n} \longrightarrow \mu$, as in the left hand term of (3.4). We are going to construct a random matrix model which will turn the proof of the convergence

$$\underbrace{\nu_n \boxplus \cdots\cdots \boxplus \nu_n}_{k_n \text{ fois}} \underset{n\to\infty}{\longrightarrow} \Lambda(\mu) \tag{3.5}$$

into a random matrices issue. For each fixed n, ν_n is the limit, as $d \to \infty$, of the empirical spectral distribution of the $d \times d$ random matrix

$$M_d(\nu_n) := U \begin{pmatrix} \lambda_1^{(\nu_n)} & & \\ & \ddots & \\ & & \lambda_d^{(\nu_n)} \end{pmatrix} U^*, \tag{3.6}$$

[2]If $\lambda > 1$, this law is equal, up to a dilation, to the Marchenko-Pastur law with $1/\lambda$ as defined at Equation (1.13).

where the $\lambda_i^{(\nu_n)}$ $(i = 1, \ldots, d)$ are i.i.d. with law ν_n and U is a Haar-distributed orthogonal or unitary matrix, independent of the $\lambda_i^{(\nu_n)}$'s. Indeed, it follows from the law of large numbers that almost surely,

$$\frac{1}{d}\sum_{i=1}^{d} \delta_{\lambda_i^{(\nu_n)}} \underset{d\to\infty}{\longrightarrow} \nu_n.$$

It then follows from the definition of $\boxplus$ that for each fixed n, if the random matrices $M_d(\nu_n)^{(1)}, \ldots, M_d(\nu_n)^{(k_n)}$ are independent copies of the matrix $M_d(\nu_n)$ defined in (3.6), then

$$\text{spectral law}\left(M_d(\nu_n)^{(1)} + \cdots\cdots + M_d(\nu_n)^{(k_n)}\right) \underset{d\to\infty}{\longrightarrow} \nu_n^{\boxplus k_n}.$$

We deduce that (3.5) can be re-written :

$$\lim_{n\to\infty}\lim_{d\to\infty}\left(\text{spectral law}\left(M_d(\nu_n)^{(1)} + \cdots\cdots + M_d(\nu_n)^{(k_n)}\right)\right) = \Lambda(\mu).$$

Up to a permutation of the limits, the convergence (3.5) is then a consequence of the following theorem [A2, Th. 3.1, Th. 6.1].

Theorem 3.2 *(a) For each fixed d, the sequence of Hermitian random $d \times d$ matrices*

$$M_d(\nu_n)^{(1)} + \cdots\cdots + M_d(\nu_n)^{(k_n)}$$

converges in law, as $n \longrightarrow \infty$, to a law $\mathbb{P}_d^\mu$ on the space of $d \times d$ Hermitian matrices, with Fourier transform

$$\int e^{i\,\mathrm{Tr}(A\,M)} \mathrm{d}\mathbb{P}_d^\mu(M) = e^{\mathbb{E}[d\,\psi_\mu(\langle u, Au\rangle)]} \qquad \textit{for each Hermitian matrix } A \in \mathbb{K}^{d\times d},$$

where u is a vector with uniform law on the unit sphere of $\mathbb{K}^d$ and Ψ_μ is the Lévy exponent *of μ, defined at Equation* (3.2).

(b) As $d \longrightarrow \infty$, the empirical spectral distribution of a $\mathbb{P}_d^\mu$-distributed random matrix converges in probability to the law $\Lambda(\mu)$.

Thanks to this theorem, the fact that the Bercovici-Pata bijection preserves limit theorems is the expression of the *commutativity* of the limits $d \to \infty$ and $n \to \infty$ in the following diagram :

$$\begin{tikzcd}
M_d(\nu_n)^{(1)} + \cdots\cdots + M_d(\nu_n)^{(k_n)} \arrow[r, "n\to\infty"] \arrow[d, "d \to \infty"] & \mathbb{P}_d^\mu \arrow[d, "d \to \infty"] \\
\text{empirical spectral law: } \mu_n^{\boxplus n} \arrow[r, "n\to\infty"] & \text{empirical spectral law: } \Psi(\mu)
\end{tikzcd}$$

Another interest of the paper [A2] is that we therein study the spectral laws of new matrix ensembles: in the preceding construction, μ can be *any* $*$-infinitely divisible law, as for example a heavy-tailed stable law, a Cauchy law... For example, this paper is, up to my knowledge, the first one giving a rigorous result for the limit spectral law of random matrices whose entries have stable non Gaussian laws (but these entries are not independent, which distinguish this model from the ones studied in [12, 9]). The same matrix models have been studied simultaneously by Cabanal-Duvillard in [28].

To conclude, let us say a few words of the proof of Theorem 3.2. Part (a) is proved using the Fourier transform, and Part (b) is proved using the method of moments when μ has moments to all orders, and then can be extended by approximation, with a control on the *rank* of the error.

3.2 $\boxplus_c$-infinitely divisible laws

In the paper [A4], I studied the infinite divisibility for the convolutions $\boxplus_c$ ($c \in [0,1]$). It appears that as for the convolutions $*$ and $\boxplus$, the laws appearing at the limit in "limit theorems" are also the ones belonging to semigroups for $\boxplus_c$ indexed by $\mathbb{R}_+$ and can be characterized by a Lévy-Kinchine formula [A4, Th. 2.2, 2.5, 2.6]:

Theorem 3.3 *For μ a law on $\mathbb{R}_+$, we have equivalence between the following propositions:*

(i) There exists a sequence (k_n) of integers with infinite limit and a sequence (ν_n) of laws on $\mathbb{R}_+$ such that $\underbrace{\nu_n \boxplus_c \cdots \cdots \boxplus_c \nu_n}_{k_n \text{ times}} \underset{n\to\infty}{\longrightarrow} \mu$.

(ii) There exists a collection $(\mu^{\boxplus_c t})_{t\in[0,+\infty)}$ of laws, starting at δ_0, such that $\mu^{\boxplus_c 1} = \mu$ and which is a semigroup for $\boxplus_c$.

(iii) There exists a finite symmetric positive measure σ on $\mathbb{R}$ such that the rectangular R-transform with ratio c of μ is given by the formula

$$C_\mu^{(c)}(z) = z \int_{t\in\mathbb{R}} \frac{1+t^2}{1-t^2 z} \mathrm{d}\sigma(t).$$

In this case, σ is unique and μ is denoted by $\nu_{\boxplus_c}^{\sigma}$.

Such laws are said to be $\boxplus_c$-*infinitely divisible.*

Part (iii) of the previous theorem allows to define a *Bercovici-Pata bijection with ratio* c from the set of symmetric $*$-infinitely divisible laws[3] to the set of $\boxplus_c$-infinitely divisible laws: this map, denoted by Λ_c, maps any law $\nu_*^{0,\sigma}$ (with symmetric σ mesure on $\mathbb{R}$) to $\nu_{\boxplus_c}^{\sigma}$. As Λ, the bijection Λ_c preserves limit theorems [A4, Th. 3.3]:

Theorem 3.4 *Let (k_n) be a sequence of positive integers with infinite limit and (ν_n) be a sequence of symmetric laws on $\mathbb{R}$. We set $|\nu_n|$ to be the law of $|X|$ for X a ν_n-distributed random variable. Then for any symmetric $*$-infinitely divisible law μ, we have*

$$\underbrace{\nu_n * \cdots\cdots * \nu_n}_{k_n \text{ times}} \longrightarrow \mu \iff \underbrace{|\nu_n| \boxplus_c \cdots\cdots \boxplus_c |\nu_n|}_{k_n \text{ times}} \longrightarrow \Lambda_c(\mu). \tag{3.7}$$

This theorem allows to see easily that for $c = 1$, Λ_c is the restriction of the "square type" Bercovici-Pata bijection Λ to the set of symmetric $*$-infinitely divisible laws.

Example 3.5 *(a) For μ the standard Gaussian law, $\Lambda_c(\mu)$ is the law of $\sqrt{x}$ for x a random variable distributed according to the Marchenko-Pastur law $L_{\mathrm{MP},\,c}$ introduced at* (1.13)*,* i.e. *$\Lambda_c(\mu)$ is the law with density*

$$\frac{\sqrt{(m_+^2 - x^2)(x^2 - m_-^2)}}{\pi c x} \mathbb{1}_{x \in [m_-, m_+]} \mathrm{d}x$$

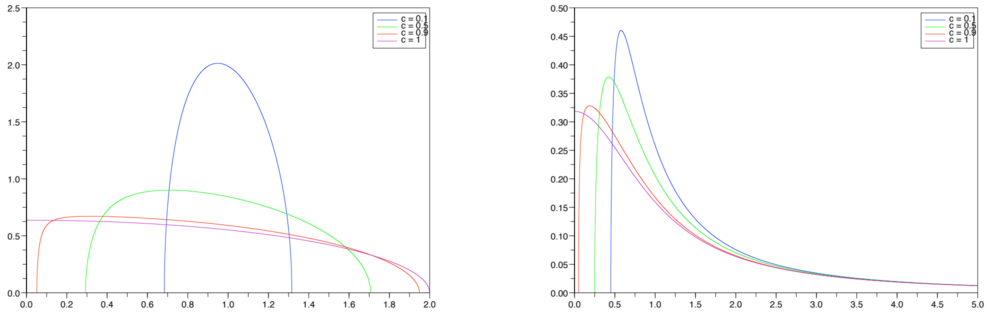
and support $[m_-, m_+]$ for $m_\pm = 1 \pm \sqrt{c}$ (if $c = 0$, this formula has to be understood as the one of the Dirac mass at 1).

(b) For μ the standard Cauchy law,

$$\Lambda_c(\mu) = \frac{\mathbb{1}_{2x > 1-c}}{\pi} \frac{\sqrt{1 - \frac{(1-c)^2}{4x^2}}}{c + x^2} \mathrm{d}x.$$

The densities of these laws are represented at Figure 3.1.

The case of the Gaussian law allows to put the semicircle law and the Marchenko-Pastur law on the same level: the semi-circle law, that is the limit distribution of the eigenvalues of Gaussian Hermitian matrices, plays the role of the Gaussian law for the convolution $\boxplus$, whereas the push-forward of the Marchenko-Pastur law $L_{\mathrm{MP},\,c}$ by the map $x \longmapsto \sqrt{x}$, giving the limit distribution of the singular values of Gaussian $n \times p$ matrices, plays the



(a) Gaussian law (b) Cauchy law

Figure 3.1: **Densities of the $\boxplus_c$-infinitely divisible analogues of the Gaussian and Cauchy laws, for $c = 0.1, 0.5, 0.9$ and 1.** One can notice a phenomenon on which we shall come back later: when $c < 1$, the supports of these laws do not contain zero. The reader who has no color version of this text can however distinguish the curves by observing that the more close c is to 1, the more extended the supports of these laws are.

role of the Gaussian law for the convolution $\boxplus_c$.

In the paper [A4], I constructed a matrix model for the bijection Λ_c of the type of the one associated to the bijection Λ and presented at the previous section: the role of the eigenvalues is payed by the singular values and the square Hermitian $d \times d$ matrix $M_d(\nu_n)$ defined at Equation (3.6) is played by the $d \times d'$ matrix

$$M_{d,d'}(\nu_n) := U \begin{pmatrix} \lambda_1^{(\nu_n)} & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \lambda_d^{(\nu_n)} & 0 & \cdots & 0 \end{pmatrix} V,$$

where the $\lambda_i^{(\nu_n)}$ $(i = 1, \ldots, d)$ are i.i.d. ν_n-distributed variables and U, V are $d \times d$, $d' \times d'$ Haar-distributed unitary matrices, independent and independent of the $\lambda_i^{(\nu_n)}$'s (the dimensions d, d' tend to infinity on such a way that $d/d' \longrightarrow c$).

3.3 Regularization properties of the free convolutions and repulsion of the singular values at zero

The classical convolution $*$ is a standard tool to regularize functions and measures. In the paper [A7], with Serban Belinschi and Alice Guionnet, we have studied the regularization

[3]It can easily be seen, via Formula (3.2), that the symmetric $*$-infinitely divisible laws are precisely the laws $\nu_*^{\gamma,\sigma}$ for which $\gamma = 0$ and σ is symmetric.

properties of the convolutions $\boxplus$ and $\boxplus_c$. Because the densities of the distributions appearing in the universe of free probabilities and random matrices often have infinite derivatives at the border of their support as shown by Biane in [19], the regularization issue for the free convolutions is linked to the positivity of the density. In order to use the semigroup properties, the regularization by infinitely divisible laws is of particular interest, and we brought a particular attention to this case.

We shall see that in the case of the rectangular free convolution $\boxplus_c$, these questions bring to light an interesting phenomenon, that one can already have observed in Figures 1.1, 1.3, 1.4 and 3.1: the density of the convolution is likely to vanish in a neighborhood of zero, which means concretely that the singular values of the sum of two random matrices are likely to avoid taking too small values.

3.3.1 Case of the "square type" free convolution $\boxplus$

The following theorem, proved in part by Bercovici and Voiculescu [17, Th. 7.4] and in part by Belinschi [8, Th. 4.1], shows that the free convolution $\boxplus$ has strong regularization properties.

Theorem 3.6 *Let μ, ν be laws on $\mathbb{R}$, none of them being supported by a single point. Then:*

(a) The law $\mu \boxplus \nu$ can be decomposed into a sum of atoms and of part which is absolutely continuous with respect to the Lebesgue measure.

(b) The atoms of $\mu \boxplus \nu$ are the real numbers a such that there exists $b, c \in \mathbb{R}$ such that $a = b + c$ and $\mu(\{b\}) + \nu(\{c\}) > 1$. In this case, $\mu(\{a\}) = \mu(\{b\}) + \nu(\{c\}) - 1$.

(c) There exists an open subset U of $\mathbb{R}$ and an analytic, positive function f on U such that the absolutely continuous part of $\mu \boxplus \nu$ is the measure $\mathbb{1}_{x \in U} f(x)\mathrm{d}x$.

This theorem, as impressive as it can be, does not solve the question of singular points, *i.e.* the ones where the density vanishes (at the border of U), sometimes with an infinite derivative. In the paper [A7], at Theorems 3.1 et 3.2, we prove that under certain technical hypotheses that we do not detail here, it is possible to improve these results.

Theorem 3.7 *Let μ be a law on $\mathbb{R}$.*

- *Under certain hypotheses about G_μ, for any law ν that is not a Dirac mass, the measure $\nu \boxplus \mu$ has a density with respect to the Lebesgue measure, that is analytic and positive on the whole real line.*

• *If μ is $\boxplus$-infinitely divisible, with semigroup $(\mu^{\boxplus t})_{t\geq 0}$, the under certain hypotheses about R_μ, for any law ν, for all $t > 0$, the measure $\nu \boxplus \mu^{\boxplus t}$ has a density with respect to the Lebesgue measure, that is analytic and positive on the whole real line.*

The Cauchy law, whose free convolution $\boxplus$ with any law coincides with its classical convolution with the same law, if of course an example of regularizing law in the sense of the previous theorem. There are other examples, given in [A7]. Unfortunately, the following proposition [A7, Prop. 3.4] show that it is impossible to regularize with laws allowing to use the method of moments.

Proposition 3.8 *Let μ be an $\boxplus$-infinitely divisible law, with semi-group $(\mu^{\boxplus t})_{t\geq 0}$. If there exists $t < 1$ such that $\mu^{\boxplus t}$ has no atom and μ has a finite second moment, then there exists a law ν such that the density of $\mu \boxplus \nu$ is neither analytical nor positive on the whole real line.*

Let us now give a few ideas about the proofs of both previous theorems.

The first one is to derive properties of probability measures out of their Cauchy transforms: it is well known that for Lebesgue-almost $x \in \mathbb{R}$, the density of the absolutely continuous part of a law μ at x is equal to

$$-\frac{1}{\pi} \lim_{y \to 0^+} G_\mu(x + iy), \tag{3.8}$$

but one can even prove [A7, Lem. 2.11], that for any interval $I \subset \mathbb{R}$, μ admits an analytic density on I if and only if G_μ, initially defined on $\mathbb{C}^+$, can be analytically extended to a neighborhood of $\mathbb{C}^+ \cup I$ (and in this case, the density is of course given by the imaginary part of this extension, up to the factor $-1/\pi$). These remarks lead a general philosophy used in these proofs: *the less G_μ explodes at the neighborhood of a real number x, the more regular μ is at x.* In other words: *the more G_μ admits analytic extensions below the real line, the more regular μ is.* The most flagrant example is the one where $\mu = \nu \boxplus \mathcal{C}$, $\mathcal{C}$ denoting the Cauchy law. In this case,

$$G_\mu(z) = G_\nu(z + i), \tag{3.9}$$

so that G_μ can be analytically extended to $\{z \in \mathbb{C}\,;\, \Im(z) > -1\}$ (and of course μ admits an analytic density on the whole real line).

From there, the main tool, in in the proofs of regularity results, will be the so-called *subordination*, proved by Biane [21, Th. 3.1] : for μ, ν laws on $\mathbb{R}$, there exists two analytic functions ω_1, ω_2 on $\mathbb{C}^+ := \{z \in \mathbb{C}\,;\, \Im(z) > 0\}$ such that for all $z \in \mathbb{C}^+$,

$$\Im(\omega_j(z)) \geq \Im(z) \qquad\qquad (j = 1, 2), \tag{3.10}$$

and

$$G_{\mu\boxplus\nu}(z) = G_\mu(\omega_1(z)) = G_\nu(\omega_2(z)). \tag{3.11}$$

Formula (3.10) means that the functions ω_j are likely to increase the imaginary part of their arguments, which allows to hope, by (3.11), that in certain cases at least, the extension of $G_{\mu\boxplus\nu}$ beyond $\mathbb{C}^+$ will be possible.

3.3.2 Case of the rectangular free convolution $\boxplus_c$: regularity and repulsion of the singular values at the origin

We shall see here that the rectangular free convolution $\boxplus_c$ shares a few regularization properties "square type" free convolution $\boxplus$, and that in this case, a supplementary phenomenon happens : the density of the convolution is likely to vanish in a neighborhood of zero, which means that the singular values of the sum of two random matrices are likely to avoid the vicinity of zero. One can have already observed this phenomenon in Figures 1.1, 1.3, 1.4 and 3.1. The repulsion of singular values at the origin plays a key role in several proofs relying on the relation between singular values and eigenvalues of non Hermitian random matrices, as in the proof, by Tao, Vu and Krishnapur, of the universality of the circle law [111] or in the proof, by Guionnet, Krishnapur and Zeitouni, of the Single Ring Theorem [62].

Let us first give our **regularity** results. They have been established in the particular case where one of the measures is $\boxplus_c$-infinitely divisible [A7, Cor. 4.4 and 4.6, Prop. 4.10].

Theorem 3.9 *Let μ be a $\boxplus_c$-infinitely divisible law such that $\mu \neq \delta_0$.*

- *Then for any law ν on $\mathbb{R}_+$, the singular part[4] of $\mu\boxplus_c\nu$ is supported by a closed set with zero Lebesgue measure and the density of its absolutely continuous part is continuous on the complementary of this closed set in $\mathbb{R}_+$.*

- *Under some supplementary hypotheses about μ, for any law ν on $\mathbb{R}_+$, $\mu\boxplus_c\nu$ is absolutely continuous, with continuous density on $\mathbb{R}_+$.*

- *Under some supplementary hypotheses about μ, for any law ν on $\mathbb{R}_+$, $\mu\boxplus_c\nu$ is absolutely continuous and its density is continuous on $\mathbb{R}_+$ and analytic on an open set U such*

[4]The *singular part* (resp. the *absolutely continuous part*) of a measure μ is μ^s (resp. μ^{ac}), where $\mu = \mu^s + \mu^{ac}$, with μ^s (resp. μ^{ac}) supported by a set of null Lebesgue measure (resp. absolutely continuous with respect to the Lebesgue measure). Note that in the general case, nothing allows to assert that the *support* of μ^s has null Lebesgue measure (the case where μ is the counting measure on the set of rational numbers is a good example).

that $\mu \boxplus_c \nu(U) = 1$.

In [A7, Prop. 4.11], we give a sufficient condition on the finite measure σ for the $\boxplus_c$-infinitely divisible law $\mu = \nu^{\sigma}_{\boxplus_c}$ to satisfy the hypotheses of the third part of the previous theorem. One can for example prove that for all $\alpha \in [1, 2)$, these hypotheses are satisfied by the $\boxplus_c$-stables laws with de parameter α, that are the images, by the Bercovici-Pata bijection Λ_c, of the classical symmetric stable laws with same parameters, and among which is the Cauchy law for $\boxplus_c$, whose density is given at Formula (3.5).

Let us now present our work on the behavior of $\mu \boxplus_c \nu$ in the neighborhood of the origin. For μ, ν laws on $\mathbb{R}_+$, it is easy to see, by linear algebra arguments, that

$$(\mu \boxplus_c \nu)(\{0\}) \geq \mu(\{0\}) + \nu(\{0\}) - 1. \tag{3.12}$$

The first part of the following theorem, taken from Proposition 4.12 of [A7], gives a kind of reciprocal to (3.12). The second part, taken from Proposition 4.13 of [A7], gives a concrete sense to the **repulsion of singular values away from zero** phenomenon.

Theorem 3.10 *Let μ, ν be laws on $\mathbb{R}_+$ such that μ is $\boxplus_c$-infinitely divisible.*

- *We have $(\mu \boxplus_c \nu)(\{0\}) = (\mu(\{0\}) + \nu(\{0\}) - 1)_+$, with $x_+ := \max\{x, 0\}$.*

- *If $\mu(\{0\}) + \nu(\{0\}) < 1$, then for a certain $\varepsilon > 0$, the law $\mu \boxplus_c \nu$ does not charge the interval $[0, \varepsilon]$.*

Let us finish with a simple consequence of this theorem. We hope that it will convince the reader of the utility of the considerations of this chapter. This result is to compare with the ones of Śniady [105] and of Haagerup [64].

Corollary 3.11 *Let $A, X \in \mathbb{K}^{n \times p}$ be independent random matrices, depending implicitly on the integers n, p which shall tend to infinity in such a way that $n/p \to c \in [0, 1]$. We suppose that the entries of X are i.i.d. centered Gaussian variables with variance σ^2/p (for $\sigma > 0$ fixed) and that the empirical singular values distribution of A converges in probability to a law $\neq \delta_0$. Then there exists $\varepsilon > 0$ such that for the convergence in probability,*

$$\sharp\{\lambda \textit{ sing. val. of } A + X \,;\, \lambda \leq \varepsilon\} = o(n).$$

Chapter 4

Eigenvectors of Wigner matrices : universality of the global fluctuations

In this chapter, I present the three main results my work of [A19] : Theorems 4.1 and 4.2 and Proposition 4.5.

4.1 Introduction

It is well known that the matrix $U_n = [u_{i,j}]_{i,j=1}^n$ whose columns are the eigenvectors of a GOE or GUE matrix X_n can be chosen to be distributed according to the Haar measure on the orthogonal or unitary group. As a consequence, much can be said about the $u_{i,j}$'s: their joint moments can be computed via the so-called *Weingarten calculus* developed in [39, 41], any finite (or not too large) set of $u_{i,j}$'s can be approximated, as $n \to \infty$, by independent Gaussian variables (see [71, 37] or Theorem 6.3 of the present text) and the global asymptotic fluctuations of the $|u_{i,j}|$'s are governed by a theorem of Donati-Martin and Rouault, who proved in [45] that as $n \to \infty$, the bivariate càdlàg process

$$\left(B^n_{s,t} := \sqrt{\frac{\beta}{2}} \sum_{\substack{1 \le i \le ns, \\ 1 \le j \le nt}} (|u_{i,j}|^2 - 1/n) \right)_{(s,t)\in[0,1]^2}$$

(where $\beta = 1$ in the real case and $\beta = 2$ in the complex case) converges in distribution, for the Skorokhod topology, to the *bivariate Brownian bridge*, *i.e.* the centered continuous Gaussian process $(B_{s,t})_{(s,t)\in[0,1]^2}$ with covariance

$$\mathbb{E}[B_{s,t}B_{s',t'}] = (\min\{s, s'\} - ss')(\min\{t, t'\} - tt'). \tag{4.1}$$

A natural question is the following:

What can be said beyond the Gaussian case, when the entries of the Wigner matrix X_n are general random variables ?

For a general Wigner matrix, the exact distribution of the matrix U_n cannot be computed and few works had been devoted to this subject until quite recently. One of the reasons is that while the eigenvalues of an Hermitian matrix admit variational characterizations as *extremums* of certain functions, the eigenvectors can be characterized as the *argmax* of these functions, hence are more sensitive to perturbations of the entries of the matrix. However, in the last three years, the eigenvectors of general Wigner matrices have been the object of a growing interest, due in part to some relations with the universality conjecture for the eigenvalues. In several papers (see, among others, [48, 49, 50]), a *delocalization property* was shown for the eigenvectors of random matrices. More recently, Knowles and Yin in [76] and Tao and Vu in [113] proved that if the first four moments of the atom distributions[1] of X_n coincide with the ones of a GOE or GUE matrix, then under some tail assumptions on these distributions, the $u_{i,j}$'s can be approximated by independent Gaussian variables as long as we only consider a finite (or not too large) set of $u_{i,j}$'s.

In this chapter, we consider the global behavior of the $|u_{i,j}|$'s, and we prove (Theorem 4.1) that for Wigner matrices whose entries have moments of all orders, the process $(B^n_{s,t})_{(s,t)\in[0,1]^2}$ has a limit in a weaker sense than for the Skorokhod topology and that this weak limit is the bivariate Brownian bridge if and only if the off-diagonal entries of the matrix have the same fourth moment as the GOE or GUE matrix (quite surprisingly, no hypothesis on the third moment is necessary). Under some additional hypotheses on the atom distributions (more coinciding moments and continuity), we prove the convergence for the Skorokhod topology (Theorem 4.2).

This result was conjectured by Chafaï, who also conjectures the same kind of universality for unitary matrices appearing in other standard decompositions, such as the singular values decomposition or the Housholder decomposition of non hermitian matrices, as long as the matrix considered has i.i.d. entries with first moments agreeing with the ones of Gaussian variables. It would also be interesting to consider the same type of question in the context of *band matrices*, connecting this problem with the so-called *Anderson conjecture* (see *e.g.* the works of Erdös and Knowles [46, 47], of Schenker [101] or of Sodin [106], or, for a short introduction, the blog note by Chafaï [36]).

4.2 Main results

For each n, let us consider a real or complex Wigner matrix

$$X_n := \frac{1}{\sqrt{n}}[x_{i,j}]_{i,j=1}^n$$

[1] The *atom distributions* of a Wigner matrix are the distributions of its entries.

such that the distributions of the $x_{i,j}$'s on the diagonal and off the diagonal do not depend on n (*Wigner matrices* have been defined in Insert 2.3, at the end of Section 2.1.2).

Let us denote by $\lambda_1 \le \cdots \le \lambda_n$ the eigenvalues of X_n and consider an orthogonal or unitary matrix $U_n = [u_{i,j}]_{i,j=1}^n$ such that

$$X_n = U_n \operatorname{diag}(\lambda_1, \dots, \lambda_n) U_n^*.$$

Note that U_n is not uniquely defined. Let us choose it in a measurable way, no matter which one.

We define the bivariate càdlàg process

$$\left(B_{s,t}^n := \sqrt{\frac{\beta}{2}} \sum_{\substack{1 \le i \le ns, \\ 1 \le j \le nt}} (|u_{i,j}|^2 - 1/n) \right)_{(s,t) \in [0,1]^2},$$

where $\beta = 1$ in the real case and $\beta = 2$ in the complex case.

The bivariate Brownian bridge has been defined in the introduction. The exact definition of the functional spaces of the two following theorems and of their topologies can be found in [A19, Sect. 4.1].

Theorem 4.1 *Suppose that the $x_{i,j}$'s have moments of all orders. Then the sequence*

$$(\text{distribution}(B^n))_{n \ge 1}$$

has a unique possible accumulation point supported by $C([0,1]^2)$. This accumulation point is the distribution of a centered Gaussian process which depends on the distributions of the $x_{i,j}$'s only through $\mathbb{E}[|x_{1,2}|^4]$, and which is the bivariate Brownian bridge if and only if $\mathbb{E}[|x_{1,2}|^4] = 4 - \beta$, as in the Gaussian case.

More precisions about the way the unique possible accumulation point depends on the fourth moment of the entries are given in Remark 4.3.

To get a stronger statement where the convergence in distribution to the bivariate Brownian bridge is actually stated, one needs stronger hypotheses.

Theorem 4.2 *Suppose that the $x_{i,j}$'s have a density and moments of all orders, matching with the ones of a GO(U)E matrix up to order 10 on the diagonal and 12 above the diagonal. Then, as $n \to \infty$, the bivariate process B^n converges in distribution, for the Skorokhod topology in $D([0,1]^2)$, to the bivariate Brownian bridge.*

Remark 4.3 Complements on Theorem 4.1. One can wonder how the unique accumulation point mentioned in Theorem 4.1 depends on the fourth moment of the entries of X_n. Let $G := (G_{s,t})_{(s,t)\in[0,1]^2}$ be distributed according to this distribution. We know that $(G_{s,t})_{(s,t)\in[0,1]^2}$ is the bivariate Brownian bridge only in the case where $\mathbb{E}[|x_{1,2}|^4] = 4-\beta$. In the other cases, defining $F_{\text{semicircle}}$ as the cumulative distribution function of the semicircle law, the covariance of the centered Gaussian process

$$\left(\int_{u=-2}^{2} u^k G_{s,F_{\text{semicircle}}(u)}\mathrm{d}u\right)_{s\in[0,1],k\geq 0}, \tag{4.2}$$

which determines completely the distribution of the process G, can be computed. However, making the covariance of G explicit out of the covariance of the process of (4.2) is a very delicate problem, and we shall only stay at a quite vague level, saying that the variances of the one-dimensional marginals of G are increasing functions of $\mathbb{E}[|x_{1,2}|^4]$. For example, that for all $0 \leq s_1, s_2 \leq 1$,

$$\mathrm{Cov}\left(\int_{u=-2}^{2} u^2 G_{s_1,F_{\text{semicircle}}(u)}\mathrm{d}u, \int_{u=-2}^{2} u^2 G_{s_2,F_{\text{semicircle}}(u)}\mathrm{d}u\right) = \frac{\mathbb{E}[|x_{1,2}|^4]-1}{4}(\min\{s_1,s_2\}-s_1s_2).$$

Remark 4.4 Comments on the hypotheses of Theorem 4.2 (1). In order to prove the convergence in the Skorokhod topology, we had to make several hypotheses on the atom distributions: absolute continuity, moments of all orders and coincidence of their 10 (on the diagonal) and 12 (above the diagonal) first moments with the ones of a GOE or GUE matrix. We needed these assumptions to control the discontinuities of the process B^n. Even though these hypotheses might not be optimal (especially the continuity one), a bound on the tails of the atom distributions seems to be necessary to avoid too large variations of the process B^n. Indeed, as illustrated by Figure 4.1, for a GOE matrix (*left* picture), $|u_{i,j}|^2$ is close to $1/n$ for all i,j with high probability, whereas when the atom distributions have not more than a second moment (*right* picture), the matrix X_n looks more like a *sparse matrix*, and so does U_n, which implies that for certain (i,j)'s, $|u_{i,j}|^2-1/n$ is not small enough. Since $|u_{i,j}|^2-1/n$ is the jump of the process B^n at $(s,t)=(i/n,j/n)$, this could be an obstruction to the existence of a continuous limit for the process B^n. That being said, we have hopes to prove the theorem under a four moments hypothesis instead of a 12 moments one (see Remark 4.6 bellow).

Note that it follows from the previous theorem that for all $0 \leq s < s' \leq 1$ and $0 \leq t < t' \leq 1$, the sequence of random variables

$$\frac{1}{\sqrt{(s'-s)(t'-t)}}\sum_{\substack{ns<i\leq ns'\\ nt<j\leq nt'}}(|u_{i,j}|^2-1/n)$$

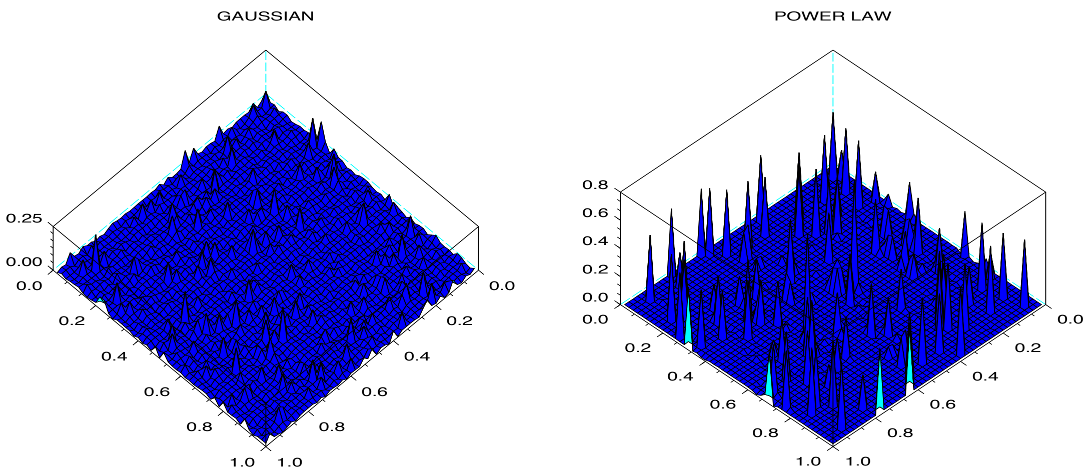


Figure 4.1: **Influence of the tails of the atom distributions of X_n on the $|u_{i,j}|$'s:** Plot of the map $(i/n, j/n) \longmapsto ||u_{i,j}|^2 - 1/n|$ for two different choices of atom distributions. *Left:* GOE matrix. *Right:* Wigner matrix with atom distribution admitting moments only up to order $2+\varepsilon$ for a small ε. For both pictures, the matrices are $n \times n$ with $n = 50$.

admits a limit in distribution as $n \to \infty$, hence is bounded in probability (in the sense of [111, Def. 1.1]). In the same way, it follows from [76] and [113] that the sequence $n|u_{i,j}|^2 - 1$ is bounded in probability. In the next proposition, we improve these assertions by making them uniform on s, s', t, t', i, j and upgrading them to the L^2 and L^4 levels.

Proposition 4.5 *Suppose that the $x_{i,j}$'s have densities and moments of all orders, matching with the ones of a GO(U)E matrix up to order 4. Then as $n \to \infty$, the sequences*

$$n|u_{i,j}|^2 - 1 \qquad \text{and} \qquad \frac{1}{\sqrt{(s'-s)(t'-t)}} \sum_{\substack{ns<i\le ns' \\ nt<j\le nt'}} (|u_{i,j}|^2 - 1/n) \tag{4.3}$$

are bounded for the respective L^4 and L^2 norms, uniformly in $s < s', t < t', i, j$.

Remark 4.6 Comments on the hypotheses of Theorem 4.2 (2). This proposition is almost sufficient to apply the tightness criterion that we use in this paper. Would the second term of (4.3) have been bounded for the $L^{2+\varepsilon}$ norm (instead of L^2), a four moments hypothesis would have been enough to prove that B^n converges in distribution, for the Skorokhod topology in $D([0,1]^2)$, to the bivariate Brownian bridge.

Let us now outline the proofs of Theorems 4.1 and 4.2 and Proposition 4.5.

Firstly, Theorem 4.2 can be deduced from Theorem 4.1 by proving that the sequence $(distribution(B^n))_{n\geq 1}$ is tight and only has $C([0,1]^2)$-supported accumulation points. This can be done via some upper bounds on the fourth moment of the increments of B^n and on its jumps (*i.e.* of its discontinuities). The proofs of these bounds and of Proposition 4.5 rely on a comparison of the eigenvectors of X_n with the ones of a GO(U)E matrix. Indeed, thanks to the Weingarten calculus, one can easily establish such bounds for Haar-distributed matrices. Such a comparison is obtained with the "one-by-one entries replacement method" developed by Tao and Vu in recent papers, such as [112, 113].

Secondly, the proof of Theorem 4.1 relies on the following remark, inspired by some ideas of Jack Silverstein (see [6, Chap. 10] and [104]): even though we do not have any "direct access" to the eigenvectors of X_n, we have access to the process $(B^n_{s,F_{\mu_{X_n}}(u)})_{s\in[0,1],u\in\mathbb{R}}$, for $F_{\mu_{X_n}}(u) := \frac{1}{n}\sharp\{i\,;\,\lambda_i \leq u\}$. Indeed,

$$B^n_{s,F_{\mu_{X_n}}(u)} = \sqrt{\frac{\beta}{2}} \sum_{1\leq i\leq ns} \sum_{\substack{1\leq j\leq n\\ \text{s.t. } \lambda_j\leq u}} (|u_{i,j}|^2 - 1/n),$$

hence for all fixed $s \in [0,1]$, the function $u \in \mathbb{R} \longmapsto B^n_{s,F_{\mu_{X_n}}(u)}$ is the cumulative distribution function of the signed measure

$$\sqrt{\frac{\beta}{2}} \sum_{1\leq i\leq ns} \sum_{j=1}^{n} (|u_{i,j}|^2 - 1/n)\delta_{\lambda_j}, \tag{4.4}$$

which can be studied via its moments

$$\sum_{1\leq i\leq ns} \left(e_i^* X_n^k e_i - \frac{1}{n} \operatorname{Tr} X_n^k \right) \qquad (k \geq 1),$$

the e_i's being the vectors of the canonical basis. From the asymptotic behavior of the *moments* of the signed measure of (4.4), one can then find out the asymptotic behavior of its *cumulative distribution function*.

Once the asymptotic distribution of the process $(B^n_{s,F_{\mu_{X_n}}(u)})_{s\in[0,1],u\in\mathbb{R}}$ identified, one can obtain the asymptotic distribution of the process $(B^n_{s,t})_{s\in[0,1],t\in[0,1]}$ because the function $F_{\mu_{X_n}}$ tends to the (non random) cumulative distribution function $F_{\text{semicircle}}$ of the semicircle law.

Chapter 5

A continuum of notions of independence notions between the classical and the free one

With Thierry Lévy, in our paper [A11], we have used the Brownian motion on the unitary group to propose an interpolation between the classical and the free notions of independence for elements of a non-commutative probability space. We are going to introduce this question from the point of view of the convolution of measures.

5.1 Convolutions

Let μ and ν be two probability measures on $\mathbb{R}$. The classical convolution of μ and ν, denoted by $\mu * \nu$, can be described in the following way. Let A et B be diagonal $n \times n$ matrices (A et B depend implicitly on the parameter n, as the matrices S, V and U_t below) whose empirical spectral laws converge respectively to μ and ν as $n \longrightarrow \infty$. Let S be the matrix of a uniform random permutation of $\{1, \dots, n\}$. Then as $n \longrightarrow \infty$, we have

$$\frac{1}{n} \sum_{\lambda \text{ eig. of } A+SBS^*} \delta_\lambda \longrightarrow \mu * \nu.$$

Let us now consider a Haar-distributed unitary matrix V. Then as $n \longrightarrow \infty$, we have

$$\frac{1}{n} \sum_{\lambda \text{ eig. of } A+VBV^*} \delta_\lambda \longrightarrow \mu \boxplus \nu.$$

There is a natural interpolation between the distribution of S and that of V: it is the family of the distributions of a unitary Brownian motion U_t ($t \in \mathbb{R}_+$) whose initial distribution is the one of S (see Insert 5.6). Then we have the following result [A11, Cor. 2.10].

Theorem 5.1 *With the notations introduced above, as $n \longrightarrow \infty$, the empirical spectral distribution of the matrix $A + U_t B U_t^*$ converges weakly in probability to a law on $\mathbb{R}$ which depends only on μ, ν and t, that we shall denote by $\mu *_t \nu$ and call the t-*free convolution *of μ and ν.*

This theorem can easily be deduced from the convergence of the empirical spectral distribution of U_t to the one of a free unitary Brownian motion taken at time t, from the convergence of the non-commutative distribution of the pair $\{A, SBS^*\}$ and from the asymptotic freeness of unitarily invariant random matrices.

A classical example where one can compute explicitly the free convolution of two measures is that where $\mu = \nu = \frac{1}{2}(\delta_1 + \delta_{-1})$. In this case, $\mu * \nu = \frac{1}{4}\delta_{-2} + \frac{1}{2}\delta_0 + \frac{1}{4}\delta_2$ and $\mu \boxplus \nu = \mathbb{1}_{[-2,2]}(x)\frac{dx}{\pi\sqrt{4-x^2}}$, a dilation of the arcsine law. One can show that for all $t > 0$,

$$\frac{\delta_1 + \delta_{-1}}{2} *_t \frac{\delta_1 + \delta_{-1}}{2} = \mathbb{1}_{[-2,2]}(x)\frac{\rho_{4t}(e^{4i \arccos \frac{x}{2}})}{\pi\sqrt{4-x^2}}\, dx, \tag{5.1}$$

where ρ_t is the function introduced at Insert 5.7 below. Figure 5.1 represents the densities of these laws, computed numerically according to the method presented at Insert 5.7.

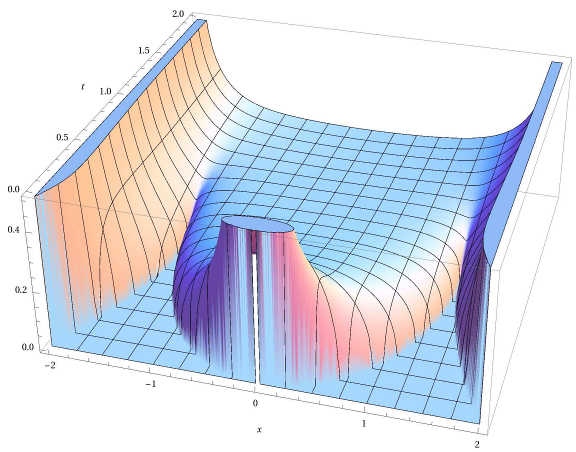


Figure 5.1: Density of $\frac{\delta_1+\delta_{-1}}{2} *_t \frac{\delta_1+\delta_{-1}}{2}$ at $x \in [-2, 2]$ as a function of x and t. One can describe completely the support of the measure for each t. The first time at which this support is the whole interval $[-2, 2]$ is $t = 1$ and the last two points to enter this support are $-\sqrt{2}$ and $\sqrt{2}$.

5.2 Dependance structures and t-freeness

Just as in the case of independence and freeness, the existence of the t-free convolution of two measures is a by-product of the existence of a structure of dependence between sub-algebras of a non-commutative probability space. By a structure of dependence, we mean the following: in a non-commutative probability space $(\mathcal{A}, \varphi)$ where two sub-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$ are given, a way of reconstructing the restriction of φ to the sub-algebra generated by $\mathcal{A}_1 \cup \mathcal{A}_2$ from the restrictions of φ to $\mathcal{A}_1$ and $\mathcal{A}_2$. In commutative terms, we would say that such a structure allows one to infer the joint distribution of a family of random variables from the knowledge of their individual distributions. More details are given at Section 9.2 of the appendix, where the reader can find the definitions of the tensor and free products of non-commutative probability spaces, as well as the ones of freeness and independence in such spaces.

We are going to define a universal model for t-freeness by defining a state on the free product of the algebras underlying two arbitrary non-commutative probability spaces. We shall use here the free product $\mathcal{A}_1 \star \mathcal{A}_2$ of two algebras, defined at Section 9.2 of the Appendix.

Definition 5.2 *Let $(\mathcal{A}_1, \varphi_1)$ and $(\mathcal{A}_2, \varphi_2)$ be two non-commutative probability spaces. Let t be a positive real number. Let $(\mathcal{U}, \tau)$ be a non-commutative probability space generated by a unitary element u_t whose distribution is the measure ν_t defined at Insert 5.7. Let f be the unique algebras morphism*

$$f : \mathcal{A}_1 \star \mathcal{A}_2 \to (\mathcal{A}_1 \otimes \mathcal{A}_2) \star \mathcal{U}$$

such that $f(a_1) = a_1 \otimes 1$ for all $a_1 \in \mathcal{A}_1$ and $f(a_2) = u_t(1 \otimes a_2)u_t^*$ for all $a_2 \in \mathcal{A}_2$. We call t-free product *of φ_1 and φ_2 the state $\varphi_1 *_t \varphi_2$ on $\mathcal{A}_1 \star \mathcal{A}_2$ defined by*

$$\varphi_1 *_t \varphi_2 = [(\varphi_1 \otimes \varphi_2) \star \tau] \circ f.$$

The fact that the measure ν_t is invariant by complex conjugation implies that the pair (u_t, u_t^{-1}) has the same distribution as the pair (u_t^{-1}, u_t), so that replacing f by the morphism f' defined by $f'(a_1) = u_t(a_1 \otimes 1)u_t^*$ and $f'(a_2) = 1 \otimes a_2$ would lead to the same definition of $\varphi_1 *_t \varphi_2$.

For $t = 0$, we recover the definition of the tensor product of two states, transported from the tensor product of the algebras to their free product by the natural morphism $\mathcal{A}_1 \star \mathcal{A}_2 \to \mathcal{A}_1 \otimes \mathcal{A}_2$. On the other hand, if $t > 0$, the element u_t is not scalar in $\mathcal{U}$ and this allows one to prove that m is injective. So, the sub-algebra of $(\mathcal{A}_1 \otimes \mathcal{A}_2) \star \mathcal{U}$ generated by $\mathcal{A}_1 \otimes 1$ and $u_t(1 \otimes \mathcal{A}_2)u_t^*$ is a realisation of the free product of $\mathcal{A}_1$ and $\mathcal{A}_2$.

Once a universal model is defined, we can define t-freeness [A11, Def. 2.5].

Definition 5.3 *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space. Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two sub-algebras of $\mathcal{A}$. Let φ_1 and φ_2 denote the restrictions of φ to $\mathcal{A}_1$ and $\mathcal{A}_2$ respectively. We say that $\mathcal{A}_1$ and $\mathcal{A}_2$ are* t-free *if the natural morphism of algebras $f : \mathcal{A}_1 \star \mathcal{A}_2 \to \mathcal{A}$ satisfies the equality $\varphi \circ f = \varphi_1 *_t \varphi_2$. We say that two subsets of $\mathcal{A}$ are t-free if the involutive sub-algebras which they generate are.*

The observation made after Definition 5.2 ensures that this definition is symmetric in $\mathcal{A}_1$ and $\mathcal{A}_2$.

We now check that t-convolution corresponds to t-freeness.

Proposition 5.4 *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space. Let $t \geq 0$ be a real number and let $a, b \in \mathcal{A}$ be two self-adjoint elements which are t-free, with respective distributions μ and ν. Then the distribution of $a + b$ is $\mu *_t \nu$.*

We can define other t-free convolutions. For example, we can multiply t-free elements: if, with the notation of the proposition above, $\mathcal{A}$ is a C^*-algebra and a, b are non-negative, we can define $\mu \odot_t \nu$ as the distribution of $\sqrt{b}a\sqrt{b}$. Also, if a and b are unitary, in which case μ and ν are measures on the unit circle, we can define $\mu \odot_t \nu$ as the distribution of ab.

Let us give a brief explanation on the concrete meaning of the t-freeness. The *freeness* characterizes the non-commutative distribution of two large Hermitian matrices A, B whose eigenvector bases are in a generic position with each other (*i.e.* one can change one into the other by a generic unitary transformation, which can for example be obtained by choosing it randomly according to the Haar measure). The **independence** characterizes the non-commutative distribution of two codiagonalisable matrices such that one can change the eigenvector basis of A ordered by increasing eigenvalues into the ordered eigenbasis of B via a generic permutation of the vectors of the basis [A11, Th. 1.8]. The *t-freeness* characterizes the non-commutative distribution of two large Hermitian matrices that are codiagonalisable *up to the conjugation of one of then by a unitary Brownian motion taken at time t.*

Besides, the t-freeness is linked to the *liberation process* defined by Voiculescu in [119, Sect. 2.1] : roughly, one can write

t-freeness = liberation process taken at time t + independence at time 0.

More specifically, if $\mathcal{A}_1$ and $\mathcal{A}_2$ are two independent algebras, then their images by the liberation process starting at $(\mathcal{A}_1, \mathcal{A}_2)$ at time t are t-free.

5.3 Differential systems

When one tries to perform computations with pairs of t-free variables, one is led to compute expressions of the form

$$\tau(a_1 u_t b_1 u_t^* \dots a_n u_t b_n u_t^*), \tag{5.2}$$

where the family $\{a_1, \dots, a_n\}$ is independent of $\{b_1, \dots, b_n\}$ and u_t is free with the union of these two families.

The best grip that one has on such expressions is to consider them as functions of t and to establish as small a differential system as possible which they satisfy. In order to differentiate with respect to t an expression like (5.2), one uses a free stochastic differential equation satisfied by the free unitary Brownian motion, which is the analogue (and in a vague sense the limit as n tends to infinity) of the stochastic differential equation (5.5) which defines the unitary Brownian motion.

This derivative involves products of expressions of the form (5.2). By considering all possible products of expressions of this form where $a_1, \ldots, a_n, u_t b_1 u_t^*, \ldots, u_t b_n u_t^*$ appear each exactly once, one obtains a finite set of functions of t which satisfies a closed differential system. By solving such systems, one obtains the following result [A11, Proposition 3.5].

Proposition 5.5 *Let us define a function $G(t,z)$ in a neighbourhood of $(0,0)$ in $\mathbb{R}_+ \times \mathbb{C}$ in the following way.*

Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space. Let a and b be two normal elements whose distributions have compact support and are symmetric (by this we mean that a and $-a$ on one hand, b and $-b$ on the other, have the same distribution). Choose $t \geq 0$. Assume that a and b are t-free. We set

$$G(t,z) = \sum_{n \geq 1} \varphi((a u_t b u_t^*)^{2n}) e^{2nt} z^n.$$

Then G is the unique solution, in a neighbourhood of $(0,0)$ in $\mathbb{R}_+ \times \mathbb{C}$, of the non-linear equation

$$\begin{cases} \partial_t G + 2z \partial_z (G^2) = 0, \\ G(0,z) = \sum_{n \geq 1} \varphi(a^{2n}) \varphi(b^{2n}) z^n. \end{cases}$$

This proposition allows one for example to prove (5.1). We can also use it to compute the distribution of the product of two t-free Bernoulli variables. The density obtained is the one represented at Figure 5.2.

Proposition 5.6 *For all $t > 0$, one has the following equality of measures on the unit circle:*

$$\frac{\delta_{-1} + \delta_1}{2} \odot_t \frac{\delta_{-1} + \delta_1}{2} = \rho_{4t}(\xi^2) d\xi,$$

where ρ_{4t} is the function introduced at Insert 5.7 below, i.e. *is the density of the measure ν_{4t} with respect to the uniform law on the unit circle.*

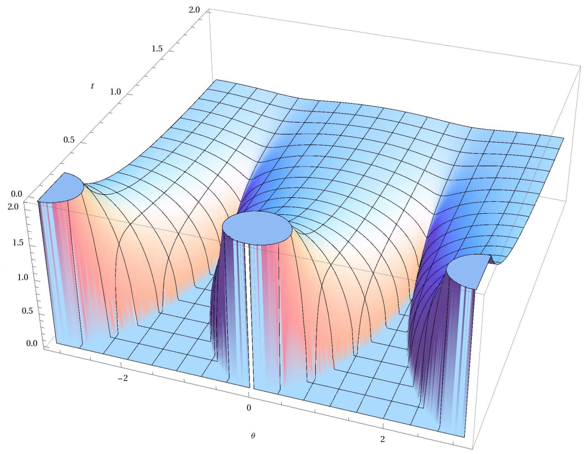


Figure 5.2: Density of $\frac{\delta_{-1}+\delta_1}{2} \odot_t \frac{\delta_{-1}+\delta_1}{2}$ at $e^{i\theta}$ as a function of θ and t. The support of the measure fills the entire circle for the first time at $t = 1$. The last points to enter the support are i and $-i$.

5.4 Non-existence of t-free cumulants

By analogy with the case of freeness, we have wondered if it was possible to find t-free cumulants, that is, multilinear forms defined on every non-commutative probability space and which would vanish as soon as they are evaluated on a set of arguments which can be split into two non-empty families which are t-free (see Section 9.6 of the appendix for an introduction to classical and free cumulants). In the t-free case for $t > 0$, we have shown that there exists nothing as simple and powerful as the free cumulants.

In order to state this result, let us present the set of multilinear forms where we looked for candidates to the role of t-free cumulants.

Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space, $n \geq 1$ and σ an element of the group $\mathfrak{S}_n$ of permutations of $\{1, \dots, n\}$. One defines an n-linear form on $\mathcal{A}$ by setting, for all $a_1, \dots, a_n \in \mathcal{A}$,

$$\varphi_\sigma(a_1, \dots, a_n) = \prod_{(i_1 \cdots i_r) \text{ cycle of } \sigma} \varphi(a_{i_1} \cdots a_{i_r}).$$

This definition makes sense thanks to the traciality of φ (*i.e.* the axiom $\varphi(xy) = \varphi(yx)$ for all x, y).

We now look for t-free cumulants in the set of linear combinations of φ_σ, $\sigma \in \mathfrak{S}_n$: this model contains at the same time the classical cumulants and the free ones. Of course, we are looking for linear combinations with the property that they vanish as soon as they are evaluated on arguments which can be split in two non-empty subsets which form two t-free families.

Definition 5.7 *Let $n \geq 2$ be an integer. Let $t \geq 0$ be a real number. A* t-free cumulant of order n *is a collection $(c(\sigma))_{\sigma \in \mathfrak{S}_n}$ of complex numbers such that*

$$\sum_{\sigma\ n-\text{cycle}} c(\sigma) \neq 0 \tag{5.3}$$

and such that the following property is satisfied in every non-commutative probability space $(\mathcal{A}, \varphi)$: for all pair $(\mathcal{A}_1, \mathcal{A}_2)$ of sub-algebras of $\mathcal{A}$ which are t-free with respect to φ, and for all $a_1, \ldots, a_n$ elements of $\mathcal{A}$ which belong each either to $\mathcal{A}_1$ or to $\mathcal{A}_2$, but neither all to $\mathcal{A}_1$ nor all to $\mathcal{A}_2$, one has

$$\sum_{\sigma \in \mathfrak{S}_n} c(\sigma) \varphi_\sigma(a_1, \ldots, a_n) = 0. \tag{5.4}$$

Let us explain why the constraint (5.3) is natural in this definition. It is clear that if $(c(\sigma))_{\sigma \in \mathfrak{S}_n}$ is a t-free cumulant, then for all $\tau \in \mathfrak{S}_n$, the collection $(c(\tau\sigma\tau^{-1}))_{\sigma \in \mathfrak{S}_n}$ is also a t-free cumulant, as well as

$$\left(\tilde{c}(\sigma) := \sum_{\tau \in \mathfrak{S}_n} c(\tau\sigma\tau^{-1}) \right)_{\sigma \in \mathfrak{S}_n}.$$

The advantage of the last one is that it is *conjugation invariant*, *i.e.* that for all σ, τ, $\tilde{c}(\sigma) = \tilde{c}(\tau\sigma\tau^{-1})$. The sum of (5.3) is then the common coefficient of all maximal length cycles, and the fact that it is non zero for all n insures that it is possible to recover the moments $\varphi(a^k)$ of an element a out of its cumulants.

Our result is the following one [A11, Th. 4.3 et 4.4]. By t-free for $t = +\infty$, we mean *free*.

Theorem 5.8 *1. For all $t \in [0, +\infty]$ and all $n \in \{2, 3, 4, 5, 6\}$, there exists a t-free cumulant of order n. If moreover one insists that this cumulant is invariant by conjugation, that is such that $c(\sigma_1\sigma_2\sigma_1^{-1}) = c(\sigma_2)$ for all $\sigma_1, \sigma_2 \in \mathfrak{S}_n$, then it is unique up to multiplication by a constant.*

2. There exists a t-free cumulant of order 7 if and only if $t = 0$ or $t = +\infty$.

We conclude this chapter with some inserts recalling basic formulas for the Îto matricial calculus, the definition of the unitary Brownian motion, and the one of the free unitary Brownian motion.

Insert 5.5 – **Matricial Îto calculus**

Let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration $K = (K_t)_{t\geq 0}$ be a standard Brownian motion on the space of skew-Hermitian $n\times n$ matrices endowed with the scalar product $A\cdot B = n\,\mathrm{Tr}(A^*B)$, i.e. a process with values in this space such that the diagonal entries of $(i\sqrt{n}K_t)_{t\geq 0}$ and the real and imaginary parts of the non diagonal entries of $(\sqrt{2n}K_t)_{t\geq 0}$are independent standard Brownian motions. Then all the matrix-valued semi-martingales X, Y of the type

$$\mathrm{d}X_t = A_t(\mathrm{d}K_t)B_t + C_t\mathrm{d}t\,, \qquad \mathrm{d}Y_t = D_t(\mathrm{d}K_t)E_t + F_t\mathrm{d}t,$$

where A, B, C, D, E, F are matricial $(\mathcal{F}_t)_{t\geq 0}$-adapted processes, we have

$$\begin{aligned}\mathrm{d}(XY)_t &= (\mathrm{d}X_t)Y_t + X_t\mathrm{d}Y_t - \frac{1}{n}\,\mathrm{Tr}(B_tD_t)A_tE_t\mathrm{d}t,\\ \mathrm{d}\langle \mathrm{Tr}(X), \mathrm{Tr}(Y)\rangle_t &= -\frac{1}{n}\,\mathrm{Tr}(B_tA_tE_tD_t)\mathrm{d}t,\end{aligned}$$

$\langle\cdot,\cdot\rangle$ denoting the quadratic variation.

INSERT 5.6 – **The three definitions of the Brownian motion on the unitary group**

For K a skew-Hermitian $n\times n$ Brownian motion as defined at Insert 5.5, we call a *unitary Brownian motion* any process $U = (U_t)_{t\geq 0}$ taking values in the space of complex $n\times n$ matrices which is a strong solution of

$$\mathrm{d}U_t = \mathrm{d}K_tU_t - \frac{1}{2}U_t\mathrm{d}t \tag{5.5}$$

such that U_0 is almost surely unitary. Thanks to the matricial Îto calculus, on can then prove that almost surely, U_t is unitary for all t, and that the left and right multiplicative increments of U_t are stationary and independent: for all t_0, the processes $(U_{t_0+t}U^*_{t_0})_{t\geq 0}$ and $(U^*_{t_0}U_{t_0+t})_{t\geq 0}$ are unitary Brownian motions, independent of $(U_s)_{0\leq s\leq t_0}$.

The process U can also be defined by rolling the unitary group $\mathcal{U}_n$ on its Lie algebra $\mathfrak{u}$ along the process K (starting at U_0, that is independent of K). Let us be more precise. For $\gamma : [0,+\infty) \to \mathfrak{u}$ a continuous, smooth by parts path starting at 0 and $u_0 \in \mathcal{U}_n$, "the wrapping" of γ on $\mathcal{U}_n$ starting at u_0 is the path w_γ defined by $w_\gamma(0) = u_0$ and $w'_\gamma(t) = w_\gamma(t)\times\gamma'(t)$. Here, U can be roughly obtained in that way with $\gamma = K$ and $u_0 = U_0$, but K is not smooth by parts. However, one can prove that if $(K^{(p)})_{p\geq 1}$is a sequence of continuous, affine by parts interpolations of K, the the sequence $(w_{K^{(p)}})_{p\geq 1}$ converges in probability to a limit process that does not depend on the choice of the interpolations and that is the process U solution of (5.5). Equivalenlty,

$$U_t = \lim_{p\to\infty} U_0 \exp K_{\frac{t}{p}} \exp(K_{\frac{2t}{p}} - K_{\frac{t}{p}})\cdots\cdots\exp(K_{\frac{pt}{p}} - K_{\frac{(p-1)t}{p}}).$$

For more details on this construction, see [70, Sect. VI.7], [97, Eq. (35.6)] ot [55].

At last, one can define U as a continuous Markov process with generator $\frac{1}{2}\Delta$, where Δ is the Laplacien on $\mathcal{U}_n$ for the Riemannian structure associated to the scalar product on its Lie algebra defined by $A\cdot B = n\,\mathrm{Tr}(A^*B)$. We shall not detail this approach, thus we do not detail it.

Insert 5.7 – **Free unitary Brownian motion**

Let U be an $n \times n$ unitary Brownian motion, as defined at Insert 5.6, starting at I. Biane, in [18], and Rains, in [95], have proved that for all $t \geq 0$, the empirical spectral distribution of U_t converges in probability, as $n \longrightarrow \infty$, to the unique law ν_t on the unit circle, invariant by $z \longmapsto \overline{z}$ and with moments

$$e^{-\frac{kt}{2}} \sum_{j=0}^{k-1} \frac{(-t)^j}{j!} \binom{k}{j+1} k^{j-1} \qquad (k \geq 0).$$

The invariance of the law of U by conjugaison by any unitary matrix and the independence of its increments allow easily to prove that the joint non-commutative distribution of $(U_t)_{t\geq 0}$ converges to the one of a process $(u_t)_{t\geq 0}$ of unitary elements[1] of a non-commutative probability space $(\mathcal{A}, \varphi)$ such that

- for all $0 \leq s \leq t$, the distribution of $u_t u_s^*$ is ν_{t-s},
- for all $m \geq 1$, for all $0 = t_0 \leq t_1 \leq \cdots \leq t_m$, the elements $u_{t_1} u_{t_0}^*, \ldots\ldots, u_{t_m} u_{t_{m-1}}^*$ are free.

Such a family $(u_t)_{t\geq 0}$ is called a *free unitary Brownian motion*.

There exists a simple expression of the support of ν_t as a function of t (see [A11, Sect. 3.3.1]) but not of its density ρ_t with respect to the uniform law on the unit circle. However, the analytic function on the unit circle κ_t whose real part is the harmonic extension of ρ_t, given by the formula $\kappa_t(z) = \int \frac{x+z}{x-z} \mathrm{d}\nu_t(x)$, satisfies the equation

$$\frac{\kappa_t(z) - 1}{\kappa_t(z) + 1} e^{\frac{t}{2}\kappa_t(z)} = z \tag{5.6}$$

(this can be obtained by the Lagrange inversion, using the formula of the moments of ν_t). Solving numerically Equation (5.6), one obtains Figure 5.3 below.

[1] A *unitary element* in a non-commutative probability space is an element u such that $uu^* = u^*u = 1$.

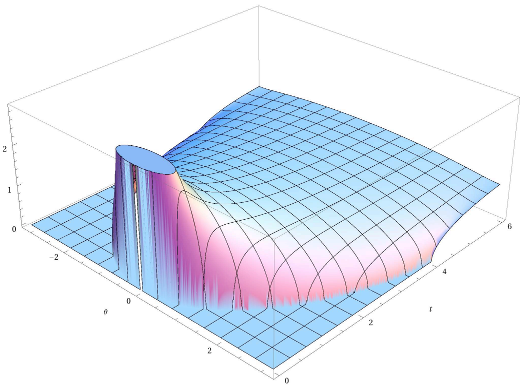


Figure 5.3: Density ρ_t of the law ν_t at $e^{i\theta}$ as a function of θ and t. We see its support filling progressively the circle up to time $t = 4$, which is the first time at which this support covers the whole circle, and then becomes analytic for $t > 4$.

Chapter 6

Central limit theorem for the Brownian motion on the unitary group

The Brownian motion on the unitary group has been introduced in the present text in Insert 5.6. Mainly due to its relations with the free unitary Brownian motion (see Insert 5.7 above) and with the two-dimentional Yang-Mills theory, the Brownian motion on large unitary groups has appeared in several papers during the last decade. Rains, in [95], Xu, in [121], Biane, in [18, 20] and Lévy and Maïda, in [77, 78], are all concerned with the asymptotics of the spectral distribution of large unitary Brownian motions. In this section, we are concerned with the asymptotic distributions of linear combinations of the entries of an $n \times n$ unitary Brownian motion as n tends to infinity.

It is clear that the same analysis would give similar results for the Brownian motion on the orthogonal group. For notational brevity, we chose to focus on the unitary group.

Our main result is the following one [A13, Th. 1.2]. Let $(e^{-t/2}V_t)_{t\geq 0}$ a Brownian motion the group of $n \times n$ matrices such that $V_0 = I$ (V_t depends implicitly on the parameter n, as the matrix A below), let (α_n) be a sequence of positive numbers with a limit $\alpha \in [0, +\infty]$ and let A a be deterministic $n \times n$ complex matrix such that for some fixed numbers a, p, q, as $n \longrightarrow \infty$,

$$\frac{1}{n}\operatorname{Tr}(A) \longrightarrow a, \quad \frac{1}{n}\operatorname{Tr}(A^2) \longrightarrow p, \quad \frac{1}{n}\operatorname{Tr}(AA^*) \longrightarrow q.$$

Let μ be the distribution of a complex-valued continuous Gaussian centered process $(X_t)_{t\geq 0}$

with independent increments such that

$$\mathbb{E}(X_t\overline{X_t}) = qt\,, \qquad \mathbb{E}(X_t^2) = \begin{cases} -pt & \text{if } \alpha = 0, \\ -p\frac{\log(\alpha t+1)}{\alpha} + a\frac{\log^2(\alpha t+1)}{2\alpha} & \text{if } 0 < \alpha < +\infty, \\ 0 & \text{if } \alpha = +\infty. \end{cases} \tag{6.1}$$

Theorem 6.1 *As $n \longrightarrow \infty$, the distribution of the process*

$$(\alpha_n^{-1/2} \operatorname{Tr}[A(V_{\log(\alpha_n t+1)} - I)])_{t\geq 0}$$

converges weakly to μ in the space $\mathcal{C}([0,\infty),\mathbb{C})$ endowed with the topology of uniform convergence on every compact subset.

Note that by the standard properties of the Gaussian spaces, this theorem implies its multidimentional version : for any fixed $d \geq 1$, if $A_1, \ldots, A_d$ are $n \times n$ matrices (depending implicitly on n) such that as $n \longrightarrow \infty$, for all i, j,

$$\frac{1}{n}\operatorname{Tr}(A_i) \longrightarrow a_i, \quad \frac{1}{n}\operatorname{Tr}(A_iA_j) \longrightarrow p_{i,j}, \quad \frac{1}{n}\operatorname{Tr}(A_iA_j^*) \longrightarrow q_{i,j},$$

then the d-dimensional complex process

$$\{\alpha_n^{-1/2}\operatorname{Tr}[A_1(V_{\log(\alpha_n t+1)} - I)], \ldots\ldots, \alpha_n^{-1/2}\operatorname{Tr}[A_d(V_{\log(\alpha_n t+1)} - I)]\}_{t\geq 0}$$

converges weakly to a distribution which can easily be expressed in terms of the a_i's, the $p_{i,j}$'s, the $q_{i,j}$'s and of α.

Recall that a *principal submatrix* of a matrix is a submatrix obtained by removing some columns, and the rows with the same indices.

Corollary 6.2 *Let us fix $p \geq 1$ and let (H_t), (S_t) be two independent standard Brownian motions on the Euclidian spaces of $p \times p$ respectively Hermitian and skew-Hermitian matrices endowed with the respective scalar products $\langle X, Y\rangle = \operatorname{Tr}(XY)/2$, $\langle X, Y\rangle = -\operatorname{Tr}(XY)/2$. Then, as n tends to infinity, the distribution of the $\mathbb{C}^{p\times p}$-valued process of the entries of any $p \times p$ principal submatrix of $\sqrt{n/\alpha_n}(V_{\log(\alpha_n t+1)} - I)_{t\geq 0}$ converges to the one of the random process $(H_{t-f_\alpha(t)} + S_{t+f_\alpha(t)})_{t\geq 0}$, where*

$$f_\alpha(t) = \begin{cases} t & \textit{if } \alpha = 0, \\ \frac{\log(\alpha t+1)}{\alpha} & \textit{if } 0 < \alpha < +\infty, \\ 0 & \textit{if } \alpha = +\infty. \end{cases}$$

Let us comment these results. The most constructive way to define the unitary Brownian motion is to consider a standard Brownian motion $(K_t)_{t\geq 0}$ on the Lie algebra of the unitary group and to take its image by the *Itô map* (whose inverse is sometimes called the *Cartan map*), i.e. to wrap it around the unitary group: the process $(U_t)_{t\geq 0}$ obtained is a unitary Brownian motion starting at I. This construction is the second of the ones presented in Insert 5.6. Our results give us an idea of the way the Itô map alterates the process $(K_t)_{t\geq 0}$ at different scales of time : for small times (i.e. when $\alpha = 0$), the limit process is still purely skew-Hermitian, whereas for large times ($\alpha = +\infty$), the limit process is a standard complex matricial Brownian motion (for intermediate scales of time, $0 < \alpha < +\infty$, the limit process is an interpolation between these extreme cases).

Moreover, the question of the choice of a rescaling of the time (depending on the dimension) raises interesting questions. There are other ways to scale the time for the Brownian motion on the unitary group. Our scaling of the time is the one for which the three limit regimes correspond respectively to small values of t, finite values of t and large values of t and for which the limit non-commutative distribution of $(e^{-t/2}V_t)_{t\geq 0}$ is the one of a free unitary Brownian motion. It also has a heuristic geometrical meaning: with this scaling, for any fixed t, the distance between V_0 and V_t has the same order as the diameter of the unitary group (see [A13, Rem. 1.1]). It means that for any fixed $t > 0$, large values of n, V_t is probably no longer too close to its departure point, while it also probably hasn't "orbited" the unitary group too many times.

To prove Theorem 6.1, we use Rebolledo's Theorem [2, Th. H.14], which states that to prove the convergence in distribution of a centered martingale to a Gaussian process with independent increaments, it suffices to prove the L^1-convergence of the bracket of the martingale to the (deterministic) bracket of the limit process. Here, the convergence of the bracket is obtained via several iterations of the so-called *matricial Îto calculus*, presented in Insert 5.5.

A by-product of Theorem 6.1 is a proof of the following theorem, about the asymptotic normality of unit vectors and unitary matrices. This result is not new (see below), but Theorem 6.1 allows to give a proof which, including the proof of Theorem 6.1, is the shortest we found in the literature. Its proof relies on the fact that the Haar measure is an invariant measure for the heat kernel on the unitary group.

Theorem 6.3 *Let U be an $n \times n$ unitary random matrix with Haar distribution and A be an $n \times n$ non-random matrix (both U and A depending implicitely on n) such that as $n \longrightarrow \infty$,*

$$\frac{1}{n}\operatorname{Tr}(AA^*) \longrightarrow q \geq 0.$$

Then as $n \longrightarrow \infty$, the distribution of $\operatorname{Tr}(AU)$ *tends to a rotation-invariant complex Gaus-*

sian distribution with variance q.

Note that just like Theorem 6.1, by standard properties of Gaussian spaces, this theorem implies its mutlidimensional version: it follows directly from Theorem 6.3 that for any $d \geq 1$, if $A_1, \ldots, A_d$ are $n \times n$ matrices (depending implicitly on n) such that as $n \longrightarrow \infty$, for all i, j,

$$\frac{1}{n} \operatorname{Tr}(A_i A_j^*) \longrightarrow q_{i,j} \in \mathbb{C},$$

then the distribution of the d-dimensional complex process

$$\{\operatorname{Tr}(A_1 U), \ldots\ldots, \operatorname{Tr}(A_d U)\}$$

converges to the one of a complex Gaussian centered vector $\{Z_1, \ldots, Z_d\}$ such that for all i, j,

$$\mathbb{E}[Z_i Z_j] = 0 \qquad \text{and} \qquad \mathbb{E}[Z_i \overline{Z_j}] = q_{i,j}.$$

The historical first result of asymptotic normality of unit vectors was due to Émile Borel, who proved a century ago, in [26], that, for a uniformly distributed point $(X_1, \ldots, X_n)$ on the unit Euclidian sphere $\mathbb{S}^{n-1}$, the scaled first coordinate $\sqrt{n} X_1$ converges weakly to the standard Gaussian distribution as the dimension n tends to infinity. As explained in the introduction of the paper [3] of Diaconis *et al.*, this says that the features of the "microcanonical" ensemble in a certain model for statistical mecanics (uniform measure on the sphere) are captured by the "canonical" ensemble (Gaussian measure). Since then, a long list of further-reaching results about the entries of uniformly distributed random orthogonal or unitary matrices have been obtained. The most recent ones are the previously cited paper of Diaconis *et al.*, the papers of Meckes and Chatterjee [82, 37], the paper of Collins and Stolz [43] and the paper of Jiang [71], where the point of view is slightly different.

Chapter 7

Small cycles of free words in random permutations

7.1 Introduction

This section is devoted to my works [N1, A12]. We consider random permutations which can be written as free words in several independent random permutations: firstly, we fix a non trivial word w in letters $g_1, g_1^{-1}, \ldots, g_k, g_k^{-1}$, secondly, for all n, we introduce a k-tuple $s_1(n), \ldots, s_k(n)$ of independent random permutations of $\{1, \ldots, n\}$, and the random permutation σ_n we are going to consider is the one obtained by replacing each letter g_i in w by $s_i(n)$. For example, for $w = g_1 g_2 g_3 g_2^{-1}$, $\sigma_n = s_1(n) \circ s_2(n) \circ s_3(n) \circ s_2(n)^{-1}$. Moreover, we allow to restrict the set of possible lengths of the cycles of the $s_i(n)$'s: we fix sets $A_1, \ldots, A_k$ of positive integers and suppose that for all i, $s_i(n)$ is uniformly distributed on the set[1] of permutations of $\{1, \ldots, n\}$ which have all their cycle lengths in A_i. For example, if $A_1 = \{1, 2\}$, $s_1(n)$ is a uniform random involution. We are interested in *small cycles* of σ_n, *i.e.* cycles whose length is fixed independently of n. Since the law of σ_n is invariant under conjugation, the positions of its cycles are uniform, and only their lengths contain some unknown randomness. So we introduce, for each positive integer ℓ, the number $N_\ell(\sigma_n)$ of cycles of length l of σ_n. We are interested in the asymptotic behavior of the $N_\ell(\sigma_n)$'s as $n \longrightarrow \infty$.

Before stating the results, let us put this work into perspective.

It is well known [4] that if σ_n is a uniform random permutation of $\{1, \ldots, n\}$, then the $N_\ell(\sigma_n)$'s ($\ell \geq 1$) are asymptotically independent and distributed according to Poisson laws with parameters $1/\ell$. Besides, the law of the "large cycles" (*i.e.* cycles whose size is proportional to n) can be expressed via the Poisson-Dirichlet law.

[1] I can be proved that for n large enough, this set is non empty as soon as n is a multiple of the largest common divisor of A_i [87, Lem. 2.3].

Nica has been the first to consider the case where σ_n is a word in independent random permutations and their inverses, distributed according to the uniform law on the symmetric group. He first proved, in [88], that the number of fixed points of such permutations is $o(n)$, which insures that independent random permutation matrices are asymptotically free. In [89], he goes further in his analysis, finding out the limit law of the number of fixed points of σ_n, and more generally of the $N_\ell(\sigma_n)$'s: he proves that if the word considered is not a power of another word, then for all $\ell \geq 1$, $N_\ell(\sigma_n)$ converges in law to the Poisson law with parameter $1/\ell$. A consequence of this result is that *second order freeness* cannot model free words in random permutations. His result has moreover been used recently by Linial and Puder in [79] where it appears that free words in random permutations play a role in the analysis of *n-lifts* of graphs.

Besides, permutations with restricted cycles lengths have been a subject of interest in the community of combinatorics for a long time (see [122]). In [123], for example, Yakymiv considered the joint limit distribution of the $N_\ell(\sigma_n)$'s in the case where σ_n is uniformly distributed in the set of permutations of $\{1, \ldots, n\}$ whose cycle lengths belong to a fixed infinite set A.

At last, in [87], Neagu somehow joins both previous approaches by proving the asymptotic freeness of the matrices of random permutations with restricted cycle lengths: it leads him to consider words in such random matrices and to prove that they only have $o(n)$ fixed points. This is an extension of Neagu's work which led us to the present work, whose results imply for example that such matrices, though asymptotically free, are not asymptotically second order free. Beyond these issues, my interest for the question comes from another, seemingly quite hard problem, that Thierry Lévy and myself have tried to solve in vain: the characterization of the words w in the letters $g_1, g_1^{-1}, \ldots, g_k, g_k^{-1}$ such that for any compact (or finite) group G, for any family $s_1, \ldots, s_k$ of independent Haar distributed random variables on G, the law of the random variable obtained by replacing each letter g_i in w by s_i is the Haar measure on G.

7.2 Case of a trivial word

We first study the case where $w = g_i$ is a word with only one letter [N1, Prop. 2.1 et Th. 2.3].

Theorem 7.1 *Fix $i \in \{1, \ldots, k\}$. Under certain technical hypotheses on A_i, as $n \longrightarrow \infty$, we have the following convergences:*

- *if A_i is infinite, then for any finite subset $\{\ell_1, \ldots, \ell_p\}$ of A_i, the random vector*
$$\{N_{\ell_1}(s_i(n)), \ldots\ldots, N_{\ell_p}(s_i(n))\}$$
converges in law to
$$\text{Poisson}(1/\ell_1) \otimes \ldots\ldots \otimes \text{Poisson}(1/\ell_p),$$

- *if A_i is finite, for $d_i := \max A_i$, for all $\ell \in A_i$, we have*

$$\frac{N_\ell(s_i(n))}{n^{\ell/d_i}} \underset{n\to\infty}{\longrightarrow} \frac{1}{\ell}$$

in all L^p spaces.

The first part of the previous theorem, which has been proved as me in the same time by Yakymiv under quite close hypotheses [123, Th. 1], signifies that if A_i is infinite, everything happens, for cycles with admitted lengths, as if there were no cycle restriction. The second part means that if A_i is finite, then $s_i(n)$ is not faraway from having order $d_i = \max A_i$ in the symmetric group: le part of $\{1,\ldots,n\}$ recovered by the cycles with length d_i has cardinality $n - O(n^\kappa)$, where $\kappa = \frac{\max(A_i\setminus\{d_i\})}{d_i} < 1$. In the case where A_i is finite, it would be interesting to understand the fluctuations of $\frac{N_\ell(s_i(n))}{n^{l/d_i}}$ around its limit. It seems that Analytic Combinatorics, as presented in the book by Flajolet and Sedgewick [54], could help studying this question.

Let us outline the proof of Theorem 7.1. We denote by $\mathfrak{S}_n^{(A_i)}$ the set of permutations of $\{1,\ldots,n\}$ whose cycle have their lengths in A_i. It can easily be seen that, for $|z| < 1$

$$\sum_{n\geq 1} \frac{\sharp\mathfrak{S}_n^{(A_i)}}{n!} z^n = \exp\left(\sum_{\ell\in A} \frac{z^\ell}{\ell}\right). \tag{7.1}$$

Moreover, Hayman's method (see [54]) allows to link the asymptotic behaviors of the coefficients of both sums in the previous equation. For example, in the case where A_i is finite, one deduce from (7.1) that denoting by p largest common divisor of A_i, as $n \longrightarrow \infty$, we have

$$\frac{\sharp\mathfrak{S}_{np-p}^{(A_i)}}{\sharp\mathfrak{S}_{np}^{(A_i)}} \sim (np)^{\frac{p}{d_i}-p}.$$

Then, the inclusion/exclusion method allows to conclude.

7.3 Non trivial words

One considers now a random permutation σ_n constructed out of a word w in the way presented at the first paragraph of Section 7.1. For each i, we suppose that A_i satisfies the following technical hypothesis:

$$A_i \text{ is finite} \qquad \text{ou} \qquad \sum_{\substack{j\geq 1\\ j\notin A_i}} \frac{1}{j} < \infty. \tag{7.2}$$

The structure of w is of course going to play a role in the behavior of the variables $N_\ell(\sigma_n)$. Thanks to the previous theorem, it seems that for each i such that A_i is finite, for $d_i = \max A_i$, $s_i(n)^{d_i} \approx \mathrm{Id}_{\{1,\dots,n\}}$, which gives the intuition that the sequences of the type $g_i^{d_i}$ in w are to neutralize. This intuition is confirmed by the following theorem (see [A12, Th. 3.4 et 3.6] for slightly more precise statements). We denote by $\mathbb{F}_k$ the free group with generators $g_1, \dots, g_k$ and by $\mathbb{F}_k/H$ its quotient by the relations $g_i^{d_i} = 1$, for the i's in $\{1, \dots, k\}$ such that A_i is finite.

Theorem 7.2 • *If the element of $\mathbb{F}_k/H$ represented by the word w has finite order L, then for large n, σ_n is not faraway from having order L, since as $n \longrightarrow \infty$, we have*

$$N_L(\sigma_n) \sim \frac{n}{L}$$

and for all $\ell \neq d$,

$$\frac{N_\ell(\sigma_n)}{n} \longrightarrow 0.$$

•*If the element of $\mathbb{F}_k/H$ represented by the word w has infinite order, then two cases can happen:*

(a) in $\mathbb{F}_k/H$, the word w does not represent that same element, up to any conjugation, as a word of the type g_i^α, with $i \in \{1, \dots, k\}$ and α an integer : then for all ℓ,

$$\liminf_{n\to\infty} \mathbb{E}[N_\ell(\sigma_n)] \geq \frac{1}{\ell}. \tag{7.3}$$

(b) up to a conjugation, w represents the same element of $\mathbb{F}_k/H$ as g_i^α, for a certain $i \in \{1, \dots, k\}$ and a certain non zero α : then (7.3) is true only for the ℓ's such that $\ell|\alpha| \in A_i$.

In the case where the order, in $\mathbb{F}_k/H$, of the element represented by w is infinite, (7.3) insures the existence of at least as much cycles with size ℓ as if σ_n had been uniformly distributed on the symmetric group (indeed, in this case, one would have had $\mathbb{E}[N_\ell(\sigma_n)] = \frac{1}{\ell}$). The following theorem [A12, Th. 3.7 et 3.8], which establishes a convergence in distribution, gives a more precise estimation.

Theorem 7.3 *For all $\ell \geq 1$, one has the convergence in distribution, as $n \longrightarrow \infty$, of the vector*

$$\{N_1(\sigma_n), \dots\dots, N_\ell(\sigma_n)\} \tag{7.4}$$

to the law

$$\mathrm{Poisson}(1/1) \otimes \dots\dots \otimes \mathrm{Poisson}(1/\ell),$$

in each of the following cases:

- *all A_i's are infinite and w is a non trivial word which is not a power of another word,*
- $w = g_1 \cdots g_k$ *with* $k > 2$.

Note that if w is a power of another word v, the decomposition of σ_n in a product of cycles can be deduced from the one of the random permutation associated to v. Besides, in the case where $w = g_1 \cdots g_k$ with $k = 2$, one can also give the limit law of the vector of (7.4), that is also a tensor product of Poisson-type laws (see [A12, Th. 3.8 and 3.12]).

The proof of both previous theorems relies on the same kind of considerations as the proof of Theorem 7.1 above, and to the study (this is the main difficulty) of certain graphs associated to the word w.

Chapter 8

Perspectives

I indicate here some of the directions I would to follow now. They appear by order of advancement of the project, which is not necessarily the one of the importance I would like to give them in the future.

Heavy tailed random matrices

By "heavy tailed random matrices", we mean random matrices whose coefficients are heavy tailed random variables. These matrices have been studied by Ben Arous and Guionnet in the article [12], and then by several authors in [9, 24, 25, 5]. A variant, allowing more computations, is proposed by Zakharevich in [124].

Such matrices, duly renormalized, are close to "sparce" matrices, *i.e.* matrices containing mainly zeros. In a work in progress with Alice Guionnet and Camille Mâle, we are trying to understand the fluctuations of the empirical spectral distribution of such matrices. The first result is that these fluctuations are in $1/\sqrt{n}$, n denoting the dimension, which corroborates the idea, already appeared in [5], that the eigenvalues of such matrices are closer to independent random variables than the ones of Wigner matrices, which are in a repulsion interaction. In the perspective of the study of the eigenvectors of such matrices with the approach of my article [A19] (see Chapter 4 in this text), it would also be interesting to study their spectral distribution with respect to a non random vector.

Besides, in a work in progress with Thierry Cabanal-Duvillard, we are trying to generalize the Marchenko-Pastur Theorem to the following context: we consider a matrix X defined by

$$X = \frac{1}{p}\sum_{j=1}^{p} C_j C_j^*,$$

where the C_j's are independent random columns, as in the original Marchenko-Pastur Theorem, but we do not suppose anymore the coefficients of the column matrices to have the same order: certain entries are widely larger than the other ones, which alters strongly the distribution of the eigenvalues of X.

Perturbations of random matrices

In a work in progress with with Nathanaël Enriquez, we are studying the following simple question:

> *What happens to the eigenvectors and eigenvalues of an Hermitian matrix X when it is submitted to a small random perturbation ?*

More specifically, we study the eigenvalues and eigenvectors of an $n \times n$ matrix of the type $X_\varepsilon = X + P_\varepsilon$, in the limit where $P_\varepsilon \longrightarrow 0$ as $n \longrightarrow \infty$. Depending on the amplitude of the pertutbation P_ε, several regimes appear at the limit.

Random matrices with correlated entries

Such matrices share with heavy tailed and band random matrices the fact of not being "white", *i.e.* to keep a strong trace of the basis it has been written in. With Djalil Chafaï and Camille Mâle, we would like, establishing a Schwinger-Dyson Formula, to obtain the convergence of the empirical spectral distribution of such matrices.

Interacting particle systems

The *simple exclusion process*, introduced by Spitzer in 1970, is one of the simplest statistical mechanics model. There exists several versions of this model, all relying on the following principle: some particles are located at various points of a line and they move to the right and/or the left, independently up to the constraint that the moves of each particle are restricted by the fact that it cannot overtake its neighbors. Two aspects of this process have drawn my attention recently.

Firstly, for several models of this type, the law of the positions of the particles coincides with the one of the largest eigenvalues of certain random matrices (see, for example, [99] or Theorem 2.7, related to these issues).

Besides, these processes induce a dynamic on Young diagrams. Philippe Biane, in his article [22], has proved that random Young diagrams and free probabilities are linked, via the representations of large symmetric groups. More specifically, he has proved that certain large Young diagrams have limit shapes that can be expressed simply via free probabilities. The dynamics of his process is not the one inherited from exclusion processes, but the tools he brought to light seem to have a certain relevance in this context too.

Both of these subjects interest me a lot. In particular, I have the project, with Damien Simon, to use Biane's ideas to find out the limit shape of the Young diagram associated to the totally asymmetric simple exclusion process (TASEP) starting with initial conditions that are more general than the well known case "all left sites occupied, no right site occupied".

Eigenvectors of band random matrices

Band matrices are one of the simplest models for the Anderson phase transition. In this context, it can be expressed in the following way. Let us consider a (large) $n \times n$ band matrix, whose band has width ℓ: if $\ell \gg \sqrt{n}$, the eigenvectors are delocalized and the eigenvalues have statistics close to the ones of Wigner matrices, whereas if $\ell \ll \sqrt{n}$, the eigenvectors are localized and the eigenvalues have statistics close to the statistics of a Poisson process. This is only a conjecture, though a few recent advances (*c.f.* the works of Erdös and Knowles [46, 47], of Schenker [101] or of Sodin [106]). I would like to try to use the method that I used in my paper [A19] to study these questions.

Maximum of correlated variables and Tracy-Widom laws

Insert 2.4 of this text, devoted to the Tracy-Widom laws, explains that these laws seem to replace the usual max-stable laws when we consider supremums of repulsing variables. I am very interested in this subject, and I would like to study its connections with the theory of convex bodies and log-concave measures (see for example the paper of Klartag [75]).

Chapter 9

Appendice : introduction aux probabilités libres

9.1 Espaces de probabilités non commutatifs et liberté

Un *espace de probabilités non commutatif* (en abrégé *e.p.n.c.*) est[1] un couple $(\mathcal{A}, \varphi)$, où $\mathcal{A}$ est une algèbre sur $\mathbb{C}$, unifère[2] et involutive[3], et φ est une forme linéaire sur $\mathcal{A}$, appelée *état*, valant 1 en $1_\mathcal{A}$ telle que $\varphi(xy) = \varphi(yx)$ et $\varphi(x^*) = \overline{\varphi(x)}$ pour tout x, y dans $\mathcal{A}$. On suppose aussi que la forme sesquilinéaire hermitienne $(x, y) \longmapsto \varphi(xy^*)$ est définie positive. On appelle *variables aléatoires non commutatives* les éléments de $\mathcal{A}$. Une variable aléatoire non commutative est dite *constante* si elle est proportionnelle à $1_\mathcal{A}$.

Soit $(\mathcal{A}, \varphi)$ un e.p.n.c. Une famille $(\mathcal{A}_i)_{i\in I}$ de sous-algèbres[4] de $\mathcal{A}$ est dite *libre* si pour tout $n \geq 1$, pour tous $i_1, \ldots, i_n \in I$ tels que $i_1 \neq i_2, i_2 \neq i_3, \ldots, i_{n-1} \neq i_n$, pour tout $(x_1, \ldots, x_n) \in \mathcal{A}_{i_1} \times \cdots \times \mathcal{A}_{i_n}$, on a

$$\varphi(x_1) = \cdots = \varphi(x_n) = 0 \Longrightarrow \varphi(x_1 \cdots x_n) = 0.$$

Une famille de parties de $\mathcal{A}$, ou de variables aléatoires non commutatives est dite *libre* si celles-ci sont contenues dans des sous-algèbres libres.

[1] Dans certains articles, toutes ces hypothèses ne sont pas faites, et dans ce cadre particulier, on parle alors de $*$-espace de probabilités non commutatif tracial et fidèle. Notons que si la forme sesquilinéaire hermitienne associée à φ n'est que supposée positive, en quotientant l'algèbre par l'idéal bilatère contenu dans $\ker\varphi$ des $a \in \mathcal{A}$ tels que $\varphi(aa^*) = 0$, on obtient une forme définie positive.

[2] Une algèbre $\mathcal{A}$ est dite *unifère* si elle possède un élément neutre $1_\mathcal{A}$ pour la multiplication.

[3] Une algèbre $\mathcal{A}$ sur $\mathbb{C}$ est dite *involutive* si elle est munie d'une involution $x \longmapsto x^*$ antilinéaire telle que pour tous $x, y \in \mathcal{A}$, $(xy)^* = y^*x^*$.

[4] Par convention, on désignera ici par *sous-algèbre* de $\mathcal{A}$ une sous-algèbre contenant $1_\mathcal{A}$ et stable par $x \longmapsto x^*$.

Si $\mathcal{A}$ est de plus munie d'une structure de C^*-algèbre (dans ce cas, φ est toujours continue car elle est positive), alors on parle de *C^*-espace de probabilités non commutatif.* On vérifie aisément que si des sous-algèbres sont libres, alors leurs adhérences le sont.

Enfin, si $\mathcal{A}$ est en plus une W^*-algèbre et si φ est normale (i.e. continue pour la topologie σ-faible d'opérateurs, ce qui équivaut au fait que si $(x_i)_{i\in I}$ est un réseau croissant d'éléments positifs de $\mathcal{A}$ qui converge faiblement vers x, alors $\varphi(x_i)$ tend vers $\varphi(x)$), on parle de *W^*-espace de probabilités non commutatif.* On vérifie, avec le théorème de densité de Kaplansky, que si des sous-algèbres sont libres, alors leurs bicommutants (i.e. W^*-algèbres engendrées) le sont.

Example 9.1 *1. Si $(\Omega, \Sigma, \mathbb{P})$ est un espace de probabilité, toute algèbre stable par conjugaison de variables aléatoires complexes admettant des moments, munie de l'état donné par l'espérance, est un e.p.n.c.. Si de plus, Ω est un espace topologique compact (resp. localement compact) et Σ est la tribu borélienne, alors en se restreignant à l'algèbre des fonctions continues (resp. à L^∞), on obtient un C^*- (resp. W^*-)e.p.n.c. agissant sur $L^2(\Omega, \mathbb{P})$. Tout C^*- (resp. W^*-)e.p.n.c. dans lequel l'algèbre est commutative est de cette forme.*

2. *Un autre exemple classique est celui de l'algèbre $M_n(\mathbb{C})$ des matrices $n \times n$ complexes munie de la trace normalisée* $\mathrm{tr} := \frac{1}{n}\,\mathrm{Tr}$.

3. *On peut faire le produit tensoriel des deux exemples précédents en considérant une algèbre de variables aléatoires définies sur un espace de probabilités classique, à valeurs dans $M_n(\mathbb{C})$, dont toutes les coordonnées possèdent des moments, que l'on munit de l'état* $\mathbb{E} \circ \mathrm{tr}$.

À moins que l'une d'entre elles ne soit constante, si deux variables aléatoires non commutatives a, b sont libres, alors elles ne commutent pas. En effet, on doit avoir, quitte à retrancher à a et b leurs images par φ, $\varphi(ab^*a^*b) = 0$ et $\varphi(aa^*b^*b) = \varphi(aa^*)\varphi(b^*b) > 0$. De plus, comme on le verra au paragraphe suivant, l'espace vectoriel qu'elles engendrent est de dimension infinie. On en déduit donc que l'on n'a pas de sous-algèbres libres non triviales dans les deux premiers exemples. Dans le troisième, cela reste vrai, mais c'est plus difficile à démontrer.

9.2 Structures de dépendance et produit libre d'espaces de probabilités non commutatifs

Une façon de décrire une structure de dépendance entre deux sous-algèbres $\mathcal{A}_1$ et $\mathcal{A}_2$ d'un espace de probabilités non-commutatif est de donner une règle qui permette de calculer $\varphi(P(a_1, a_2))$ pour tous $a_1 \in \mathcal{A}_1$, $a_2 \in \mathcal{A}_2$ et tout polynôme à deux variables non commutatives P connaissant les moments de a_1 et a_2. C'est, indirectement, de cette façon que nous avons défini la liberté. C'est aussi de cette nature qu'est la règle $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$ qui caractérise l'indépendance de deux variables aléatoires. Une telle description n'assure cependant pas l'existence d'exemples non triviaux de la structure de dépendance qu'on veut définir.

En probabilités classiques, la construction qui assure l'existence de variables indépendantes de lois arbitraires est celle du produit cartésien d'espaces mesurables muni du produit tensoriel des lois. Du point de vue non pas des espaces de probabilités mais des variables aléatoires, cette construction correspond au produit tensoriel des algèbres. Considérons par exemple un espace de probabilités $(\Omega, \mathbb{P})$ et notons $(\mathcal{A}, \varphi)$ l'espace de probabilités non-commutatif (commutatif !) $(L^\infty(\Omega, \mathbb{P}), \mathbb{E})$. Donnons-nous deux variables aléatoires bornées X et Y sur Ω. Notons $\mathcal{A}_X$ la sous-algèbre de $\mathcal{A}$ engendrée par X, c'est-à-dire l'ensemble des polynômes en X, et φ_X la forme linéaire induite par l'espérance sur $\mathcal{A}_X$. Alors $(\mathcal{A}_X, \varphi_X)$ est un e.p.n.c.. Considérons de façon analogue $(\mathcal{A}_Y, \varphi_Y)$. Il existe un unique morphisme d'algèbre $f : \mathcal{A}_X \otimes \mathcal{A}_Y \to \mathcal{A}$ tel que $f(X \otimes 1) = X$ et $f(1 \otimes Y) = Y$, et l'image de ce morphisme est exactement l'algèbre engendrée par X et Y. On peut *définir* un état $\varphi_X \otimes \varphi_Y$ sur $\mathcal{A}_X \otimes \mathcal{A}_Y$ en posant $(\varphi_X \otimes \varphi_Y)(a \otimes b) = \varphi_X(a)\varphi_Y(b)$. Cette étape correspond à la construction du produit tensoriel de deux mesures. On peut alors, pour tout élément c de $\mathcal{A}_X \otimes \mathcal{A}_Y$, comparer son espérance dans $\mathcal{A}_X \otimes \mathcal{A}_Y$ et dans $\mathcal{A}$, *i.e.* comparer $(\varphi_X \otimes \varphi_Y)(c)$ et $\varphi(f(c))$. La proposition suivante est une caractérisation purement algébrique de l'indépendance :

> *Les variables X et Y sont indépendantes si et seulement si le morphisme d'algèbres $f : \mathcal{A}_X \otimes \mathcal{A}_Y \to \mathcal{A}$ préserve les espérances, c'est-à-dire si $\varphi_X \otimes \varphi_Y = \varphi \circ f$.*

En généralisant de façon évidente la définition de $\varphi_X \otimes \varphi_Y$ à un produit quelconque $\otimes_{i \in I} \varphi_{|A_i}$, nous en tirons la définition générale suivante.

Definition 9.2 *Soit $(\mathcal{A}, \varphi)$ un espace de probabilités non-commutatif et $(\mathcal{A}_i)_{i \in I}$ une famille de sous-algèbres de $\mathcal{A}$. On dit que cette famille est* indépendante *si pour tout $i \neq j$, tout élément de $\mathcal{A}_i$ commute à tout élément de $\mathcal{A}_j$ et si le morphisme d'algèbres naturel $f : \otimes_{i \in I} \mathcal{A}_i \to \mathcal{A}$ satisfait l'égalité $\varphi \circ f = \otimes_{i \in I} \varphi_{|A_i}$.*

La construction algébrique qui correspond à la liberté est le produit *libre* d'algèbres. Si $(\mathcal{A}_i)_{i \in I}$ est une famille d'algèbres unifères, leur produit libre est la plus grosse algèbre unifère engendrée par les $\mathcal{A}_i$. Plus précisément, le produit libre des algèbres $(\mathcal{A}_i)_{i \in I}$ est défini à isomorphisme près comme l'unique algèbre $\mathcal{A}$ telle que pour tout i, on a une injection d'algèbres $\iota_i : \mathcal{A}_i \hookrightarrow \mathcal{A}$, et telle que si $\mathcal{B}$ est une algèbre unifère telle que pour tout i, il existe un morphisme $f_i : \mathcal{A}_i \to \mathcal{B}$, alors il existe un unique morphisme $f : \mathcal{A} \to \mathcal{B}$ tel que pour tout i, $f \circ \iota_i = f_i$. Cette algèbre, que l'on note $\star_{i \in I} \mathcal{A}_i$, se construit aisément : si pour tout i, $\mathcal{A}_i^\circ$ est un supplémentaire, dans $\mathcal{A}_i$, de $\mathbb{C} \cdot 1_{\mathcal{A}_i}$, le produit libre des $\mathcal{A}_i$ est l'espace vectoriel

$$\mathbb{C} \cdot 1 \oplus \bigoplus_{n \geq 1} \bigoplus_{i_1 \neq \cdots \neq i_n} \mathcal{A}_{i_1}^\circ \otimes \cdots \otimes \mathcal{A}_{i_n}^\circ, \tag{9.1}$$

sur laquelle le produit se définit de façon naturelle (en tenant compte du fait que le produit de deux éléments d'un même $\mathcal{A}_i^\circ$ n'appartient pas forcément à $\mathcal{A}_i^\circ$). Si les algèbres $\mathcal{A}_i$ sont des $*$-algèbres, on peut munir le produit libre d'une involution qui en fera un *produit libre de $*$-algèbres.*

Étant donné une famille $(\mathcal{A}_i, \varphi_i)_{i\in I}$ d'espaces de probabilités non commutatifs, il existe sur $\star_{i\in I}\,\mathcal{A}_i$ un unique état ψ conférant au produit une structure d'e.p.n.c. tel que pour tout i, $\varphi_i = \psi \circ \iota_i$ et tel que les sous-algèbres $\iota_i(\mathcal{A}_i)$ sont libres dans le produit libre muni de l'état ψ. Cet état ψ, noté $\star_{i\in I}\,\varphi_i$, est défini de la façon suivante : dans la construction précédente, on choisit, pour tout i, $\mathcal{A}_i^\circ = \ker \varphi_i$, et on définit ψ par $\psi(1) = 1$ et ψ est nulle sur la somme située à droite de $\mathbb{C} \cdot 1$ dans (9.1). Notons que cet état sur le produit libre des algèbres unifères est, en un sens, le plus naturel. L'e.p.n.c. obtenu est appelé *produit libre* des $(\mathcal{A}_i, \varphi_i)$.

L'analogue de la caractérisation de l'indépendance donnée plus haut est alors le suivant :

> *Une famille $(\mathcal{A}_i)_{i\in I}$ de sous-algèbres d'un même e.p.n.c. $(\mathcal{A}, \phi)$ est libre si et seulement si en notant $\psi = \star_{i\in I}\,\varphi_i$, le morphisme canonique f du produit libre des $\mathcal{A}_i$ dans $\mathcal{A}$ satisfait $\psi = \phi \circ f$.*

Example 9.3 (Algèbre de groupes) *Pour G un groupe, on définit définit* l'agèbre de G $\mathbb{C}[G] = \mathbb{C}^{(G)}$ *de G comme l'ensemble des fonctions $G \to \mathbb{C}$ de support fini (une telle fonction est notée $\sum_{g\in G} x_g \cdot g$) muni de sa structure d'espace vectoriel complexe canonique, du produit de convolution :*

$$\sum_{g\in G} x_g \cdot g \;\times\; \sum_{g\in G} y_g \cdot g = \sum_{g\in G} \left(\sum_{h,k\in G,\, hk=g} x_h y_k \right) \cdot g$$

et de la conjugaison

$$\left(\sum_{g\in G} x_g \cdot g \right)^* = \sum_{g\in G} \overline{x_g} \cdot g^{-1}.$$

On peut munir $\mathbb{C}[G]$ d'une structure d'e.p.n.c. avec l'état tracial φ_G défini par

$$\varphi_G \left(\sum_{g\in G} x_g \cdot g \right) = x_e,$$

où e désigne l'élément neutre de G.

Pour $(G_i)_{i\in I}$ une famille de groupes, l'e.p.n.c. associé au groupe produit cartésien *des G_i s'identifie naturellement à l'e.p.n.c. $(\otimes_{i\in I}\mathbb{C}[G_i], \otimes_{i\in I}\varphi_{G_i})$, alors que l'e.p.n.c. associé au groupe* produit libre *des G_i s'identifie naturellement à l'e.p.n.c. $(\star_{i\in I}\mathbb{C}[G_i], \star_{i\in I}\varphi_{G_i})$.*

Disons maintenant quelques mots sur les pendants topologiques de ces constructions. On peut définir, de la même façon, le produit tensoriel et le produit libre de C^*- (resp. W^*-) e.p.n.c.. Dans l'idée, la construction est dans la continuité de la précédente : on considère le produit tensoriel ou libre des e.p.n.c., on le fait agir sur un espace de Hilbert, et on passe à l'adhérence. Cela dit, la construction détaillée est loin d'être évidente. Donnons ici un résultat central dans cette construction : on peut montrer, avec la construction GNS, que tout C^*-e.p.n.c. (resp. tout W^*-e.p.n.c.) se plonge, en tant que C^*-algèbre (resp. W^*-algèbre), dans l'algèbre $\mathcal{B}(H)$ des opérateurs bornés sur un certain espace de Hilbert dans lequel on peut trouver un vecteur ξ tel que $\varphi(\cdot) = \langle \cdot \xi, \xi \rangle$.

9.3 Distributions de variables aléatoires non commutatives

Soit I un ensemble. Définissons l'algèbre $\mathbb{C}\langle X_i, i \in I \rangle$ des polynômes à variables non commutatives indexées par I : il s'agit tout simplement du produit libre des algèbres unifères $\mathbb{C}[X_i]$, où i varie dans I.

La *distribution* d'une famille $(a_i)_{i\in I}$ de variables aléatoires non commutatives d'un même e.p.n.c. $(\mathcal{A}, \varphi)$, est la forme linéaire

$$\mu_{(a_i)_{i\in I}} : P \in \mathbb{C}\langle X_i, i \in I \rangle \mapsto \varphi(P(a_i, i \in I)) \in \mathbb{C}.$$

Notons que par hypothèse, la suite des moments $\varphi(a^k)$, $k \geq 0$ d'un élément auto-adjoint a de $\mathcal{A}$ est une suite positive au sens de [1], donc la suite des moments d'une mesure de probabilité sur la droite réelle : μ_a est l'intégration par rapport à une loi de probabilité sur $\mathbb{R}$, que l'on appellera, lorsqu'elle est unique, *loi spectrale*, ou *distribution* de a. Elle est unique lorsque l'on travaille dans un C^*-e.p.n.c.. L'élément a est dit *positif* si μ_a est l'intégration par rapport à une mesure portée par $\mathbb{R}^+$.

Example 9.4 *1. Si $(\mathcal{A}, \varphi)$ est $(M_n(\mathbb{C}), \mathrm{tr})$, alors pour toute matrice auto-adjointe a, μ_a est la loi empirique sur le spectre de a.*

2. Si $\mathcal{A}$ est une algèbre de variables aléatoires définies sur un espace de probabilités classique, à valeurs dans $M_n(\mathbb{C})$, dont tous les coefficients possèdent des moments exponentiels, que l'on munit de l'état $\mathbb{E} \circ \mathrm{tr}$, alors pour toute matrice aléatoire hermitienne a de $\mathcal{A}$, μ_a est l'espérance de la loi spectrale de a, i.e. la mesure d'intensité de la mesure aléatoire donnée par la loi spectrale de a.

3. Si $\mathcal{A}$ est une sous-algèbre de $\mathcal{B}(H)$ pour un espace de Hilbert H et si $\varphi(\cdot) = \langle \cdot \xi, \xi \rangle$, μ_a est la mesure de probabilité dont la valeur sur tout borélien B est $\langle p_B(a)\xi, \xi \rangle$, où $p_B(a)$ est le projecteur spectral sur B pour a. Cette mesure est donc portée par le spectre de a.

Remarque 9.5 On peut montrer que si deux familles $(a_i, a_i^*)_{i\in I}$ et $(b_i, b_i^*)_{i\in I}$ de deux e.p.n.c. (resp. C^*-e.p.n.c., W^*-e.p.n.c.) $(\mathcal{A}, \varphi)$ et $(\mathcal{B}, \psi)$ ont la même distribution et engendrent $\mathcal{A}$ et $\mathcal{B}$ comme $*$-algèbres unifères (resp. comme C^*-algèbres, comme W^*-algèbres), alors on peut trouver un isomorphisme de $*$-algèbres (resp. un isomorphisme isométrique de C^*-algèbres, un isomorphisme isométrique de C^*-algèbres continu pour les topologies σ-faibles d'opérateurs) entre $\mathcal{A}$ et $\mathcal{B}$ qui envoie chaque a_i sur b_i, et préserve les états. Cette remarque nous éclaire sur l'intérêt que peut avoir cette théorie dans la perspective de la classification de ces algèbres.

Notons que si $(\mathcal{A}_i)_{i\in I}$ est une famille libre de sous-algèbres de $(\mathcal{A}, \varphi)$ et que si pour tout i, $a_i \in \mathcal{A}_i$, alors la distribution de la famille $(a_i)_{i\in I}$ dans $(\mathcal{A}, \varphi)$ et dans le produit libre des $(\mathcal{A}_i, \varphi_{|\mathcal{A}_i})$ est la même (par hypothèse de préservation de l'état). On en déduit le résultat suivant : *la distribution d'une famille libre ne dépend que des distributions individuelles.* Ce résultat est l'analogue du résultat de probabilités classiques qui dit que la loi jointe d'une famille de variables aléatoires indépendantes ne dépend que de leurs lois individuelles. Il sera central pour répondre à la question posée plus haut à propos des matrices aléatoires.

9.4 Liberté asymptotique des matrices aléatoires carrées

Si pour tout $n \geq 1$, $(a_i(n))_{i\in I}$ est une famille d'éléments d'un e.p.n.c. $(\mathcal{A}(n), \varphi_n)$, on dit que $(a_i(n))_{i\in I}$ *converge en distribution* vers une famille $(a_i)_{i\in I}$ d'éléments d'un e.p.n.c. $(\mathcal{A}, \varphi)$ si les distributions convergent point par point. La famille $(a_i(n))_{i\in I}$ est dite *asymptotiquement libre* si elle converge en distribution vers une famille libre.

On étend la notion de convergence en distribution (et donc la notion de liberté asymptotique) aux matrices aléatoires de la façon suivante : si pour tout $n \geq 1$, $(A_i(n))_{i\in I}$ est une famille de matrices aléatoires $n \times n$, on dit que la famille *converge en distribution en probabilité* vers une famille $(a_i)_{i\in I}$ d'éléments d'un e.p.n.c. $(\mathcal{A}, \varphi)$ si pour tout $P \in \mathbb{C}\langle X_i, i \in I\rangle$, la variable aléatoire classique donnée par la trace normalisée tr $P(A_i(n), i \in I)$ converge en probabilité vers le nombre $\varphi(P(a_i, i \in I))$.

Les résultats liant probabilités libres et matrices aléatoires caractérisent la distribution non commutative **jointe** de familles de matrices aléatoires à partir de lois spectrales **individuelles** [120, 118].

Theorem 9.6 *Soit, pour tout $n \geq 1$, $(H_i)_{i\in I}$ une famille de matrices aléatoires $n \times n$ hermitiennes*[5]*, dont les coordonnées ont des moments à tous les ordres. On suppose :*

(a) la famille $(H_i)_{i\in I}$ est indépendante,

[5] La dépendance en n des matrices $H_i = H_i(n)$ est maintenue implicite ici afin de ne pas alourdir les notations. Il en sera de même des matrices D_j de la remarque 9.8 et pour A et B au paragraphe 9.5.

(b) pour tout $i \in I$, H_i a une loi invariante par conjugaison par les matrices unitaires ou bien est une matrice de Wigner,

(c) pour tout $i \in I$, lorsque $n \longrightarrow \infty$, la loi spectrale empirique de H_i converge faiblement en probabilité vers une loi à support compact déterministe.

Alors la famille $(H_i)_{i\in I}$ est asymptotiquement libre : il existe une famille libre $(a_i)_{i\in I}$ d'éléments auto-adjoints d'un e.p.n.c. $(\mathcal{A}, \varphi)$ telle que pour tout $P \in \mathbb{C}\langle X_i, i \in I\rangle$, la variable aléatoire $\operatorname{tr} P(H_i, i \in I)$ *converge en probabilité, lorsque $n \longrightarrow \infty$, vers le nombre $\varphi(P(a_i, i \in I))$.*

Remarque 9.7 Ce théorème reste vrai en remplaçant l'hypothèse (ii) par le fait que les matrices sont du type $H_i = X_i X_i^*$, où X_i est une matrice rectangulaire $n \times p$, à coefficients i.i.d., avec $n/p \longrightarrow c > 0$ [68, 31, 32].

Remarque 9.8 (Liberté asymptotique de matrices aléatoires avec des matrices déterministes) Ce théorème peut être amélioré de la façon suivante : si pour tout $n \geq 1$, on considère aussi une famille $(D_j)_{j\in J}$ de matrices déterministes $n \times n$, qui converge en distribution, lorsque $n \longrightarrow \infty$, vers une famille $(d_j)_{j\in J}$, alors la famille

$$(H_i)_{i\in I}, (D_j)_{j\in J},$$

indexée par l'union disjointe $I \cup J$, converge en distribution en probabilité vers la famille

$$(a_i)_{i\in I},\ (d_j)_{j\in J},$$

où les ensembles $\{a_i\}_{i\in I}$, $\{d_j\,;\, j \in J\}$ sont libres.

9.5 Convolutions libres $\boxplus$ et $\boxtimes$

Le théorème précédent permet d'affirmer que pour A, B des matrices aléatoires hermitiennes $n \times n$ indépendantes telles que A ou B est invariante, en loi, par conjugaison par n'importe quelle matrice unitaire et telles qu'il existe μ_1, μ_2 lois sur $\mathbb{R}$ telles que pour la convergence faible en probabilités,

$$\frac{1}{n} \sum_{\lambda \text{ val. pr. de } A} \delta_\lambda \longrightarrow \mu_1 \qquad \text{et} \qquad \frac{1}{n} \sum_{\lambda \text{ val. pr. de } B} \delta_\lambda \longrightarrow \mu_2,$$

il existe une loi μ sur $\mathbb{R}$, ne dépendant que de μ_1 et μ_2 telle que pour la convergence faible en probabilités,

$$\frac{1}{n} \sum_{\lambda \text{ val. pr. de } A+B} \delta_\lambda \longrightarrow \mu.$$

La loi μ ainsi définie est alors appelée la *convolution additive libre* des lois μ_1 et μ_2. On la note $\mu_1 \boxplus \mu_2$.

Lorsque A et B sont en plus supposées positives, il existe aussi une loi, notée $\mu_1 \boxtimes \mu_2$ et appelée *convolution multiplicative libre* de μ_1 et μ_2, telle que pour la convergence faible en probabilités,

$$\frac{1}{n} \sum_{\lambda \text{ val. pr. de } AB} \delta_\lambda \longrightarrow \mu_1 \boxtimes \mu_2.$$

9.6 Cumulants libres et R-transformée

Il est naturel, une fois les convolutions libres définies, de chercher des transformations intégrales qui les linéarisent. Elles sont notamment utiles pour étendre à l'ensemble des lois de probabilités des résultats qui, comme ceux du paragraphe précédent, se démontrent d'abord pour les lois à support compact. Dans le contexte des probabilités classiques, la convolution additive est linéarisée par le logarithme de la transformée de Fourier. Les coefficients du développement en série entière du logarithme de la transformée de Fourier de variables aléatoires classiques sont donc des fonctions additives de variables aléatoires indépendantes. Ils sont appelés *cumulants classiques* des variables aléatoires. L'objet de ce paragraphe est de rappeler la définition de ces cumulants classiques, puis de présenter leurs analogues libres. Ils ont été construits et étudiés essentiellement par Speicher et Nica [108, 109, 90], dans des travaux qui furent le point de départ de l'approche combinatoire des probabilités libres, laquelle s'est révélée très fructueuse.

9.6.1 Cumulants classiques

Soit μ une mesure de probabilités à support compact sur $\mathbb{R}$. Les cumulants (classiques) de μ sont les nombres $(c_n^*(\mu))_{n\geq 1}$ définis par l'égalité

$$\log \int_{\mathbb{R}} e^{zt} \mathrm{d}\mu(t) = \sum_{n\geq 1} c_n^*(\mu) \frac{z^n}{n!}.$$

Ils linéarisent la convolution : si μ et ν sont à support compact, on a $c_n^*(\mu * \nu) = c_n^*(\mu) + c_n^*(\nu)$ pour tout $n \geq 1$.

Les cumulants sont reliés de façon combinatoire aux moments. Introduisons l'ensemble $\mathrm{Part}(n)$ des partitions de $\{1, \dots, n\}$. Pour tout $n \geq 1$, notons $m_n(\mu) = \int_{\mathbb{R}} t^n \mathrm{d}\mu(t)$ le n-ième moment de μ. Alors la relation qui lie moments et cumulants est la suivante: pour tout $n \geq 1$,

$$m_n(\mu) = \sum_{\pi \in \mathrm{Part}(n)} \prod_{\substack{B \text{ bloc de } \pi \\ B = \{i_1 < \cdots < i_r\}}} c_r^*(\mu). \tag{9.2}$$

Cette relation se généralise à plusieurs variables de la façon suivante : pour $X_1, \ldots, X_d$ variables aléatoires réelles bornées, les cumulants $c^*(X_{i_1}, \ldots, X_{i_n})$ sont définis par la relation

$$\log \mathbb{E}[e^{z_1 X_1 + \cdots + z_d X_d}] = \sum_{n \geq 1} \frac{1}{n!} \sum_{1 \leq i_1, \ldots, i_n \leq d} c^*(X_{i_1}, \ldots, X_{i_n}) z_{i_1} \cdots z_{i_n} \tag{9.3}$$

et par le fait que $c^*(X_{i_1}, \ldots, X_{i_n})$ ne dépend pas de l'ordre de ses arguments. L'intérêt des cumulants vient alors du fait que, en plus de leur rôle dans la formule (9.3) du logarithme de la transformée de Fourier, ils caractérisent l'indépendance : on voit immédiatement par (9.3) que $X_1, \ldots, X_d$ *sont indépendants if and only if leurs cumulants mixtes s'annulent*, i.e. si $c^*(X_{i_1}, \ldots, X_{i_n}) = 0$ dès que l'on n'a pas $i_1 = \cdots = i_n$.

Voici la version multidimensionnelle de la relation (9.2) : pour $Z_1, \ldots, Z_n$ variables aléatoires réelles bornées,

$$\mathbb{E}[Z_1 \cdots Z_n] = \sum_{\pi \in \mathrm{Part}(n)} \prod_{\substack{B \text{ bloc de } \pi \\ B = \{i_1 < \cdots < i_r\}}} c^*(Z_{i_1}, \ldots, Z_{i_r}). \tag{9.4}$$

En introduisant la fonction de Möbius[6] $\text{Möb}(\cdot, \cdot)$ du treilli $\mathrm{Part}(n)$ muni de la relation d'ordre partiel pour laquelle $\pi \leq \pi'$ si tout bloc de π est contenu dans un bloc de π', la formule (9.4) s'inverse de la façon suivante :

$$c^*[Z_1, \ldots, Z_n] = \sum_{\pi \in \mathrm{Part}(n)} \text{Möb}(\pi, \mathbf{1}_n) \prod_{\substack{B \text{ bloc de } \pi \\ B = \{i_1 < \cdots < i_r\}}} \mathbb{E}[Z_{i_1} \cdots Z_{i_r}], \tag{9.5}$$

où $\mathbf{1}_n$ désigne le plus grand élément de $\mathrm{Part}(n)$, $\{\{1, \ldots, n\}\}$.

9.6.2 Cumulants libres

C'est en adaptant la formule (9.4) au cadre non commutatif des probabilités libres que l'on définit les cumulants libres. Pour $(\mathcal{A}, \varphi)$ espace de probabilités non-commutatif et $a_1, \ldots, a_n \in \mathcal{A}$, la non-commutativité entraîne que l'ordre des éléments dans les expressions du type $\varphi(a_1 \cdots a_n)$ compte. C'est ce qui amène à apporter plus d'attention à la structure des partitions utilisées. On introduit alors les *partitions non croisées* de $\{1, \ldots, n\}$: une partition π de $\{1, \ldots, n\}$ est dite *non croisée* s'il n'existe pas quatre éléments $x < y < z < t$ de E vérifiant $x \overset{\pi}{\sim} z \overset{\pi}{\nsim} y \overset{\pi}{\sim} t$.

On note $\mathrm{NC}(n)$ l'ensemble des partitions non croisées de $\{1, \ldots, n\}$, que l'on munit de l'ordre induit par celui de $\mathrm{Part}(n)$. On note $\text{Möb}_{\mathrm{NC}}(\cdot, \cdot)$ sa fonction de Möbius.

[6] La définition de la fonction de Möbius d'un ensemble fini partiellement ordonné se trouve par exemple au chapitre 10 de [90].

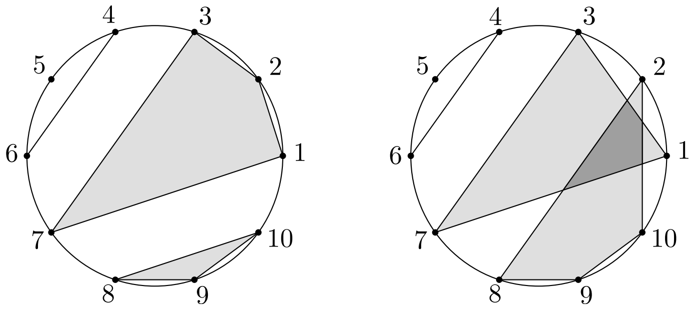


Figure 9.1: La partition $\{\{1,2,3,7\},\{4,6\},\{5\},\{8,9,10\}\}$ est non-croisée alors que $\{\{1,3,7\},\{2,8,9,10\},\{4,6\},\{5\}\}$ est croisée.

On définit alors, récursivement, pour tout $n \geq 1$, la fonction n-linéaire k_n sur $\mathcal{A}^n$ par la formule :

$$\forall a_1,\ldots,a_n \in \mathcal{A}, \qquad \varphi(a_1\cdots a_n) = \sum_{\pi\in \mathrm{NC}(n)} \prod_{\substack{B \text{ bloc de } \pi \\ B=\{i_1<\cdots<i_r\}}} k_r(a_{i_1},\ldots,a_{i_r}),$$

ou, de façon équivalente,

$$k_n(a_1,\ldots,a_n) = \sum_{\pi\in \mathrm{NC}(n)} \text{Möb}_{\mathrm{NC}}(\pi,\mathbf{1}_n) \prod_{\substack{B \text{ bloc de } \pi \\ B=\{i_1<\cdots<i_r\}}} \varphi(a_{i_1}\cdots a_{i_r}). \tag{9.6}$$

Pour $a \in \mathcal{A}$, les nombres $k_n(a,\ldots,a)$, notés $k_n(a)$, sont appelés les *cumulants libres* de a.

Proposition 9.9 *Soit $(\mathcal{A},\varphi)$ un espace de probabilités non commutatif. Deux sous-algèbres $\mathcal{A}_1$ et $\mathcal{A}_2$ de $\mathcal{A}$ sont libres si et seulement si pour tout $n \geq 1$ et tous $a_1,\ldots,a_n$ éléments de $\mathcal{A}_1 \cup \mathcal{A}_2$ qui n'appartiennent ni tous à $\mathcal{A}_1$, ni tous à $\mathcal{A}_2$, on a $k_n(a_1,\ldots,a_n) = 0$.*

En particulier, si a et b sont deux éléments libres de $\mathcal{A}$, on a pour tout $n \geq 1$ l'égalité $k_n(a+b) = k_n(a) + k_n(b)$.

9.6.3 La R-transformée

Ainsi, les cumulants libres linéarisent la convolution libre et caractérisent la liberté. Une transformation intégrale analytique dont les coefficients seraient les cumulants libres permettrait donc de linéariser la convolution additive libre.

Le théorème suivant donne le lien entre la série formelle dont les coefficients sont les *moments* $\varphi(a^n)$ d'une variable aléatoire non commutative a et la série formelle dont les coefficients sont ses cumulants libres.

Theorem 9.10 *Soit $a \in \mathcal{A}$. Alors les séries formelles*

$$G_a(z) := \sum_{n \geq 0} \varphi(a^n) \left(\frac{1}{z}\right)^{n+1} \quad \textit{et } R_a(z) := \sum_{n \geq 0} k_{n+1}(a) z^n$$

sont liées par la relation

$$R_a(z) = G_a^{-1}(z) - \frac{1}{z},$$

où $G_a^{-1}(\cdot)$ désigne l'inverse de $G_a(\cdot)$ pour la composition.

En remarquant que si μ est une loi à support compact dont les moments concident avec ceux de a, $G_a(z)$ concide avec la *transformée de Cauchy* de μ

$$G_\mu(z) := \int_{x \in \mathbb{R}} \frac{\mathrm{d}\mu(x)}{z - x} \qquad (\text{pour } z \notin \text{support}(\mu)), \tag{9.7}$$

on obtient alors (en étendant par densité aux lois à support non compact[7]) la transformation qui linéarise $\boxplus$: c'est la R-transformée de μ, définie par

$$R_\mu(z) = G_\mu^{-1}(z) - \frac{1}{z}. \tag{9.8}$$

Example 9.11 *Pour μ respectivement égale à δ_a, $\frac{1}{2}(\delta_0+\delta_1)$, $\frac{2}{\pi r^2}\sqrt{r^2-(x-m)^2}1_{|x-m|\leq r}\mathrm{d}x$, $L_{\mathrm{MP},c}$, $\frac{\mathrm{d}x}{\pi(1+x^2)}$, on a $R_\mu(z) = a$, $\frac{z-1+\sqrt{1+z^2}}{2z}$, $m + \frac{r^2}{4}z$, $\frac{1}{1-cz}$, $-i$. On peut facilement en déduire que la convolution libre et la convolution classique de la loi de Cauchy avec n'importe quelle loi sont les mêmes, et que la convolution libre de deux lois de Bernouilli symétriques est la loi d'arcsinus $\frac{\mathrm{d}x}{\pi\sqrt{x(2-x)}}$ sur $[0,2]$.*

De la même façon, la convolution multiplicative libre $\boxtimes$ se calcule via une transformation intégrale : en posant

$$T_\mu(z) = \int_{x \in \mathbb{R}} \frac{x\mathrm{d}\mu(x)}{z - x} \qquad (\text{pour } z \notin \text{support}(\mu)), \tag{9.9}$$

et en définissant la S-*transformée* de μ par la formule

$$S_\mu(z) = (1+z)/(zT_\mu^{-1}(z)), \tag{9.10}$$

on a la formule $S_{\mu \boxtimes \nu}(z) = S_\mu(z) S_\nu(z)$ pour toutes lois μ, ν sur $\mathbb{R}_+$.

[7]Dans le cas où μ est à support compact, les fonctions G_μ et R_μ sont analytiques sur des voisinages respectivement de ∞ et de zéro dans $\mathbb{C}$ dont les amplitudes sont contrôlées par le support de μ. Dans le cas général, on a encore des fonctions analytiques, mais elles sont définies sur des *voisinages non tangentiels* de ∞ et de zéro . La notion la plus adéquate est alors celle de *germes* de fonctions analytiques.

9.7 Probabilités libres à valeurs opérateurs

Les notions d'espérance conditionnelle et d'indépendance conditionnelle ont aussi leurs analogues libres. Nous en présentons ici brièvement la théorie [120, 117, 109, 103], qui est la clé de la modélisation que l'on fait du comportement asymptotique des matrices aléatoires rectangulaires.

Si $\mathcal{A}$ est une $*$-algèbre unifère et $\mathcal{D}$ une sous-algèbre de $\mathcal{A}$, une application linéaire $\varphi_{\mathcal{D}} : \mathcal{A} \to \mathcal{D}$ qui envoie $1_{\mathcal{A}}$ sur $1_{\mathcal{A}}$ et satisfait, pour tout $a \in \mathcal{A}, d, d' \in \mathcal{D}$, $\varphi_{\mathcal{D}}(dad') = d\varphi_{\mathcal{D}}(a)d'$, est appelée une *espérance conditionnelle* de $\mathcal{A}$ dans $\mathcal{B}$. Le couple $(\mathcal{A}, \varphi_{\mathcal{D}})$ est alors un *espace de probabilités non commutatif* $\mathcal{D}$*-valué.*

De nombreuses notions des probabilités libres ont leurs analogues "$\mathcal{B}$-valués".

Une famille $(\mathcal{A}_i)_{1\leq i\leq r}$ de sous $*$-algèbres unifères de $\mathcal{A}$ qui contiennent toutes $\mathcal{D}$ est dite *libre avec amalgamation sur* $\mathcal{D}$ si pour tout $n \geq 1$, pour tout $i_1, \ldots, i_n \in I$ tels que $i_1 \neq i_2 \neq \cdots \neq i_n$, pour tout $(x_1, \ldots, x_n) \in \mathcal{A}_{i_1} \times \cdots \times \mathcal{A}_{i_n}$, on a

$$\varphi_{\mathcal{D}}(x_1) = \cdots = \varphi_{\mathcal{D}}(x_n) = 0 \Longrightarrow \varphi_{\mathcal{D}}(x_1 \cdots x_n) = 0.$$

Une famille de parties de $\mathcal{A}$, ou de variables aléatoires non commutatives est dite *libre avec amalgamation sur* $\mathcal{D}$ si elles sont contenues dans des sous-algèbres qui le sont.

Si $(a_i)_{i\in I}$ est une famille d'éléments d'un e.p.n.c. $\mathcal{D}$-valué $(\mathcal{A}, \varphi_{\mathcal{D}})$, la $\mathcal{D}$*-distribution* de la famille $(a_i)_{i\in I}$ est l'application

$$\mu^{\mathcal{D}}_{(a_i)_{i\in I}} : P \in \mathcal{D}\langle X_i, i \in I\rangle \mapsto \varphi_{\mathcal{D}}(P(a_i, i \in I)) \in \mathcal{D},$$

où $\mathcal{D}\langle X_i, i \in I\rangle$ est l'algèbre des polynômes à coefficients dans $\mathcal{D}$, et à variables non commutatives (ni entre elles, ni avec les éléments de $\mathcal{D}$), indexées par I (c'est l'ensemble des sommes finies de termes du type $d_0X_{i_1}d_1X_{i_2} \cdots X_{i_k}d_k$, avec $k \geq 0$, $i_1, \ldots, i_k \in I$ et $d_0, \ldots, d_k \in \mathcal{D}$).

On peut définir, de la même façon qu'au paragraphe 9.2, le *produit libre avec amalgamation sur* $\mathcal{D}$ de n'importe quelle famille d'espaces de probabilités non commutatifs $\mathcal{D}$-valués. On en déduit le même genre de proriétés pour la liberté avec amalgamation qu'au paragraphe 9.2, par exemple le fait de pouvoir construire à sa guise des variables libres avec amalgamation de $\mathcal{D}$-distributions individuelles prescrites, et le fait que *la* $\mathcal{D}$*-distribution d'une famille libre avec amalgamation ne dépend que des* $\mathcal{D}$*-distributions individuelles.* Notons que si l'on a un état φ sur $\mathcal{A}$ tel que $\varphi = \varphi \circ \varphi_{\mathcal{D}}$, la distribution (avec φ) d'une famille d'éléments libres avec amalgamation est entièrement déterminée par les $\mathcal{D}$-distributions individuelles, mais non par les distributions individuelles avec φ.

Enfin, lorsque $\mathcal{D}$ est de dimension finie, on peut définir, comme au paragraphe 9.4, la *convergence en* $\mathcal{D}$*-distribution* et la *convergence en* $\mathcal{D}$*-distribution en probabilité* pour des suites d'éléments de $\mathcal{D}$-espaces de probabilités non commutatifs.

List of published or submitted works

[A1] F. Benaych-Georges *Failure of the Raikov theorem for free random variables.* Séminaire de Probabilités XXXVIII, (2005) 313–320.

[A2] F. Benaych-Georges *Classical and free infinitely divisible distributions and random matrices.* Ann. Probab. Vol. 33, no. 3 (2005) 1134–1170.

[A3] F. Benaych-Georges *Taylor expansions of R-transforms, application to supports and moments.* Indiana Univ. Math. J., Vol. 55, no. 2 (2006) 465–483.

[A4] F. Benaych-Georges *Infinitely divisible distributions for rectangular free convolution: classification and matricial interpretation.* Probab. Theory Related Fields, Vol. 139, issue 1–2 (2007), 143–189.

[A5] F. Benaych-Georges *Rectangular random matrices, related convolution.* Probab. Theory Related Fields, Vol. 144, issue 3 (2009), 471–515.

[A6] F. Benaych-Georges, I. Nechita *A permutation model for free random variables and its classical analogue.* Pacific J. Math., 242 (2009), no. 1, 33–51.

[A7] S.T. Belinschi, F. Benaych-Georges, A. Guionnet *Regularization by Free Additive Convolution, Square and Rectangular Cases.* Complex Anal. Oper. Theory, Vol. 3, no. 3 (2009) 611–660.

[A8] F. Benaych-Georges *Rectangular random matrices, related free entropy and free Fisher's information.* J. Operator Theory, Vol. 62, no. 2 (2009) 371–419.

[A9] F. Benaych-Georges, T. Cabanal-Duvillard *A matrix interpolation between classical and free max operations: I. The univariate case.* J. Theoret. Probab., Vol. 23, no. 2 (2010) 447–465.

[A10] F. Benaych-Georges *On a surprising relation between the Marchenko-Pastur law, rectangular and square free convolutions.* Ann. Inst. Henri Poincaré Probab. Stat., Vol. 46, no. 3 (2010), 644–652.

[A11] F. Benaych-Georges, T. Lévy *A continuous semigroup of notions of independence between the classical and the free one.* Ann. Probab., Vol. 39, no. 3 (2011), 904–938.

[A12] F. Benaych-Georges *Cycles of free words in several independent random permutations with restricted cycle lengths.* Indiana Univ. Math. J., Vol. 59, no. 5 (2010) 1547–1586.

[A13] F. Benaych-Georges *Finite dimensional projections of the Brownian motion on large unitary groups.* To appear in Bull. Soc. Math. France.

[A14] F. Benaych-Georges, R. N. Rao *The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices.* Adv. Math. (2011), Vol. 227, no. 1, 494–521.

[A15] F. Benaych-Georges, R. N. Rao *The singular values and vectors of low rank perturbations of large rectangular random matrices.* Preprint (2011).

[A16] F. Benaych-Georges *Rectangular R-transform at the limit of rectangular spherical integrals.* J. Theoret. Probab., Vol. 24, no. 4 (2011) 969–987.

[A17] F. Benaych-Georges, A. Guionnet, M. Maïda *Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices.* Electron. J. Prob., Vol. 16 (2011), Paper no. 60, 1621–1662.

[A18] F. Benaych-Georges, A. Guionnet, M. Maïda *Large deviations of extreme eigenvalues of finite rank deformations of deterministic matrices.* To appear in Probab. Theory Related Fields (2011).

[A19] F. Benaych-Georges *Eigenvectors of Wigner matrices: universality of global fluctuations.* Preprint (2011).

[A20] F. Benaych-Georges, F. Chapon *Random right eigenvalues of Gaussian quaternionic matrices.* To appear in Random Matrices: Theory and Applications (RMTA).

Unpublished notes:

[N1] F. Benaych-Georges *Cycles of random permutations with restricted cycle lengths.* ArXiv:0712.1903.

[N2] F. Benaych-Georges, N. Enriquez *Perturbations of diagonal matrices by band random matrices.* ArXiv:1104.1983.

Conference act:

[AC1] F. Benaych-Georges, R. N. Rao. *The breakdown point of signal subspace methods.* Proceedings of the 2010 IEEE Sensor, Array, Multichannel Processing Workshop. October 2010, Jerusalem, Israel.

Bibliography

[1] N.I. Akhiezer *The classical moment problem.* Oliver & Boyd, 1965.

[2] G. Anderson, A. Guionnet, O. Zeitouni *An Introduction to Random Matrices.* Cambridge studies in advanced mathematics, 118 (2009).

[3] A. D'Aristotile, P. Diaconis, C. Newman *Brownian motion and the classical groups.* With Probability, Statisitica and their applications: Papers in Honor of Rabii Bhattacharaya. Edited by K. Athreya *et al.* 97–116. Beechwood, OH: Institute of Mathematical Statistics, 2003.

[4] R. Arratia, A.D. Barbour, S. Tavaré *Logarithmic combinatorial structures: a probabilistic approach.* EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2003.

[5] A. Auffinger, G. Ben Arous, S. Péché *Poisson convergence for the largest eigenvalues of heavy tailed random matrices.* Ann. Inst. Henri Poincaré Probab. Stat. 45 (2009), no. 3, 58–610.

[6] Z.D. Bai, J.W. Silverstein *Spectral analysis of large dimensional random matrices.* Second Edition, Springer, New York, 2009.

[7] J. Baik, G. Ben Arous, S. Péché *Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices.* Ann. Prob. 33 (2005) 1643–1697.

[8] S.T. Belinschi. *The Lebesgue decomposition of the free additive convolution of two probability distributions*, Probab. Theory Related Fields 142 (2008), no. 1-2, 125–150.

[9] S.T. Belinschi, A. Dembo, A. Guionnet *Spectral measure of heavy tailed band and covariance random matrices.* Comm. Math. Phys. 289 (2009), no. 3, 1023–10.

[10] G. Ben Arous, A. Dembo, A. Guionnet *Aging of spherical spin glasses.* Probab. Theory Related Fields 120 (2001), no. 1, 1–67.

[11] G. Ben Arous, A. Guionnet *Large deviations for Wigner's law and Voiculescu 's non-commutative entropy.* Probab. Theory Related Fields, 108, (1997) 517–542.

[12] G. Ben Arous, A. Guionnet *The spectrum of heavy tailed random matrices.* Comm. Math. Phys. 278 (2008), no. 3, 715–751.

[13] G. Ben Arous, O. Zeitouni *Large deviations from the circular law.* ESAIM Probab. Statist. 2 (1998), 123–134.

[14] H. Bercovici, V. Pata, with an appendix by P. Biane *Stable laws and domains of attraction in free probability theory.* Annals of Mathematics, 149. (1999) 1023–1060.

[15] H. Bercovici, D. Voiculescu *Lévy-Hinchin type theorems for multiplicative and additive free convolution* Pacific J. Math. 153 (1992), no. 2, 217–248.

[16] H. Bercovici, D. Voiculescu *Free convolution of measures with unbounded supports.* Indiana Univ. Math. J. 42 (1993) 733–773.

[17] H. Bercovici, D. Voiculescu *Regularity questions for free convolution.* Nonselfadjoint operator algebras, operator theory, and related topics, 37–47, Oper. Theory Adv. Appl. 104, Birkhäuser, Basel, 1998.

[18] P. Biane *Brownian motion, free stochastic calculus and random matrices.* Free probability theory (Waterloo, ON, 1995), Fields Inst. Commun., 12, Amer. Math. Soc., Providence, RI, 1997.

[19] P. Biane *On the Free convolution by a semi-circular distribution.* Indiana Univ. Math. J., Vol. 46, (1997), 705–718.

[20] P. Biane *Segal-Bargmann transform, functional calculus on matrix spaces and the theory of semi-circular and circular systems.* J. Funct. Anal. 144 (1997), no. 1, 232–286.

[21] P. Biane *Processes with free increments.* Math. Z., 227, (1998), no. 1, 143–174.

[22] P. Biane *Representations of symmetric groups and free probability*, Adv. Math. 138 (1998), no. 1, 126–181.

[23] P. Biane, M. Capitaine, A. Guionnet *Large deviation bounds for matrix Brownian motion.* Invent. Math. 152 (2003), no. 2, 433?459.

[24] C. Bordenave, P. Caputo, D. Chafaï *Spectrum of large random reversible Markov chains – heavy-tailed weights on the complete graph.* To appear in Ann. Probab.

[25] C. Bordenave, P. Caputo, D. Chafaï *Spectrum of non-Hermitian heavy tailed random matrices.* To appear in Comm. Math. Phys.

[26] É. Borel *Sur les principes de la théorie cinétique des gaz.* Annales de l'École Normale Supérieure 23 (1906), 9–32.

[27] A. Borodin, S. Péché *Airy kernel with two sets of parameters in directed percolation and random matrix theory.* J. Stat. Phys. 132 (2008), no. 2, 275–290.

[28] T. Cabanal-Duvillard *A matrix representation of the Bercovici-Pata bijection.* Electron. J. Probab. 10 (2005), no. 18, 632–661.

[29] T. Cabanal Duvillard, A. Guionnet *Large deviations upper bounds for the laws of matrix-valued processes and non-communicative entropies.* Ann. Probab. 29 (2001), no. 3, 1205–1261.

[30] M. Capitaine *Additive/multiplicative free subordination property and limiting eigenvectors of spiked additive deformations of Wigner matrices and spiked sample covariance matrices.* Preprint, 2011.

[31] M. Capitaine, M. Casalis *Asymptotic freeness by generalized moments for Gaussian and Wishart matrices. Application to beta random matrices.* Indiana Univ. Math. J. 53 (2004), no. 2, 397–431.

[32] M. Capitaine, C. Donati-Martin *Strong asymptotic freeness for Wigner and Wishart matrices.* Indiana Univ. Math. J. 56 (2007), no. 2, 767–803.

[33] M. Capitaine, C. Donati-Martin, D. Féral *The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations.* Ann. Probab. 37 (2009)1–47.

[34] M. Capitaine, C. Donati-Martin, D. Féral *Central limit theorems for eigenvalues of deformations of Wigner matrices.* A paraître aux Annales de L'Institut Henri Poincaré (2011).

[35] M. Capitaine, C. Donati-Martin, D. Féral, M. Février *Free convolution with a semi-circular convolution and eigenvalues of spiked deformations of Wigner matrices.* Preprint 2010.

[36] D. Chafaï *Anderson localization.* Blog note, `http://blog.djalil.chafai.net/2010/09/07/anderson-localization/`

[37] S. Chatterjee, E. Meckes *Multivariate normal approximation using exchangeable pairs.* ALEA 4, 257–283 (2008).

[38] S. Chatterjee, S.R.S. Varadhan *Large deviations for random matrices.* arXiv.

[39] B. Collins *Moments and Cumulants of Polynomial Random Variables on Unitary Groups, the Itzykson-Zuber Integral, and Free Probability.* Int. Math. Res. Not., No. 17, 2003.

[40] B. Collins, J.A. Mingo, P. Śniady, R. Speicher *Second order freeness and fluctuations of random matrices. III. Higher order freeness and free cumulants.* Doc. Math. 12 (2007), 1–70.

[41] B. Collins, P. Śniady *Integration with respect to the Haar measure on unitary, orthogonal and symplectic group.* Comm. Math. Phys. 264 (2006), no. 3, 773–795.

[42] B. Collins, P. Śniady *New scaling of Itzykson-Zuber integrals.* Annales de l'IHP Prob. Stats. 43, 2 (2007), 139–146.

[43] B. Collins, M. Stolz *Borel theorems for random matrices from the classical compact symmetric spaces.* Annals of Probability Volume 36, Number 3 (2008), 876–895.

[44] A.B. Dieker, J. Warren *On the largest-eigenvalue process for generalized Wishart random matrices.* ALEA Lat. Am. J. Probab. Math. Stat. 6 (2009), 369–376.

[45] C. Donati-Martin, A. Rouault *Truncations of Haar unitary matrices, traces and bivariate Brownian bridge.* ArXiv:1007:1366.

[46] L. Erdös, A. Knowles *Quantum Diffusion and Eigenfunction Delocalization in a Random Band Matrix Model.* To appear in Commun. Math. Phys.

[47] L. Erdös, A. Knowles *Quantum Diffusion and Delocalization for Band Matrices with General Distribution.* To appear in Annales Inst. H. Poincaré.

[48] L. Erdös, B. Schlein, H.-T. Yau *Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices.* Ann. Probab. 37 (2009), no. 3, 815–852.

[49] L. Erdös, B. Schlein, H.-T. Yau *Local semicircle law and complete delocalization for Wigner random matrices.* Comm. Math. Phys. 287 (2009), no. 2, 641–655.

[50] L. Erdös, B. Schlein, H.-T. Yau *Wegner estimate and level repulsion for Wigner random matrices.* To appear in Int. Math. Res. Not.

[51] D. Féral *On large deviations for the spectral measure of discrete Coulomb gas.* Séminaire de Probabilités XLI, 19–49, Lecture Notes in Math., 1934, Springer, Berlin, 2008.

[52] D. Féral, S. Péché *The largest eigenvalue of rank one deformation of large Wigner matrices.* Comm. Math. Phys. 272 (2007)185–228.

[53] A. Fey, R. van der Hofstad, M. Klok *Large deviations for eigenvalues of sample covariance matrices, with applications to mobile communication systems.* Adv. in Appl. Probab. 40, (2008) 1048–1071.

[54] P. Flajolet, R. Sedgewick *Analytic combinatorics.* Cambridge University Press, 2009.

[55] P.K. Friz, H. Oberhauser *Rough path limits of Wong-Zakai type with modified drift term.* Journal of Functional Analysis 256 (2009), 3236–3256.

[56] S. Geman *A limit theorem for the norm of random matrices.* Ann. Probab., 8, 252–261, 1980.

[57] A. Ghaderipoor, C. Tellambura *Generalization of some integrals over unitary matrices by character expansion of groups.* J. Math. Physics, 49 073519 (2008).

[58] V. Gnedenko, A.N. Kolmogorov *Limit distributions for sums of independent random variables.* Adisson-Wesley Publ. Co., Cambridge, Mass., 1954.

[59] D. Gregoratti, W. Hachem, X. Mestre *Space-time bock coding for the DF relay channel.* Prépublication, 2011.

[60] A. Guionnet *Large random matrices: lectures on macroscopic asymptotics.* École d'Été de probabilités de Saint-Flour XXXVI–2006, Lect. Notes Math. 1957 (2009).

[61] A. Guionnet, M. Maïda *A Fourier view on the R-transform and related asymptotics of spherical integrals.* J. Funct. Anal. 222, 2, (2005), 435–490.

[62] A. Guionnet, M. Krishnapur, O. Zeitouni *The single ring theorem.* To appear in Ann. Math.

[63] A. Guionnet, O. Zeitouni *Large deviations asymptotics for spherical integrals.* J. Funct. Anal. 188, 2 (2002), 461–515.

[64] U. Haagerup *Unpublished lecture notes at MSRI.* (2001)

[65] U. Haagerup, F. Larsen *Brown's spectral distribution measure for R-diagonal elements in finite von Neumann algebras* Journ. Functional Analysis. 176, 331-367 (2000).

[66] W. Hachem, P. Loubaton, X. Mestre, J. Najim, P. Vallet *A Subspace Estimator for Fixed Rank Perturbations of Large Random Matrices.* Prépublication, 2011.

[67] D.L. Hanson, F.T. Wright *A bound on tail probabilities for quadratic forms in independent random variables.* Ann. Math. Statist., 42, (1971) 1079–1083.

[68] F. Hiai, D. Petz *The semicircle law, free random variables, and entropy.* Amer. Math. Soc., Mathematical Surveys and Monographs Volume 77, 2000.

[69] D.C. Hoyle, M. Rattray. *Statistical mechanics of learning multiple orthogonal signals: asymptotic theory and fluctuation effects.* Phys. Rev. E (3), 75(1):016101, 13, 2007.

[70] N. Ikeda, S. Watanabe *Stochastic differential equations and diffussion processes.* North-Holland, 1981.

[71] T. Jiang *How many entries of a typical orthogonal matrix can be approximated by independent normals?* Ann. Probab. 34(4), 1497–1529. 2006.

[72] K. Johansson *Shape fluctuations and random matrices.* Comm. Math. Phys. 209 (2000), no. 2, 437–476.

[73] I.M. Johnstone *On the distribution of the largest Principal Component.* Ann. Statist. 29 (2001) 295–327.

[74] Y. Kabashima *Inference from correlated patterns: a unified theory for perceptron learning and linear vector channels.* Journal of Physics: Conference Series 95 (2008) 012001.

[75] B. Klartag *A central limit theorem for convex sets.* Invent. Math. 168 (2007), no. 1, 91–131.

[76] A. Knowles, J. Yin *Eigenvector Distribution of Wigner Matrices.* ArXiv:1102.0057.

[77] T. Lévy *Schur-Weyl duality and the heat kernel measure on the unitary group.* Adv. Math. 218 (2008), no. 2, 537–575.

[78] T. Lévy, M. Maïda *Central limit theorem for the heat kernel measure on the unitary group.* J. Funct. Anal. 259 (2010), no. 12, 3163–3204.

[79] N. Linial, D. Puder. *Word maps and spectra of random graph lifts.* Random Structures and Algorithms 37 (2010), no. 1, 100–135.

[80] M. Maïda *Large deviations for the largest eigenvalue of rank one deformations of Gaussian ensembles.* Elec. J. Probab. 12, (2007) 1131–1150.

[81] V.A. Marchenko, L.A. Pastur *The distribution of eigenvalues in certain sets of random matrices.* Mat. Sb. 72 (1967) 507-536; traduc. anglaise USSR Sb. 1 (1967) 457–483.

[82] E. Meckes *Linear functions on the classical matrix groups.* Trans. Amer. Math. Soc. 360 (2008), no. 10, 5355–5366.

[83] J.A. Mingo, A. Nica *Annular noncrossing permutations and partitions, and second-order asymptotics for random matrices.* Int. Math. Res. Not. 2004, no. 28, 1413–1460.

[84] J.A. Mingo, R. Speicher *Second order freeness and fluctuations of random matrices. I. Gaussian and Wishart matrices and cyclic Fock spaces.* J. Funct. Anal. 235 (2006), no. 1, 226–270.

[85] J.A. Mingo, P. Śniady, R. Speicher *Second order freeness and fluctuations of random matrices. II. Unitary random matrices.* Adv. Math. 209 (2007), no. 1, 212–240.

[86] B. Nadler *Finite sample approximation results for principal component analysis: a matrix perturbation approach.* Ann. Statist., 36(6):2791–2817, 2008.

[87] M. Neagu *Asymptotic freeness of random permutation matrices with restricted cycles lengths.* Indiana Univ. Math. J. 56 (2007), 2017–2049.

[88] A. Nica *Asymptotically free families of random unitaries in symmetric groups.* Pacific J. Math. 157 (1993), no. 2, 29–310.

[89] A. Nica *On the number of cycles of a given length of a free word in several random permutations.* Random Structures and Algorithms, Vol. 5, No. 5, 703–730, 1994.

[90] A. Nica, R. Speicher *Lectures on the combinatorics of free probability.* London Mathematical Society Lecture Note Series, 335. Cambridge University Press, Cambridge, 2006.

[91] D. Paul *Asymptotics of sample eigenstructure for a large dimensional spiked covariance model.* Statist. Sinica, 17(4):1617–1642, 2007.

[92] S. Péché *The largest eigenvalue of small rank perturbations of Hermitian random matrices.* Probab. Theory Related Fields, 134, (2006) 127–173.

[93] S. Péché *Universality results for the largest eigenvalues of some sample covariance matrix ensembles.* Probab. Theory Related Fields 143 (2009) 481–516.

[94] S. Péché *The edge of the spectrum of random matrices.* Habilitation à diriger des recherches.

[95] E.M. Rains *Combinatorial properties of Brownian motion on compact classical groups.* J. Theo. Prob., Vol. 10, No. 3, 259–279. 1997.

[96] J.A. Ramírez, B. Rider, B. Virág *Beta ensembles, stochastic Airy spectrum, and a diffusion.* To appear in J. Amer. Math. Soc.

[97] L.C.G. Rogers, D. Williams *Diffusions, Markov processes and martingales: Volume 2, Ito calculus.* J. Wiley & sons, 1987.

[98] A. Ruzmaikina *Universality of the edge distribution of eigenvalues of Wigner random matrices with polynomially decaying distributions of entries.* Comm. Math. Phys. 261 (2006) 277–296.

[99] T. Sasamoto *Fluctuations of the one-dimensional asymmetric exclusion process using random matrix techniques.* J. Stat. Mech. Theory Exp. 2007, no. 7, P07007, 31 pp. (electronic).

[100] K.I. Sato *Lévy processes and infinitely divisible distributions.* Volume 68 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1999.

[101] J. Schenker *Eigenvector localization for random band matrices with power law band width.* Comm. Math. Phys. 290, (2009) 1065–1097.

[102] B. Schlittgen, T. Wettig *Generalizations of some integrals over the unitary group.* J. Phys. A, 36, 3195 (2003).

[103] D. Shlyakhtenko *Random Gaussian band matrices and freeness with amalgamation.* Internat. Math. Res. Notices 1996, no. 20, 1013–1025.

[104] J.W. Silverstein *Weak convergence of random functions defined by the eigenvectors of sample covariance matrices.* Ann. Probab. 18 (1990), no. 3, 1174–1194.

[105] P. Śniady *Random regularization of Brown spectral measure.* J. Funct. Anal. 193 (2002) 291–313.

[106] S. Sodin *The spectral edge of some random band matrices.* Ann. of Math. (2) 172 (2010), no. 3, 2223–2251.

[107] A. Soshnikov *Universality at the edge of the spectrum in Wigner random matrices.* Comm. Math. Phys. 207 (1999) 697–733.

[108] R. Speicher *Multiplicative functions on the lattice of non-crossing partitions and free convolution.* Math. Annalen 298 (1994) 611–628.

[109] R. Speicher *Combinatorial theory of the free product with amalgamation and operator-valued free probability theory.* Mem. Amer. Math. Soc. 132 (1998), no. 627.

[110] F. Spitzer *Interaction of Markov processes.* Advances in Math. 5 1970 246–290 (1970).

[111] T. Tao, V. Vu (with an appendix by M. Krishnapur) *Random matrices: Universality of ESDs and the circular law.* Ann. Probab. Vol. 38, no. 5 (2010), 2023–2065.

[112] T. Tao, V. Vu *Random matrices: universality of local eigenvalue statistics.* ArXiv 0906.0510, to appear in Acta Mathematica.

[113] T. Tao, V. Vu *Random matrices: universality of eigenvectors.* ArXiv 1103.2801.

[114] C. Tracy, H. Widom *Level spacing distribution and Airy kernel.* Commun. Math. Phys. 159 (1994) 151–174.

[115] C. Tracy, H. Widom *On orthogonal and symplectic matrix ensembles.* Commun. Math. Phys. 177 (1996), 727–754.

[116] D. Voiculescu *Limit laws for random matrices and free products.* Invent. Math. 104 (1991), no. 1, 201–220.

[117] D. Voiculescu *Operations on certain non-commutative operator-valued random variables.* Recent advances in operator algebras (Orléans, 1992). Astérisque No. 232 (1995), 243–275.

[118] D. Voiculescu *A strengthened asymptotic freeness result for random matrices with applications to free entropy.* Internat. Math. Res. Notices 1998, no. 1, 41–63.

[119] D. Voiculescu *The analogues of entropy and of Fisher's information measure in free probability theory. VI. Liberation and mutual free information.* Adv. Math. 146 (1999), no. 2, 101–166.

[120] D. Voiculescu, K. Dykema, A. Nica *Free random variables.* CRM Monograghs Series No.1, Amer. Math. Soc., Providence, RI, 1992.

[121] F. Xu *A random matrix model from two-dimentional Yang-Mills theory.* Comm. Math. Phys. 190 (2) 287–307. 1997.

[122] A.L. Yakymiv *The distribution of lengths of the m-th maximal cycle in a random A permutation.* Discrete Math. Appl. 15 (5), 527–546 (2005).

[123] A.L. Yakymiv *Random A-Permutations: Convergence to a Poisson Process.* Mathematical Notes, 2007, Vol. 81, No. 6, pp. 840–846.

[124] I. Zakharevich *A generalization of Wigner's law.* Commun. Math. Phys. 268(2), 403–414 (2006).

[125] P. Zinn-Justin, J.-B. Zuber *On some integrals over the $U(N)$ unitary group and their large N limit.* J. Phys. A 36, 12 (2003) 3173–3193.

[126] A. Zvonkin *Matrix integrals and map enumeration: an accessible introduction.* Math. Comput. Modelling 26, 8–10 (1997), 281–304. Combinatorics and physics (Marseilles, 1995).