# Une introduction à la théorie des matrices aléatoires

Michaël Ulrich *

2012

## Table des matières

**1 Introduction** **2**

**2 Quelques notions utiles par la suite** **3**

**3 Etude des matrices de Wigner centrées** **5**
3.1 Définition des matrices de Wigner . . . . . 5
3.2 Loi des matrices de Wigner et de leur distribution spectrale . . . 5
3.2.1 Algébre multilinéaire . . . . . 7
3.2.2 Suite de la preuve . . . . . 8
3.3 Convergence de la mesure spectrale . . . . . 11
3.4 Principe de grandes déviations pour les matrices de Wigner . . . 15
3.4.1 Pour s'affranchir de M . . . . . 18
3.4.2 Le principe du minimax pour aller plus vite . . . . . 18
3.4.3 Tension exponentielle et fin de la borne sup . . . . . 20
3.4.4 Preuve de la borne inf . . . . . 20

**4 Les intégrales sphériques** **22**
4.1 Présentation et motivation . . . . . 22
4.2 Théorème de base . . . . . 23

**5 Grandes déviations pour des matrices de Wigner non centrées** **27**
5.1 Exemple d'application : Mesure empirique du mouvement brownien 27
5.1.1 Définitions utiles et martingales . . . . . 28
5.1.2 Borne supérieure . . . . . 29
5.1.3 Borne inférieure . . . . . 30
5.2 Principe de grande déviation pour une matrice de Wigner non-centrée . . . . . 31
5.2.1 Quelques définitions . . . . . 31

*Mémoire réalisé sous la direction de M. Lorenzo Zambotti, à qui je souhaite adresser mes sincères remerciements pour l'aide apportée.

5.2.2 Prérequis . . . . . 32
5.2.3 Borne supérieure . . . . . 32
5.2.4 Borne inférieure : le cas de $\mathcal{MC}$ . . . . . 33
5.2.5 Borne inférieure : le cas de $\mathcal{A}$ . . . . . 34
5.2.6 Borne inférieure : conclusion . . . . . 36

**6 Introduction : le concept de théorie des probabilités de variables non-commutatives 36**
6.1 Idée générale . . . . . 36
6.2 Définitions . . . . . 37
6.3 Liberté . . . . . 41
6.4 Transformée-R . . . . . 42
6.5 Calcul stochastique libre . . . . . 44
6.5.1 Construction du mouvement brownien classique . . . . . 44
6.5.2 Construction du mouvement brownien libre . . . . . 45
6.6 Probabilités libres et matrices aléatoires . . . . . 46
6.6.1 Intérêt . . . . . 46
6.6.2 Démonstration des résultats . . . . . 46

**7 Conclusion 48**

**Références 49**

# 1 Introduction

Le concept de matrices aléatoires est connu depuis les années 30 et leur introduction par Wishart, qui étudiait des questions liées aux statistiques. Le physicien Wigner, lui-aussi, a été amené dans les années 50 à introduire des matrices aléatoires, dans le cadre de la mécanique quantique cette fois. Les équations qui interviennent dans cette branche de la physique font en effet intervenir le Hamiltonien, qui est une matrice. Dans certains cas complexes, il est cependant trés difficile de connaitre exactement cette matrice. L'idée de Wigner fut de choisir cette matrice aléatoirement (tout en respectant les contraintes connues).
Dans les deux cas, dans les questions statistiques de Wishart et en mécanique quantique, ce sont les valeurs propres des matrices qui sont particuliérement intéressantes. Nous verrons qu'il est possible de dire des choses intéressantes à ce sujet. Il est ainsi par exemple possible d'obtenir des résultats de grandes déviations. Nous introduirons également le concept de probabilités libres, une théorie analogue aux probabilités "classiques" construite à l'aide d'algèbre d'opérateurs et offrant un cadre conceptuel intéressant pour l'étude de matrices aléatoires.

## 2 Quelques notions utiles par la suite

**Définition 2.1.** *Soient $\mu_n$ une suite de mesures de probabilités et $\mu$ une autre telle mesure. Alors, on dit que $\mu_n$ converge étroitement vers $\mu$ si pour toute fonction continue bornée $\phi$, on a :*

$$\int \phi d\mu_n \to \int \phi d\mu$$

On rappelle le théorème de portmanteau :

**Théorème 2.1.** *Il y a équivalence entre :*

1. *$\mu_n$ converge étroitement vers $\mu$*
2. *$\liminf_{n\to\infty} \mu_n(O) \geq \mu(O)$ pour tout ouvert $O$*
3. *$\limsup_{n\to\infty} \mu_n(F) \leq \mu(F)$ pour tout fermé $R$*
4. *$\lim_{n\to\infty} \mu_n(A) = \mu(A)$ pour tout borélien $A$ dont le bord est de $\mu$-mesure nulle*

Nous supposerons dans la suite être dans un espace polonais $E$, ie $E$ est métrique séparable. Alors, on a le résultat suivant :

**Théorème 2.2** (Lévy-Prokhorov)**.** *La topologie de la convergence étroite est métrisable, par exemple à l'aide de la distance de Lévy-Prokhorov. Soient $\mu$ et $\nu$ deux mesures de probabilités sur $E$, on définit leur distance de la manière suivante :*

$$d(\mu,\nu) := \inf\{\epsilon > 0 : \mu(A) \leq \nu(A^\epsilon) + \epsilon \text{ et } \nu(A) \leq \mu(A^\epsilon) + \epsilon \text{ pour tout } A \text{ borélien de } E\}$$

*où $A^\epsilon$ désigne le $\epsilon$-voisinage de $A$.*

*Démonstration.* c'est une distance, en effet :

- $d$ est clairement à valeurs positives au sens large.
- Pour une probabilité $\mu$, on a $d(\mu,\mu) \leq \epsilon$ pour tout $\epsilon > 0$ car $A \subset A^\epsilon$. Donc $d(\mu,\mu) = 0$.
- Soit $\mu$ tel que $d(\mu,\nu) = 0$. Alors, soit $A$ un borélien et $\epsilon > 0$, on a :

  $$\mu(A) \leq \nu(A^\epsilon) + \epsilon \leq \mu\left(A^{2\epsilon}\right) + 2\epsilon$$

  La dernière inégalité provenant de l'inégalité triangulaire pour la distance de $E$. Or, les $A^\epsilon$ sont décroissants donc $\mu(A^\epsilon)$ tend vers $\mu(A)$ quand $\epsilon$ tend vers 0. D'où $\mu(A) = \nu(A)$ et donc finalement $\mu = \nu$.
- Les deux inégalités définissant $d$ sont clairement symétriques en $\mu$ et $\nu$ et donc $d$ est aussi symétrique.
- Soient $\mu$, $\nu$ et $\eta$ trois mesures de probabilités. On pose en outre, $\epsilon_1 > d(\mu,\nu)$, $\epsilon_2 > d(\nu,\eta)$. Alors, pour $A$ un borélien de $E$, on a :

  $$\mu(A) \leq \nu(A^{\epsilon_1}) + \epsilon_1 \leq \eta\left(A^{\epsilon_1+\epsilon_2}\right) + \epsilon_1 + \epsilon_2$$

Ici encore, c'est l'inégalité triangulaire pour la distance de $E$ qui se cache. L'inégalité dans le sens inverse s'obtenant de la même façon, on a finalement :

$$d(\mu,\nu) \leq \epsilon_1 + \epsilon_2$$

D'où l'inégalité triangulaire voulue en prenant l'infimum sur les $\epsilon_1$ et les $\epsilon_2$.

Il reste à présent à montrer que $d$ est compatible avec la convergence étroite. Remarquons que les

$$U(\mu,\mathcal{F},\epsilon) := \{\nu \text{ tq } |\mu(A) - \nu - (A| < \epsilon \text{ pour tout } A \in \mathcal{F}\}$$

où $\mathcal{F}$ est une famille finie de boréliens dont le bord est de $\mu$-mesure nulle, est une base pour la topologie engendrée par la convergence étroite.

On va chercher à montrer que les $\{\nu : d(\nu,\mu) < \epsilon\}$ contiennent un ensemble $U$ du type précédent, ce qui signifiera alors que la convergence étroite implique la convergence au sens de $d$. Soit $0 < \delta < \epsilon/3$ et $K$ un compact tel que $\mu(K) > 1-\delta$. Comme $E$ est séparable et par compacité, on peut recouvrir $K$ par un nombre fini de boules $(B_i)_{i=1\ldots k}$ de rayon $\delta$. Posons $\mathcal{D}$ l'ensemble des réunions d'un nombre quelconque de ces boules. On pose en outre $D_0 = \bigcup_0^k B_i$. Notons au passage que, quitte à augmenter un peu $\delta$, on peut toujours supposer que la $\mu$-mesure du bord des $B_i$ est nulle.

Alors, $U(\mu,\mathcal{D},\delta)$ convient. En effet, si $\nu$ y appartient, on remarque déjà que :

$$|\nu(D_0 - \mu(D_0)| < \delta$$

et donc $\nu(D_0) > 1 - 2\delta$.

Si, par ailleurs, $A$ est un borélien, on peut poser $B$ la réunion des $B_i$ dont l'intersection avec $A$ est non vide. On a :

$$A = B \cup D_0^c$$

D'où :

$$\nu(A) \leq 2\delta + \nu(B) \leq 3\delta + \mu(B) \leq \epsilon + \mu(A^\epsilon)$$

La dernière inégalité provient du fait que les boules sont de rayon infrieur à $\delta$ et donc $B \subset A^\delta \subset A^\epsilon$. L'autre inégalité intervenant dans la définition de $d$ se montre de la même façon.

Supposons maintenant que $\mu_n$ converge vers $\mu$ au sens de $d$. Soit $f$ une fonction continue bornée positive et $0 < \epsilon$. Soit $N$ tel que, pour tout $N \geq N$, on ait :

$$d(\mu_n,\mu) < \frac{\epsilon}{2||f||_\infty}$$

Posons $E_a = \{f \geq a\}$. On a :

$$\int f d\mu_n = \int_0^{||f||_\infty} \mu_n(E_a)\, da \leq \frac{\epsilon}{2} + \int_0^{||f||_\infty} \mu\left(E_a^{\frac{\epsilon}{2||f||_\infty}}\right) da$$

On sait que $\mu\left(E_a^\delta \backslash E_a\right)$ tend vers 0 quand $\delta$ décroit vers 0. Donc, quitte à prendre une condition plus forte sur $d(\mu_n, \mu)$, on peut supposer que :

$$\mu\left(E_a^\delta \backslash E_a\right) < \frac{\epsilon}{2||f||_\infty}$$

Donc :

$$\int f d\mu_n \leq \epsilon + \int f d\mu$$

L'inégalité dans l'autre sens s'obtient de la même façon. □

# 3 Etude des matrices de Wigner centrées

## 3.1 Définition des matrices de Wigner

Nous allons tout d'abord définir une gamme particulière de matrices aléatoires, les matrices de Wigner.

**Définition 3.1** (Matrices de Wigner)**.** *Soient $N \in \mathbb{N}$ et $\beta = 1$ (respectivement $\beta = 2$ ou 4). On se donne $\left(g_{kl}^i\right)_{1 \leq k \leq l \leq \infty, i \leq \beta}$ des variables aléatoires gaussiennes centrées réduites et indépendantes.*
*On définit de plus, selon la valeur de $\beta$ :*

- ***Cas** $\beta = 1$ **(Ensemble Orthogonal Gaussien, GOE)** : $e_1^1 = 1$*
- ***Cas** $\beta = 2$ **(Ensemble Unitaire Gaussien, GUE)** : $e_2^1 = 1$ et $e_2^2 = i$.*
- ***Cas** $\beta = 4$ **(Ensemble Simplectique Gaussien, GSE)** :*

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

*Nous sommes prêts pour définir la matrice de Wigner $X^{N,\beta} = \left(X_{ij}^{N,\beta}\right)_{ij}$ comme une matrice auto-adjointe telle que :*

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

*et*

$$X_{kk}^{N,\beta} = \sqrt{\frac{2}{\beta N}} g_{kk}^1 e_\beta^1, 1 \leq k \leq N$$

## 3.2 Loi des matrices de Wigner et de leur distribution spectrale

Nous écrirons dans la suite $X$ au lieu de $X^{N,\beta}$ lorsqu'il n'y a pas de confusion possible.
Nous commencerons par montrer un résultat fondamental d'invariance.

**Théorème 3.1** (Invariance sous l'action du groupe orthogonal)**.** *Soit $U$ une matrice orthogonale. Alors $Z :=^t UXU$ a même loi que $X$.*

*Démonstration.* On peut écrire $X$ sous la forme $X = \frac{1}{2}\left({}^tY + Y\right)$, où $Y$ est une matrice aléatoire dont les entrées sont des gaussiennes centrées indépendantes de variance $\frac{2}{N}$.
Posons donc $A := {}^tUYU$. Alors $Z = \frac{1}{2}\left({}^tA + A\right)$ est clairement symétrique et nous pouvons donc nous intéresser à $A$.

- Les entrées de $A$ sont gaussiennes par combinaison linéaire de gaussiennes indépendantes.
- $$\mathbb{E}\left(a_{ij}\right) = \mathbb{E}\left(\sum_{kl} u_{ki} y_{kl} u_{lj}\right) = \sum_{kl} 0 = 0$$
  avec des notations évidentes.
- $$\mathbb{E}\left(a_{ij}^2\right) = 0 + \mathbb{E}\left(\sum_{kl} u_{ki}^2 y_{kl}^2 u_{lj}^2\right) = \sum_{kl} u_{ki}^2 \frac{2}{N} u_{lj}^2 = \frac{2}{N}$$
  en utilisant l'indépendance des $y_{kl}$ et le fait que $U$ appartient au groupe orthogonal.

On a donc bien l'invariance. □

**Théorème 3.2** (Densité de la loi des coefficients)**.** *La loi des coefficients de $X$ admet une densité par rapport à la mesure de Lebesgue, qui est :*

$$\prod_{j=1}^{N} \sqrt{\frac{N}{4\pi}} e^{-x_{jj}^2 N/4} \prod_{1 \leq j < k \leq N} \sqrt{\frac{N}{2\pi}} e^{-x_{jk}^2 N/2}$$

*Ce qui s'écrit encore, en rassemblant tous les $x_{ij}$ :*

$$\alpha \prod_{i,j} e^{-\frac{N}{4} Tr X^2}$$

*où $\alpha$ est un coefficient de normalisation.*

*Démonstration.* C'est clair ! □

On souhaite à présent trouver la loi des valeurs propres de la matrice de Wigner. Pour cela, on commence par remarquer que $X$ dépend de $N\left(N+1\right)$ variable indépendantes alors qu'il n'y a que $N$ valeurs propres (théorème spectral). Comme la matrice est entièrement déterminée par la donnée des valeurs propres et des vecteurs propres, cela signifie qu'il y a $N\left(N-1\right)/2$ coefficients indépendants des vecteurs propres, que l'on notera $p_j$. On peut donc réecrire la densité par changement de variable, en fonction d'une jacobienne $J$ :

$$e^{-\frac{1}{2}\sum \lambda_i^2} |J| \prod d\lambda_j \prod dp_j$$

Dans toute la suite, nous raisonnerons au signe près, car seule la valeure absolue de la jacobienne nous intéresse ! !
Introduisons des notations :

**Définition 3.2.** *Soit $A$ une matrice $N \times N$ :*

$$(A) := \Lambda_{1 \leq j \leq k \leq N} a_{jk}$$

*et*

$$dA := \begin{pmatrix} dx_{11} & \dots & dx_{1N} \\ \vdots & & \vdots \\ dx_{N1} & \dots & dx_{NN} \end{pmatrix}$$

Nous allons courtement rappeler quelques notions d'algébre multi-linéaire, en particulier par rapport au wedge :

### 3.2.1 Algébre multilinéaire

Soit $E$ un espace vectoriel, on note $\bigwedge^p E^*$ l'espace des $p$-formes linéaires alternées sur $E$, ie des formes p-linéaires qui sont nuls si au moins deux vecteurs d'entrée sont identiques. On note $\bigwedge^* E^* := \bigoplus_p \bigwedge^p E^*$. On définit alors ce qu'on appelle le produit extérieur :

**Définition 3.3.** *Soient $\omega$ et $\eta$ deux formes, respectivement $p$ et $q$, linéaires alternées. Alors, on définit :*

$$\omega \wedge \eta \left(x_1, \dots, x_{_{p+q}}\right) := \frac{1}{p!q!} \sum_{\sigma \in S_{p+q}} \epsilon_\sigma \omega \left(x_{\sigma(1)}, \dots, x_{\sigma(p)}\right) \eta \left(x_{\sigma(p+1)}, \dots, x_{\sigma(p+q)}\right)$$

On a alors les proprietés suivantes :

**Théorème 3.3.** *Le produit extérieur est associatif et admet 1 comme élément neutre (en identifiant $\mathbb{R}$ avec $\bigwedge^0 E^*$).*
*Il est également anticommutatif :*

$$\omega \wedge \eta = (-1)^{pq} \eta \wedge \omega$$

*Il munit l'espace $\bigwedge^* E^*$ d'une structure d'algèbre graduée, ie :*

$$\bigwedge^p E^* \wedge \bigwedge^q E^* \subset \bigwedge^{p+q} E^*$$

*Démonstration.* Le caractère gradué et l'élément neutre sont évidents.
Montrons l'associativité (avec des notations évidentes) :

$$\omega \wedge (\eta \wedge \theta)\,(x_1, \dots, x_{p+q+r}) = \frac{1}{p!\,(q+r)!} \sum_{\pi \in S_{p+q+r}} \epsilon_\pi \omega \left(x_{\pi(1)}, \dots, x_{\pi(p)}\right) (\eta \wedge \theta) \left(x_{\pi(p+1)}, \dots, x_{\pi(p+q+r)}\right)$$

$$= \frac{1}{p!\,(q+r)} \frac{1}{q!r!} \sum_{\pi \in S_{p+q+r}, \sigma \in \tilde{S}_{q+r}} \epsilon_\pi \epsilon_\sigma \omega \left(x_{\pi(1)}, \dots, x_{\pi(p)}\right) \eta \left(x_{\pi \circ \sigma(p+1)}, \dots, x_{\pi \circ \sigma(p+q)}\right)$$
$$\times \theta \left(x_{\pi \circ \sigma(p+q+1)}, \dots, x_{\pi \circ \sigma(p+q+r)}\right)$$

où $\tilde{S}_{q+r}$ s'identifie avec $S_{q+r}$ mais agissant sur $\{p+1,\dots,p+q+r\}$.
Or, $\pi\circ\sigma$ s'identifie avec une permutation de $S_{p+q+r}$ si on prolonge $\sigma$ par l'identité sur $\{1,\dots,p\}$. On peut donc faire le changement de variable $\hat{\pi} := \pi\circ\sigma$.
On remarque que pour un $\hat{\pi}$ fixé, il y a exactement $(q+r)$ couples $(\pi,\sigma)$ qui lui correspondent. il suffit en effet de se donner un $\sigma$ quelconque puis de prendre $\pi = \hat{\pi}\circ\sigma^{-1}$.
On obtient finalement :

$$\frac{1}{p!q!r!}\sum_{\hat{\pi}\in S_{p+q+r}} \epsilon_{\hat{\pi}}\omega\left(x_{\hat{\pi}(1)},\dots,x_{\hat{\pi}(p)}\right)\eta\left(x_{\hat{\pi}(p+1)},\dots,x_{\hat{\pi}(p+q)}\right)\theta\left(x_{\hat{\pi}(p+q+1)},\dots,x_{\hat{\pi}(p+q+r)}\right)$$

En partant de $(\omega\wedge\eta)\wedge\theta$, on voit facilement que l'on aboutit a la même formule. D'où l'associativité.
Pour l'anti-commutativité, posons $\tau$ la permutation de $S_{p+q}$ telle que $\tau(i) = i+p \mod (p+q)$. Alors sa signature est $(-1)^{pq}$ et :

$$\begin{aligned}(\omega\wedge\eta)(x_1,\dots,x_{p+q}) &= \frac{1}{p!q!}\sum_{\sigma\in S_{p+q}}\epsilon_\sigma\omega\left(x_{\sigma(1)},\dots,x_{\sigma(p)}\right)\eta\left(x_{\sigma(p+1)},\dots,x_{\sigma(p+q)}\right)\\ &= \frac{1}{p!q!}\sum_{\sigma\in S_{p+q}}\epsilon_{\sigma\circ\tau}\omega\left(x_{\sigma\circ\tau(1)},\dots,x_{\sigma\circ\tau(p)}\right)\eta\left(x_{\sigma\circ\tau(p+1)},\dots,x_{\sigma\circ\tau(p+q)}\right)\\ &= (-1)^{pq}(\eta\wedge\omega)(x_1,\dots,x_{p+q})\end{aligned}$$

□

Nous sommes dans des espaces à dimension finie et on peut donc identifier $E$ avec son bidual. De cette manière, on pose $\bigwedge^* E := \bigwedge^* E^{**}$, ie on prend l'algèbre engendrée par les formes multilinéaires sur le dual de $E$. On l'appelle algèbre des multivecteurs de $E$. C'est dans ce sens-là que nous avons pris le wedge dans la définition 3.2.

### 3.2.2 Suite de la preuve

Prouvons d'abord une proposition qui va nous être utile :

**Théorème 3.4.** *Soient $A$ et $M$ des matrices réelles $N\times N$ et $M$ symétrique.*

$$\left({}^tAdMA\right) = (detA)^{N+1}(dM)$$

*Démonstration.* On peut écrire $({}^tAdMA) = p(A)(dM)$ avec $p(A)$ un polynome en $A$ par multilinéarité du wedge.
Prenons alors $B$ une autre matrice, et on obtient :

$$\left({}^tB{}^tAdMAB\right) = p(B)(A)(dM)$$

D'où $p(AB) = p(A)p(B)$, ce que seul une puissance du déterminant peut satisfaire, cf. le lemme suivant.
En prenant $diag(a,1,\dots,1)$, où $a$ est réel, on peut conclure. □

**Lemme 3.1.** *Soit $p$ un polynôme en les coefficients d'une matrice $n \times n$ tel que $p(A)\,p(B) = p(AB)$. Alors, $p$ est une puissance entière du déterminant.*

*Démonstration.* Il suffit de montrer qu'un tel polynôme de degré minimale est le déterminant, car alors des divions euclidiennes de polynômes permettent de conclure.
On a :

$$p(I) = p(I)^2$$

et comme $p$ n'est pas le polynôme nul, $p(I) = 1$. Soit

$$U = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

$$W_k(l) = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & \dots & l & \dots & 0 \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

où le $l$ se trouve sur la diagonale en $k$ième position. Enfin :

$$V_j(k) = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 & 0 \\ \vdots & & & & & \vdots \\ \vdots & \dots & 1 & k & 0 & \dots \\ \vdots & \dots & 0 & 1 & 0 & \dots \\ \vdots & & & & & \vdots \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$$

où le $k$ se trouve à a la $j$ième ligne.
$U^2 = 1$ et donc $p(U) = \pm 1$. En outre, $V_j(k)\,V_j(-k) = I$ et donc $p(V_j(k))\,p(V_j(-k)) = 1$. Mais les polynômes sont de même degré, qu'il s'agisse de $k$ ou de $-k$. Ces deux polynômes sont donc des constantes. Or, $p(V_j(0)) = 1$ et donc $p(V_j(k)) = 1$ pour tout $j, k$. De plus, $UW_1(l)\,U = W_2$ et donc $p$ prend la même valeur en $W_1$ et en $W_2$. Un raisonnement analogue montre que $p$ prend la même valeur sur tous les $W_i(l)$ ($l$ étant fixé). On remarque que :

$$W_i(l)\,W_i(1/l) = I$$

$$p(W_i(l))\,p(W_i(1/l)) = 1$$

Or, $W(l)$ est un polynôme en $l$ que l'on peut noter $P$ par exemple. Soit $r$ sa valuation. Le coefficient de $X^{n-l}$ dans $P(X)P(1/X)$ vaut $a_n a_r$ et comme, par définition, $a_r$ n'est pas nul, on a $a_n = 0$. On peut continuer ainsi de proche en proche et on obtient que $P$ est de la forme $cX^i$. Mais comme $p(W_i(1)) = 1$, on a $c = 1$.
On remarque à présent que l'on peut remplacer $U$ par n'importe quelle matrice modélisant une transposition et le résultat serait le même. Or à partir de telles matrices, on forme toutes les matrices de permutations et, en rajoutant les matrices du type $V$ et $W$, on engendre finalement toutes les matrices. On connait donc la valeur de $p$ sur un ensemble générateur. le déterminant est un polynôme convenable de ce type avec $i = 1$. Si $p$ est un tel polynôme non trivial et de degré minimal, il doit être associé à $i = 1$ et c'est donc le déterminant. □

Nous pouvons enfin démontrer le résultat voulu :

**Théorème 3.5** (Densité de la loi des valeurs propres)**.** *La loi des valeurs propres, classés dans l'ordre croissant, de $X$ admet une densité par rapport à la mesure de Lebesgue, qui est donnée par :*

$$\alpha \exp\left(-\frac{N\beta}{4}\sum_j \lambda_j^2\right) \prod_{1\leq j<k\leq N} |\lambda_k - \lambda_j|^\beta$$

*où $\alpha$ est un coefficent de normalisation*

*Démonstration.* Montrons le dans le cas $\beta = 1$.
Soit $f$ une fonction bornée en les valeurs propres. Les valeurs propres elles-mêmes sont des fonction des coefficients :

$$\lambda_k = \lambda_k\left((x_{ij})_{ij}\right),\ \forall 1 \leq k \leq N$$

Il suffit de montrer que :

$$\alpha \int_{\mathbb{R}^N} f(\lambda_1, \ldots, \lambda_N) \prod_{i<j} |\lambda_i - \lambda_j|\, d\lambda_1 \lambda_N = \int_{\mathrm{GUE}} f(\lambda_1(x_{ij}), \ldots, \lambda_N(x_{ij}))\, dX^N$$

En d'autres termes, il s'agit juste de faire un changement de variable. Nous allons montrer que ce changement de variables s'opère avec le jacobien voulu.
D'après le théorème spectrale, on peut se donner une matrice diagonale $L$ et une matrice orthogonale $R$ telles que $X = RL\ {}^tR$. D'où :

$$dX = dRL\ {}^tR + RdL\ {}^tR + RL\ {}^tdR$$

D'où :

$${}^tRdXR = {}^tRdRL + L\ {}^tdRR + dL = {}^tRdRL - L\ {}^tRdR + dL$$

en tenant compte du fait que ${}^tRR = Id$, ce qui donne, en différentiant, ${}^tdRR = -\ {}^tRdR$.

Or le résultat du théorème précédant montre que $({}^tRdXR) = (X)$ au signe près, car $R$ est orthogonale.
De plus, en utilisant l'antisymétrie de ${}^tdRR$, on obtient :

$$ {}^tRdRL - L\ {}^tRdR + dL = \begin{pmatrix} d\lambda_1 & \dots & (\lambda_N - \lambda_1)\ {}^t\vec{r_1}.d\vec{r}_N \\ (\lambda_2 - \lambda_1)\ {}^t\vec{r_1}.d\vec{r}_2 & & \vdots \\ \vdots & & \vdots \\ (\lambda_N - \lambda_1)\ {}^t\vec{r_1}.d\vec{r}_N & \dots & d\lambda_N \end{pmatrix} $$

où $\vec{r_i}$ est le $i^{ime}$ vecteur colonne de $R$.
D'où :

$$ J\Lambda_i d\lambda_i \Lambda_j dp_j = (dX) = \prod_{1\leq j<k\leq N} (\lambda_k - \lambda_j)\, \Lambda_j d\lambda_j \left({}^tRdR\right) $$

La première égalité vient en exprimant $dx_{i,j}$ en fonction des $d\lambda_k$ et $dp_k$ et en utilisant que le wedge est alterné et linéaire, on voit apparaitre le déterminant de la matrice jacobienne.
D'où le résultat ! □

## 3.3 Convergence de la mesure spectrale

Avant d'étudier l'existence d'un principe de grandes déviations pour la mesure spectrale, montrons que celle-ci converge vers une loi que l'on appelle loi semi-circulaire. On se limite au cas $\beta = 2$.
On pose la mesure spectrale :

$$ \hat{\mu}^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i} $$

L'idée va être de calculer les moments de $\hat{\mu}^N$. On pose ainsi :

$$ m_k^N := \int x^k \hat{\mu}^N (dx) = \frac{1}{N} \sum \lambda_i^k = \frac{1}{N} Tr \left[ \left(X^N\right)^k \right] $$

Si on prend l'espérance de cette quantité, on arrive à la formule :

$$ \mathbb{E}\left(m_k^N\right) = \frac{1}{N} \sum_{i_1,\dots,i_k} \mathbb{E}\left[X^N_{i_1 i_2} \dots X^N_{i_k i_1}\right] \tag{1} $$

Nous allons à présent étudier cette quantité d'un peu plus près. On se rappelle que l'on a :

$$ \mathbb{E}\left[X_{ij}^N\right] = 0 \tag{2} $$

$$ \mathbb{E}\left[\left(X_{ij}^N\right)^2\right] = 0 \tag{3} $$

$$ \mathbb{E}\left[|X_{ij}^N|^2\right] = \frac{1}{N} \tag{4} $$

Par conséquent, s'il existe un terme dans la somme 1 tel qu'un certain $(i_j, i_{j+1})$ n'y apparait qu'une seule fois, alors l'espérance de ce terme est nulle.
On considère un polygone à $k$ cotés, indicés $i_1, \ldots, i_k$. Avec chaque terme de 1, on contruit une relation d'équivalence sur les côtés du polygone : On dit que les côtés $(i_l, i_{l+1})$ et $(i_p, i_{p+1})$ sont équivalents si et seulement si $i_l = i_{p+1}$ et $i_{l+1} = i_p$. Pour que le terme soit non nul, chaque classe d'équivalence doit

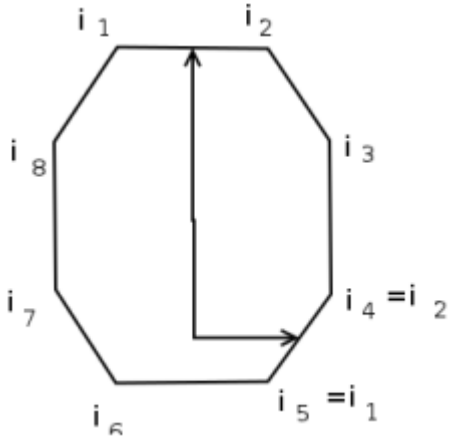


FIGURE 1 – Exemple d'appariement

contenir au moins deux côtés. Il y a donc au plus $k/2$ classes. En d'autres mots, il y a au plus $k/2$ indices distincts.
Supposons une relation d'équivalence fixée avec $l \leq k/2$ classes d'équivalences. Il y a alors

$$N\,(N-1)\ldots(N-l+1)$$

choix d'indices possibles. On peut majorer cette quantité par $N^l$.
Or,

$$\begin{aligned}\left|\mathbb{E}\left(X^N_{i_1 i_2}\ldots X^N_{i_k i_1}\right)\right| &\leq \mathbb{E}\left(\left|X^N_{i_1 i_2}\right|^k\right)^{1/k}\ldots\mathbb{E}\left(\left|X^N_{i_k i_1}\right|^k\right)^{1/k}\\ &\leq \mathbb{E}\left(\left|X^N_{12}\right|^k\right)\end{aligned}$$

où l'on a utilisé Hölder généralisé et le fait que tous les $X^N_{ij}$ ont même loi.
Donc, en sommant sur tous les termes donnant lieu à cette relation d'équivalence que l'on s'était fixée, on a (si $S$ est cette somme) :

$$S \leq \frac{\text{constante}}{N^{k/2}} N^l$$

où le $N^{k/2}$ provient de (4).
Cette quantité tend vers 0 quand $N$ tend vers l'infini dès que $l < k/2$. Ne contribuent donc asymptotiquement à la somme $m^N_k$ que les termes correspondant à

une partition des côtés du polygone en $k/2$ classes. On peut de plus supposer $k$ pair, sans quoi les termes s'évanouissent à l'infini.
Reprenons notre polygone et fusionnons les arêtes d'une même classe d'équivalence. On obtient alors un graphe à $k/2$ arêtes et $p$ sommets. Comme ce graphe est connexe, on a forcément $p \leq k/2 + 1$.
Or, le nombre de sommets obtenus après fusion des côtés correspond au nombre

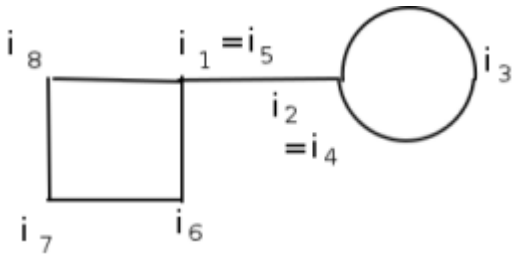


FIGURE 2 – Fusion des côtés équivalents du polygone de la figure 1

d'indices différents. Mais, si un terme a, disons, $r$ indices différents, alors ce terme intervient $N(N-1)\dots(N-r+1)$ fois dans la somme. Le terme dominant correspond au nombre maximum de sommets, c'est-à-dire $k/2+1$. Cela revient à dire que le graphe est un arbre.
Ce terme dominant de $m_{2k}^N$ va donc être :

$$\begin{aligned} & \frac{1}{N} \sum_{\text{polygones qui se réduisent en arbres}} \mathbb{E}\left(X_{i_1 i_2} \dots X_{i_k i_1}\right) \\ = \quad & \frac{1}{N} \sum_{arbres} \prod_{1}^{k} \mathbb{E}\left(X_{\alpha\beta} X_{\beta\alpha}\right) \\ = \quad & \frac{1}{N} \frac{1}{N^k} \times \sharp\{\text{nombre d'appariements non-croisés des côtés d'un polygone à } 2k \text{ côtés}\} \\ & \times N(N-1)\dots(N-k) \end{aligned}$$

Lorsque $N$ tend vers l'infini, cela tend exactement vers le nombre d'appariements non-croisés, que l'on note $c_k$.

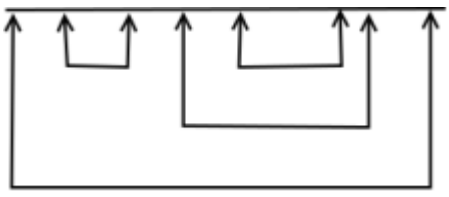

FIGURE 3 – Exemple d'appariements non croisés

**Lemme 3.2.** *Les $c_k$ sont les nombres dits de Catalan :*

$$c_k = \frac{1}{k+1}\binom{2k}{k}$$

*De plus, les $c_k$ sont les moments de la loi semi-circulaire, c'est-à-dire la loi de densité :*

$$\frac{1}{\pi}\sqrt{4-x^2}$$

*Démonstration.* Si l'on fait des appariements non-croisés sur $2k$ côtés, on peut d'abord choisir le côté que l'on décide d'associer au côté 1. Notons ce côté $l$. Il faut que $l$ soit pair, sans quoi on ne peut pas construire d'appariement non croisé avec les côtés restants. Avant l'indice $l$, il reste $l-2$ côtés à apparier et après $l$, il y en a encore $2k-l$.

$$c_k = \sum_{l=2, l \text{ pair}}^{2k} c_{l/2-1} c_{k-l/2}$$

Posons $\phi$ la fonction génératrice des $c_k$ :

$$\begin{aligned}
\phi(z) &= \sum_{k\geq 2} z^{2k} c_k \\
&= 1 + \sum_{k\geq 1} z^{2k} \sum_{l=1}^{k} c_{l-1} c_{k-l} \\
&= 1 + \sum_{k,l\geq 1, l\leq k} z^2 z^{2l-2} c_{l-1} z^{2k-2l} c_{k-l} \\
&= 1 + z^2 \phi(z)^2
\end{aligned}$$

En résolvant cette équation, on obtient :

$$\phi(z) = \frac{1 \pm \sqrt{1-4z^2}}{2z^2}$$

pour $-1/2 \leq z \leq 1/2$. On choisit le signe "moins" car il faut que $\phi(0) = 1$. Faisons un développement de Taylor :

$$\phi(z) = \frac{1}{2z^2} \sum_{k=1}^{\infty} \frac{(2k-2)!}{k!\,(k-1)!\,2^{2k-1}} 4^k z^{2k}$$

Par conséquent, le coefficient de $z^{2k}$ est :

$$\frac{(2k)!4^{k+1}}{2k!\,(k+1)!\,2^{2k+1}} = \frac{(2k)!}{k!k!\,(k+1)} = \frac{1}{k+1}\binom{2k}{k}$$

En outre, les moments de la loi semi-circulaire sont (les moments impairs, les pairs sont nuls pour des raisons évidentes de parité de la densité) :

$$\begin{aligned}\frac{1}{\pi}\int_{-2}^{2}\sqrt{4-x^2}x^{2k}dx &= \frac{1}{\pi}\int_0^{\pi}2\sin\theta\left(2\cos\theta\right)^{2k}2\sin\theta d\theta\\ &= \frac{4}{\pi}\int_0^{\pi}2^{2k}\sin^2\theta\cos^{2k}\theta d\theta\\ &= \frac{1}{\pi}\int_0^{\pi}\left[4\left(2\cos\theta\right)^{2k}-\left(2\cos\theta\right)^{2k+2}\right]d\theta\end{aligned}$$

Si on développe à préset les cosinus en $e^{i\theta}+e^{-i\theta}$ puis en développant les exponentielles en séries de puissances, on trouve les nombres de Catalan. □

Ce que l'on vient de montrer est que pour tout $k$ :

$$\mathbb{E}\left(\int x^k\hat{\mu}^N\left(dx\right)\right)\to\frac{1}{\pi}\int_0^{\pi}x^k\sqrt{4-x^2}dx$$

Regardons la variance :

$$\mathbb{E}\left[\left(\frac{1}{N}Tr\left(X_N^k\right)\right)^2\right]=\frac{1}{N^2}\mathbb{E}\left[\sum M_{i_1i_2}\ldots M_{i_ki_1}M_{j_1j_2}\ldots M_{j_kj_1}\right]$$

On peut alors appliquer une méthode analogue à ce que l'on a fait précédemment, avec les polygones, pour montrer que cette variance tend vers 0.
On a donc une convergence en probabilité de $m_k$ vers le $k$ième moment de la loi semi-circulaire.

## 3.4 Principe de grandes déviations pour les matrices de Wigner

On cherche à démontrer le résultat de grandes déviations suivant :

**Théorème 3.6** (Principe de Grande déviation)**.** *On pose $\mathcal{P}\left(\mathbb{R}\right)$ l'espace des mesures de probabilité sur $\mathbb{R}$ et on se donne la fonction suivante sur cet espace : $I_\beta\left(\mu\right):=\frac{\beta}{4}\int x^2d\mu\left(x\right)-\frac{\beta}{2}\Sigma\left(\mu\right)-\frac{3}{8}\beta$ avec $\Sigma\left(\mu\right):=\int\int\log|x-y|d\mu\left(x\right)d\mu\left(y\right)$. Alors,*

- *$I_\beta$ est bien définie, à valeurs dans $[0,\infty]$*
- *$I_\beta$ est une bonne fonction de taux*
- *$\hat{\mu}$ satisfait un principe de grandes déviations à vitesse $N^2$ et de bonne fonction de taux $I_\beta$*

Une méthode à laquelle on pourrait penser serait la méthode de Laplace, que nous rappelons rapidement ici, sans démonstration :

**Théorème 3.7.** *Soit $\left(\mu_n\right)_n$ une suite de mesures de probabilités satisfaisant un principe de grande déviation de vitesse $a_n$ et de bonne fonction de taux $I$. Soit $F$ une fonction à valeure réelle continue et majorée, alors posons :*

$$\nu_n\left(dx\right):=\alpha\exp\left(a_nF\left(x\right)\right)\mu_n\left(dx\right)$$

*avec $\alpha$ un coefficient de normalisation.Posons aussi :*

$$J(x) := I(x) - F(x) - \sup_{y \in X} \{F(y) - I(y)\}$$

*Alors $(\nu_n)$ satisfait un principe de grande déviation à vitesse $a_n$ et de bonne fonction de taux $J$.*

Si nous cherchions à appliquer cette méthode, nous devrions d'abord trouver une fonction $F$ convenable. On peut en effet, réecrire cette densité sous la forme :

**Théorème 3.8.** *Soit $f(x,y) := -\log|x-y| + \frac{1}{4}\left(x^2+y^2\right)$. Alors, la densité des valeurs propres s'écrit aussi, à un coefficent multiplicatif de normalisation près :*

$$\exp\left(-\frac{\beta}{2}N^2 \int_{x \neq y} f(x,y)\, d\hat{\mu}^N(x)\, d\hat{\mu}^N(y)\right) \prod_{i=1}^{N} e^{-\frac{\beta}{4}\lambda_i^2}$$

*Démonstration.* Montrons le dans le cas $\beta = 1$.On a d'une part :

$$\int_{x \neq y} \log|x-y|\, d\hat{\mu}(x)\, d\hat{\mu}(y) = \frac{1}{N^2} \sum_{i \neq j} \log|\lambda_i - \lambda_j| = \frac{2}{N^2} \sum_{i<j} \log|\lambda_j - \lambda_i|$$

En remplaçant dans l'expression, on obtient donc la partie déterminant de Van Der Monde de la densité spectrale. De plus :

$$\int_{x \neq y} \frac{1}{4}\left(x^2+y^2\right) d\hat{\mu}(x)\, d\hat{\mu}(y) = \frac{N^2}{4} \sum_{i \neq j} \left(\lambda_j^2 + \lambda_i^2\right) = 2\frac{N-1}{4} N^2 \sum_i \lambda_i^2$$

ce qui permet d'obtenir la partie "exponentielle de la trace". □

Le problème, cependant, est que le logarithme explose lorsque $x$ est proche de $y$ et on ne peut donc pas appliquer la méthode de Laplace de cette manière (car $F$ doit être majoré).
Pour alléger les notations, nous définirons dans la suite

$$J_\beta(\mu) := \frac{\beta}{2} \int \int_{x \neq y} f(x,y)\, d\mu(x)\, d\mu(y)$$

ainsi que $Z_\beta^N$ la constante de normalisation et $Q_\beta^N$ la densité non normalisée.
On peut se restreindre à l'étude des grandes déviations de $Q$, comme le montre le résultat suivant :

**Théorème 3.9.** *Suppsosons avoir prouvé que la mesure spectrale satisfait sous $Q^N$ un principe de grande déviation à vitesse $a_N$ et avec bonne fonction de taux $J$. Alors :*

$$\lim_{N \to \infty} \frac{1}{N^2} \log Z_\beta^N = -\frac{\beta}{2} \inf_{\nu \in P(\mathbb{R})} \int \int_{x \neq y} f(x,y)\, d\nu(x)\, d\nu(y) = \frac{3}{8}\beta$$

*Cela permet donc de conclure quant au principe de grande déviation pour la mesure spectrale (normalisée).*

*Démonstration.* La premiére inégalité est une conséquence directe du principe de grandes déviations pour $Q$.
La seconde inégalité est admise. □

**Théorème 3.10.** *$I_\beta$ est une bonne fonction de taux.*

*Démonstration.* On remarque que si $g$ est une fonction continue bornée, alors, par le théorème de convergence dominée, $\int g(x,y)\,d\mu(x)\,d\mu(y)$ est continue en $\mu$ (pour la topologie de la convergence en loi). Or, $f(x,y)$ peut être vu comme la limite croissante, donc le sup, de $f(x,y)\,1_{|x|\leq n,|y|\leq n,|x-y|\geq\frac{1}{n}}$. Cela signifie que $I_\beta$ est le supremum de fonctions continues en $\mu$ et est donc semi-continue inférieurement.
De plus, on a l'inégalité suivante (en prenant $m=-\inf f$) pour tout $A$ positif :

$$\begin{aligned}\frac{2}{\beta}I_\beta(\mu)+\frac{3}{8}+m &= \int (f+m)\,d\mu^{\otimes 2}\\ &\geq \left[\mu\left({}^c[-A,A]\right)\inf_{{}^c[-A,A]}(f+m)\right]^2\end{aligned}$$

D'où

$$\{I_\beta\leq M\}\subset K_{2M}$$

où

$$K_M:=\bigcap_{A>0}\left\{\mu\in P(\mathbb{R}),\mu([-A,A])\leq\sqrt{\frac{\frac{2}{\beta}M+m+\frac{3}{8}}{\inf{}_{{}^c[-A,A]}(f+m)}}\right\}$$

On note par ailleurs que les $K_M$ sont compacte car ils sont tendus par définition. Les sous-niveau de $I_\beta$ sont donc compactes : c'est une bonne fonction de taux. □

Nous allons maintenant prouver la borne supérieure.
Remarquons tout d'abord que, si on note $\Delta=\{(x,x),x\in\mathbb{R}\}$, on a :

$$\hat{\mu}^N\otimes\hat{\mu}^N(\Delta)=\int_{x=y}d\hat{\mu}^N(x)\,d\hat{\mu}^N(y)=\frac{1}{N^2}\sum_{i=j}1=\frac{1}{N}$$

Tout cela ayant lieu $Q$-presque sûrement et en tenant compte du fait que, presque sûrement, $\lambda_i\neq\lambda_j$.
On se donne à présent $M$ réel positif et on écrit, $Q$-presque sûrement :

$$\begin{aligned}\int\int_{x\neq y}f(x,y)\,d\hat{\mu}^N(x)\,d\hat{\mu}^N(y)\geq &\int\int_{x\neq y}f(x,y)\wedge M\,d\hat{\mu}^N(x)\,d\hat{\mu}^N(y)\\ =&\int\int f(x,y)\wedge M\,d\hat{\mu}^N(x)\,d\hat{\mu}^N(y)-\frac{M}{N}\end{aligned}$$

Donc, en utilisant le théorème 3.8, on obtient, pour tout $A$ dans $P(\mathbb{R})$ :

$$\begin{aligned} Q\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}\in A\right) &= \int_{\mathbb{R}^N}\exp\left(-\frac{\beta}{N}\int_{x\neq y}f(x,y)\,d\hat{\mu}(x)\,d\hat{\mu}(y)\right)\prod_i\exp\left(-\frac{\beta}{4}\lambda_i^2\right)1_{\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}\in A}d\lambda_i \\ &\leq \frac{1}{\sqrt{\beta\pi}^N}\exp\left(-\frac{\beta N^2}{2}\inf_{\mu\in A}\left\{\int f(x,y)\wedge M d\mu(x)\right)d\mu(y)\right\}+NM \end{aligned} \tag{6}$$

D'où

$$\limsup_{N\to\infty}\frac{1}{N^2}\log\left(Q\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}\in A\right)\right)\leq-\frac{\beta}{2}\inf_{\mu\in A}\left\{\int f(x,y)\wedge M d\mu(x)\,d\mu(y)\right\}$$

Il suffit juste de s'affranchir du $M$.

### 3.4.1 Pour s'affranchir de M

Restreignons-nous pour un moment à des $A$ de la forme $B(\mu,\delta)$, ie des boules ouverte. Par semi-continuité inférieure, on a :

$$\limsup_{\delta\to0}\limsup_{N\to\infty}\frac{1}{N^2}\log Q\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}\in B(\mu,\delta)\right)\leq-\int\int f(x,y)\wedge M d\mu(x)\,d\mu(y)$$

La convergence monotone, pour $M\to\infty$ permet d'obtenir (quitte à séparer les intégrales en partie positive et partie négative) :

$$\limsup_{\delta\to0}\limsup_{N\to\infty}\frac{1}{N^2}\log Q\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}\in B(\mu,\delta)\right)\leq-\int\int f(x,y)\,d\mu(x)\,d\mu(y)$$

Il faut cependant encore passer à des $A$ plus généraux. On suppose déjà dans un premier temps que $A$ est compacte. On peut donc le recouvrir par un nombre fini de boules de rayon $\delta$ : $B(\mu_1,\delta),\ldots,B(\mu_k,\delta)$. D'où :

$$\limsup_{N\to\infty}\frac{1}{N^2}\log Q\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i}\in A\right)\leq-\inf_{i=1\ldots k}\int\int f(x,y)\,d\mu_i(x)\,d\mu_i(y)$$

En prenant des $\delta$ tendant vers 0, on a ce qu'on veut dans le cas $A$ compact.

### 3.4.2 Le principe du minimax pour aller plus vite

En fait, ce qui se passe lorsqu'on veut faire tendre $M$ vers l'infini revient souvent dans les principes de grandes déviations et on aurait pu directement appliquer le résultat suivant :

**Théorème 3.11.** *Soit $E$ un espace polonais et $P_n$ une suite de mesures de probabilités. sur $E$. Soit $(J_M)_{M\in\mathbb{U}}$ une famille de fonctions semi-continues supérieurement. Supposons que l'on ait pour tout ouvert $O$ :*

$$\limsup_{N\to\infty} \frac{1}{a_N} \log P_N(O) \leq \inf_{M\in\mathbb{U}} \sup_{x\in O} J_M(x)$$

*Alors, on a pour tout compact $K$ :*

$$\limsup_{N\to\infty} \frac{1}{a_N} \log P_N(K) \leq \sup_{x\in K} \inf_{M\in\mathbb{U}} J_M(x)$$

*Ce qui permet souvent de conclure pour la borne supérieure d'un principe de grandes déviations (en prenant pour $J_M$ les opposés de fonctions de taux).*

Pour démontrer ce résultat, il faut passer par un lemme :

**Lemme 3.3.** *Soit $K$ un compact de $E$. Avec les mêmes notations qu'avant, on a :*

$$\inf_{O_1,\ldots,O_n} \max_{1\leq j\leq n} \inf_{M\in\mathbb{U}} \sup_{y\in O_j} J_M(y) \leq \sup_{x\in K} \inf_{M\in\mathbb{U}} J_M(x)$$

*avec les $O_1,\ldots,O_n$ parcourant tous les recouvrements de $K$.*

*Démonstration.* Soit $R$ le membre de droite de l'inégalité désirée. Soit $\epsilon$ strictement positif.
On a pour tout $x$ $\inf_M J_M(x) \leq R$ et par conséquent, on peut se donner $M(x)$ tel que

$$J_{M(x)}(x) \leq R + \epsilon$$

Par semi-continuité, on a pour tout $y$ dans un voisinage ouvert $O_x$ de $x$ :

$$\sup_{y\in O_x} J_{M(x)}(y) \leq R + 2\epsilon$$

Par compacité, on peut recouvrir $K$ par un nombre fini $O_{x_1},\ldots,O_{x_n}$ de ces ouverts. Le résultat précédent permet alors simplement d'obtenir le membre de gauche de l'inégalité désirée. □

*Démonstration du théorème.* On se donne $O_1,\ldots,O_n$ un recouvrement ouvert de $K$. On a que :

$$\limsup_{N\to\infty} \frac{1}{a_N} \log P_N(K) \leq \max_i \limsup_{N\to\infty} \frac{1}{a_N} \log P_N(O_i) \leq \max_i \inf_{M\in\mathbb{U}} \sup_{x\in O_i} J_M(x)$$

En prenant l'infimum sur les recouvrements ouverts de $K$ ; on peut appliquer le lemme, ce qui donne immédiatement le résultat. □

### 3.4.3 Tension exponentielle et fin de la borne sup

Si on peut montrer la tension exponentielle de $Q$, alors on pourra passer d'un principe de grandes déviations faible à un principe de grandes déviations. C'est le cas, car :

**Théorème 3.12** (Tension exponentielle)**.** *La suite des mesures $Q_\beta^N$ est exponentiellement tendue.*

*Démonstration.* On remarque que :

$$\limsup_{N\to\infty} \frac{1}{N^2} \log Q^N \left( \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i} \in A \right) \leq -\frac{\beta}{2} \inf_{\mu\in A} \left\{ \int f \wedge M d\mu^{\otimes 2} \right\}$$

On reprends $K_M$ comme dans la preuve du caractère "bonne fonction de taux" de $I_\beta$. Alors, en reprenant cette preuve, on voit que :

$$K_{2M}^c \subset \{I_\beta > M\}$$

et donc :

$$\limsup_{N\to\infty} \frac{1}{N^2} \log Q^N \left( \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i} \in A \right) \leq -M\beta/2$$

D'où la tension exponentielle. □

### 3.4.4 Preuve de la borne inf

Nous allons à présent nous atteler à la démonstration de la borne inférieure. Nous allons pour cela utiliser une approximation de mesure :

**Théorème 3.13.** *Soit $\nu$ une mesure de probabilité sur $\mathbb{R}$. On peut supposer que $\nu$ n'a pas d'atomes, car sinon $I_\beta = \infty$ et la borne inférieure est évidente...On pose :*

$$x^{1,N} := \inf \left\{ x \text{ tel que } \nu\left(]-\infty, x]\right) \geq \frac{1}{N+1} \right\}$$

*et pour $1 \leq i \leq N-1$,*

$$x^{i+1,N} := \inf \left\{ x^{i,N} \text{ tel que } \nu\left(]x^{i,N}, x]\right) \geq \frac{1}{N+1} \right\}$$

*Et on pose $\nu^N := \frac{1}{N} \sum_{i=1}^N \delta_{x^{i,N}}$.*
*Alors, $\nu^N$ converge vers $\nu$ (pour la topologie de la convergence en loi, comme toujours).*

Remaque : On a pris $\frac{1}{N+1}$ dans la définition au lieu de $\frac{1}{N}$ pour que $x^{N,N}$ soit toujours bien défini.

*Démonstration.* Soit $\phi$ une fonction continue bornée.

$$\int \phi d\nu^N = \frac{1}{N}\sum_i \phi\left(x^{i,N}\right) = \int \phi^N d\nu \to \int \phi d\nu$$

avec $\phi^N(x) := \phi\left(x^{i,N}\right)$ pour $x \in [x^{i,N}, x^{i+1,N}[$, et convergence dominée. □

On pose $\Delta_N := \{\lambda_1 \leq \lambda_2 \leq \ldots \lambda_N\}$. Alors :

$$Q_\beta^N\left(\hat{\mu}^N \in B(\nu,\delta)\right) \geq Z_\beta^N \mathbb{P}\left(\forall i \left|\lambda_i - x^{i,N}\right| < \frac{\delta}{2} \text{ et } \lambda_1 \leq \lambda_2 \leq \ldots \lambda_N\right)$$

En utilisant l'invariance par translation de la mesure de Lebesque et quitte à permuter l'ordre des $\lambda_i$, le membre de droite vaut :

$$\int_{\Delta_N \cap \left\{|\lambda_i| < \frac{\delta}{2}\right\}} \prod_{i<j} \left|x^{i,N} - x^{j,N} + \lambda_i - \lambda_j\right|^\beta \exp\left(-\frac{N}{2}\sum_{i=1}^N \left(x^{i,N} + \lambda_i\right)^2\right) \prod_{i=1}^N d\lambda_i$$

Or, nous avons le grand avantage de connaitre l'ordre des $\lambda_i$ et des $x^{i,N}$. Donc :

$$\left|x^{i,N} - x^{j,N} + \lambda_i - \lambda_j\right| > \max\left\{\left|x^{j,N} - x^{i,N}\right|, |\lambda_j - \lambda_i|\right\}$$

D'où l'on peut écrire :

$$\begin{aligned} Q_\beta^N\left(\hat{\mu}^N \in B(\nu,\delta)\right) \quad \geq \quad & \prod_{i+1<j} \left|x^{i,N} - x^{j,N}\right|^\beta \\ \times \quad & \prod_{i=1}^N \left|x^{i+1,N} - x^{i,N}\right|^{\beta/2} \exp\left(-\frac{N}{2}\sum_{i=1}^N \left(\left|x^{i,N}\right| + \delta/2\right)^2\right) \\ \times \quad & \int_{\Delta_N \cap [-\frac{\delta}{2},\frac{\delta}{2}]^N} \prod_{i=1}^N |\lambda_{i+1} - \lambda_i|^{\beta/2} \prod_{i=1}^N d\lambda_i \end{aligned}$$

**Théorème 3.14.** *On a la minoration suivante*

$$\int_{\Delta_N \cap [-\frac{\delta}{2},\frac{\delta}{2}]^N} \prod_{i=1}^{N-1} |\lambda_{i+1} - \lambda_i|^{\beta/2} \prod_{i=1}^N d\lambda_i \geq \left(\frac{1}{\beta/2+1}\right)^{N-1} \left(\frac{\delta}{2N}\right)^{(\beta/2+1)(N-1)+1}$$

*Démonstration.*

$$I := \int_{\Delta_N \cap [-\frac{\delta}{2},\frac{\delta}{2}]^N} \prod_{i=1}^{N-1} |\lambda_{i+1} - \lambda_i|^{\beta/2} \prod_{i=1}^N d\lambda_i \geq \int_{[-\delta/2, -\delta/2+\delta/N] \times [0,\delta/N]^{N-1}} \prod_{i=2}^N u_i^{\beta/2} \prod_{i=1}^N du_i$$

Obtenu en faisant le changement de variables $u_1 := \lambda_1$ et pour $i = 2 \ldots N$, $u_i := \lambda_i - \lambda_{i-1}$ et en restreignant les $u_i$ à l'intervalle $[0, \delta/N]$ car $\lambda_k = \sum_{i=1}^k u_i$. Bref :

$$I \geq \frac{\delta}{N} \prod_{i=2}^N \frac{1}{\beta/2+1} \left(\frac{\delta}{N}\right)^{\beta/2+1}$$

D'où le résultat voulu.

□

**Théorème 3.15.** *On a :*

$$\frac{1}{2(N+1)}\sum_{i=1}^{N}\left(x^{i,N}\right)^2-\frac{\beta}{2(N+1)^2}\sum_{i\neq j}\log\left|x^{i,N}-x^{j,N}\right|$$

$$\leq\frac{1}{2}\int x^2d\nu(x)-\beta\int_{x^{1,N}\leq x<y\leq x^{N,N}}\log(y-x)\,d\nu(x)\,d\nu(y)$$

*Démonstration.* On pose $I_i:=[x^{i,N},x^{i+1,N}]$ et alors, en utilisant la croissance de la fonction logarithme :

$$\int_{x^{1,N}\leq x<y\leq x^{N,N}}\log(y-x)\,d\nu(x)\,d\nu(y)\leq\sum_{1\leq i\leq j\leq N-1}\log\left(x^{j+1,N}-x^{i,N}\right)\nu^{\otimes 2}\left((x,y)\in I_i\times I_j\text{ et }x<j\right)$$

$$=\frac{1}{(N+1)^2}\sum_{i<j}\log\left|x^{i,N}-x^{j+1,N}\right|+\frac{1}{2(N+1)^2}\sum_{i}\log\left|x^{i+1,N}-x^{i,N}\right|$$

Le terme de droite venant par symétrie de $x<y$ et $y<x$ et parce que $\nu$ n'a pas d'atomes (par hypothèse).
On peut appliquer le même type de raisonnement à l'autre intégrale : $\int x^2d\nu(x)$ et on obtient le résultat désiré. □

Donc, en combinant 3.14 et 3.15, on peut écrire (le premier terme vient de la somme de Riemann) :

$$\liminf_{N\to\infty}\frac{1}{N^2}\log Q\left(\hat{\mu}^N\in B(\nu,\delta)\right)$$

$$\geq-\delta\int|x|\,d\nu(x)-\delta^2+\liminf_{N\to\infty}\left\{\beta\int_{x^{1,N}\leq x<y\leq x^{N,N}}\log(x-y)\,d\nu(x)\,d\nu(y)\right\}-\frac{1}{2}\int x^2d\nu(x)$$

On conclut en remarquant que la limite inférieure correspond à $\int\int fd\nu^{\otimes 2}$ (car $x^{1,N}$ tend vers 0 quand $N$ tend vers $\infty$ et de même $x^{N,N}$ tend vers $\infty$) et en faisant tendre $\delta$ vers 0.
En effet, prouver la borne inférieure d'un principe de grande déviation revient à la prouver pour des boules ouvertes arbitrairement petites.

# 4 Les intégrales sphériques

On s'intéresse à présent à une classe particuliére d'int{egrales, qui vont avoir un certain nombre de rapports avec les matrices al{eatoires.

## 4.1 Présentation et motivation

**Définition 4.1** (Intégrales sphériques)**.** *Soit $m_N$ la mesure de Haar sur le groupe unitaire (si $\beta=2$, pour les autres valeurs de $\beta$, $m_N$ est la mesure*

*de Haar sur le groupe orthogonal, respectivement symplectique). Soit $D_N$ une matrice symétrique, $E_N$ une matrice autoadjointe. Alors, on définit :*

$$I_N^{(\beta)}(D_N, E_N) := \int \exp\left(\frac{N\beta}{2} Tr\left(U D_N U^* E_N\right)\right) dm_N(U)$$

L'intérêt que présentent ces intégrales en rapport avec les matrices aléatoires apparait lorsque nous considérons par exemple des matrices de Wigner non centrées $Y_N := D_N + X_N$, où $X_N$ est une matrice de Wigner centrée et $D_N$ est diagonale déterministe. La densité des coefficients de $Y_N$ est du type :

$$\exp\left(-\frac{N}{4} Tr\left(Y_N^2\right)\right) = \exp\left(-\frac{N}{4} Tr\left(X_N^2\right)\right) \exp\left(-\frac{N}{4} Tr D_N^2\right) \exp\left(-\frac{N}{2} Tr\left(D_N X_N\right)\right)$$

On voit donc, du fait du dernier facteur, apparaitre quelque chose qui ressemble aux intégrales sphériques car on peut décomposer $X$ sous la forme $X_N = UVU^*$ avec V diagonale. Dans toute la suite, nous ferons également les hypothèses suivantes : [H]

1. Il existe un nombre réel positif $d_m$ tel que $D_N$ ait toute ses valeurs propres au dedans d'une boule de centre 0 et de rayon $d_m$.
2. $\hat{\mu}_{D_N}^N$ converge étroitement vers une probabilité $\mu_D$.
3. $\hat{\mu}_{E_N}^N$ converge étroitement vers une probabilité $\mu_E$
4. $\int x^2 d\hat{\mu}_{E_n}^N(x)$ est uniformément borné.

## 4.2 Théorème de base

Le reste de cette partie consacrée aux intégrales sphérique nous permettra de montrer le théorème suivant, ou du moins d'en esquisser une idée de preuve. On pose d'abord :

$$\bar{I}^{(\beta)}(\mu_D, \mu_E) := \limsup_{N\to\infty} \frac{1}{N^2} \log I_N^{(\beta)}(D_N, E_N)$$

et

$$\underline{I}^{(\beta)}(\mu_D, \mu_E) := \liminf_{N\to\infty} \frac{1}{N^2} \log I_N^{(\beta)}(D_N, E_N)$$

N.B. Le point 1 du théorème justifiera la notation que l'on vient d'introduire, en particulier le fait que ces quantités sont indépendantes des $D_N$ ou $E_N$ choisis, tant que la mesure spectrale converge étroitement vers $\mu_D$, $\mu_E$.

**Théorème 4.1.** *1. Il existe une fonction $g : \mathbb{R}_+^* \times \mathbb{R}_+ \to \mathbb{R}_+^*$, qui ne dépend que de $\mu_E$, tendant vers 0 quand $\delta$ tend vers 0 à $L$ fixé, telle que, si $\hat{E}_N$ et $\bar{E}_N$ sont deux matrices vérifiants :*

$$d\left(\hat{\mu}_{\bar{E}_N}^N, \mu_E\right) + d\left(\hat{\mu}_{\hat{E}_N}^N, \mu_E\right) \leq \delta/2$$

*et*

$$\int x^2 d\hat{\mu}^N_{\hat{E}_N}(x) + \int x^2 d\hat{\mu}^N_{\bar{E}_N}(x) \leq L$$

*Alors :*

$$\limsup_{N\to\infty} |\frac{1}{N^2} \log \frac{I_N^{(\beta)}\left(D_N, \hat{E}_N\right)}{I_N^{(\beta)}\left(D_N, \bar{E}_N\right)}| \leq g(\delta, L)$$

2. *Pour toute mesure $\mu$ de probabilité, on a :*

$$\inf_{\delta 0} \liminf_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}\left(d\left(\hat{\mu}^N_Y, \mu\right) < \delta\right) = \inf_{\delta\to 0} \limsup_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}\left(d\left(\hat{\mu}^N_Y, \mu\right) < \delta\right)$$

*avec $Y_N$ défini comme avant, c'est-à-dire comme la somme d'une matrice de Wigner centrée et d'une matrice diagonale d'eterministe. On appelle alors cette quantité $-J_\beta(\mu_D, \mu)$.*

3. *Avec $I_\beta(\mu)$ défini comme dans la partie sur les matrices de Wigner, on a :*

$$\begin{aligned} \bar{I}^{(\beta)}(\mu_D, \mu_E) &= \underline{I}^{(\beta)}(\mu_D, \mu_E) \\ &= -J_\beta(\mu_D, \mu_E) + I_\beta(\mu_E) - \inf_{\mu\in\mathcal{P}(\mathbb{R})} I_\beta(\mu) + \frac{\beta}{4}\int x^2 d\mu_D(x) \end{aligned}$$

*Démonstration.* Nous allons démontrer ce théorème dans un cadre un peu plus restrictif, à savoir que nous faisons l'hypothèse que $\bar{E}_N$ et $\hat{E}_N$ sont uniformément bornés par un certain $M$ (au sens de la norme sup sur les matrices, par exemple). La preuve se fait par étape :

1. Nous allons découper l'intervalle $[-M, M]$ en morceaux suffisament petits ; plus précisément, si $\delta'$ est un réel strictement positif, on partitionne l'intervalle en $(A_i)_{i\in\mathcal{I}}$ dont la longueur appartient à $[\delta', 2\delta']$ et dont les extrémités sont des points de continuité de $\mu_E$.
   On peut alors poser :

$$\bar{I}_k := \left\{i : \bar{E}_N(i) \in A_k\right\}$$

$$\hat{I}_k := \left\{i : \hat{E}_N(i) \in A_k\right\}$$

   Avec $M(i)$ qui dénote la $i$ème valeur propre de la matrice $M$.
   Comme on a en outre, par hypothèse, convergence étroite de $\hat{\mu}^N_{\hat{E}_N}$ (respectivement $\hat{\mu}^N_{\bar{E}_N}$) vers $\mu_E$ et que la frontière des $A_i$ est de masse nulle sous $\mu_E$, on a pour $N$ assez grand :

$$|\mu_E(A_j) - |\hat{I}_j|/N| + |\mu_E(A_j) - |\bar{I}_j|/N| \leq \delta \tag{7}$$

2. On souhaite contruire une permutation $\sigma$ telle que l'on ait :

$$|\hat{E}(\sigma(i)) - \bar{E}(i)| \leq 2\delta$$

pour "beaucoup" d'indices.
On construit cette permutation comme suit :
Soit un indice $k$. Si $|\bar{I}_k| \leq |\hat{I}_k|$, alors on pose $\tilde{I}_k := \bar{I}_k$ et on prend $\sigma$ qui envoie $\tilde{I}_k$ sur $\hat{I}_k$.
Si, par contre, $|\bar{I}_k| > |\hat{I}_k|$, alors, on se donne $\tilde{I}_k$ une partie de $\bar{I}_k$ mais de taille $|\hat{I}_k|$. Et on définit alors la permutation $\sigma_N$ pour qu'elle envoie, comme avant, $\tilde{I}_k$ sur $\hat{I}_k$.
On a alors, si $\mathcal{I}_0$ est l'ensemble des indices vérifiants la proprieté voulue :

$$\begin{aligned}
|\mathcal{I}_0| &\geq |\cup_k \sigma_N\left(\tilde{I}_k\right)| \\
&= \sum_k |\sigma_N\left(\tilde{I}_k\right)| \\
&= N - \sum_k |\bar{I}_k \backslash \tilde{I}_k| \\
&\geq N - \max_k \left(|\bar{I}_k| - |\tilde{I}_k|\right) |\mathcal{I}| \\
&\geq N - \delta N \frac{M}{\delta'}
\end{aligned}$$

où la dernière ligne vient de la relation 7.

3. On remarque à présent que $I_N^{(\beta)}\left(D_N, \bar{E}_N\right)$ est invariant par permutation des éléments de $D_N$. En effet, soit $\sigma$ une permutation de $\{1, \ldots, N\}$ :

$$\begin{aligned}
I_N^{(\beta)}\left(D_N, \bar{E}_N\right) &= \int \exp\left[\frac{N\beta}{2} Tr\left(U D_N U^* \bar{E}_N\right)\right] dm_N(U) \\
&= \int \exp\left[\frac{N\beta}{2} \sum_{i,j,k} u_{ij} D_N(j) \bar{u}_{kj} \bar{E}_N(k,i)\right] dm_N(U) \\
&= \int \exp\left[\frac{N\beta}{2} \sum_{i,j,k} u_{i\sigma(j)} D_N(\sigma(j)) \bar{u}_{k\sigma(j)} \bar{E}_N(k,i)\right] dm_N(U) \\
&= \int \exp\left[\frac{N\beta}{2} \sum_{i,j,k} u_{ij} D_N(\sigma(j)) \bar{u}_{kj} \bar{E}_N(k,i)\right] dm_N(U)
\end{aligned}$$

où la dernière égalité provient du fait que l'on a juste permuté les colonnes de $U$ et comme $m$ est une mesure de Haar, elle est invariante par multiplication à gauche d'une matrice unitaire, a fortiori donc par une matrice de permutation !
On remarquera aussi au passage que, quitte à décomposer $\bar{E}_N = U \bar{E}'_N U^*$ et comme la mesure de Haar est invariante par translation à gauche, on peut supposer que $\bar{E}_N$ est diagonale et alors la somme dans l'exponentielle s'écrit plus simplement sous la forme :

$$\sum_{i,k} |u_{ik}|^2 D_N(k) \bar{E}_N(i)$$

4. Fin de la preuve du point 1 du théorème.
On peut donc écrire, en prenant la permutation $\sigma$ qui a été construite au point 2 :

$$\begin{aligned}
& \sum_{i,k} |u_{i,k}|^2 D_N(k) \bar{E}_N(\sigma(i)) \\
= & \sum_{i\in\mathcal{I}_0,k} |u_{i,k}|^2 D_N(k) \bar{E}_N(\sigma(i) + \sum_{i\notin\mathcal{I}_0,k} |u_{i,k}|^2 D_N(k) \bar{E}_N(\sigma(i)) \\
\leq & \sum_{i,k} |u_{i,k}|^2 D_N(k) \left[\hat{E}_N(i) + 2\delta\right] + d_m M [\delta N M/\delta'] \\
\leq & \sum_{i,k} |u_{i,k}|^2 D_N(k) \hat{E}_N(\sigma(i)) + \mathcal{O}\left(d_m N \frac{M^2\delta}{\delta'}\right)
\end{aligned}$$

On remarquera que l'on a utilisé à la troisième ligne le fait que $U$ est une matrice unitaire et donc que $\sum_i |u_{i,k}|^2 = 1$, ainsi que le fait que $D_N$ est borné par $d_m$.
En jouant sur $\delta'$, qui est arbitraire, on peut toujours supposer que $d_m \frac{M^2\delta}{\delta'} = \sqrt{\delta}$ et alors, on a :

$$I_N^{(\beta)}\left(D_N, \bar{E}_N\right) \leq e^{N^2\sqrt{\delta}} I_N^{(\beta)}\left(D_N, \hat{E}_N\right)$$

et donc, comme on obtient une in{egalité du même type en échangeant les rôles de $\bar{E}$ et de $\hat{E}$, on a montré le point 1 du théorème (dans le cas plus restrictif où les $\bar{E}_N, \hat{E}_N$ sont bornés) avec $g(\delta, L) = \sqrt{\delta}$.

5. Il reste à prouver les points 2 et 3 du théorème.
On utilise pour cela le lien évident qui unit matrices de Wigner non-centrées et intégrales sphériques. On rappelle que $Y_N = D_N + X_N$ où $X_N$ est une matrice de Wigner centrée. On a, en faisant appel à ce que nous savions déjà sur $X_N$ et sa densité, puis en développant :

$$\begin{aligned}
& \mathbb{P}\left(d\left(\hat{\mu}_Y^N, \mu\right) < \delta\right) \\
= & 1/Z \int_{d\left(\hat{\mu}_Y^N,\mu\right)<\delta} \exp\left[-\frac{N\beta}{4} Tr X_N^2\right] dY \\
= & 1/Z \exp\left[-\frac{N\beta}{4} Tr D_N^2\right] \times \\
& \int\int_{d\left(\hat{\mu}_Y^N,\mu\right)<\delta} \exp\left[-\frac{N\beta}{2} Tr\left(U D(\lambda) U^* D_N\right) - \frac{N\beta}{4}\sum_{i=1}^N \lambda_i^2\right] \Delta(\lambda)^\beta \prod_{i=1}^N d\lambda_i dm(U) \\
= & 1/Z \exp\left[-\frac{N\beta}{4} Tr D_N^2\right] \int_{d\left(\frac{1}{N}\sum_i \delta_{\lambda_i},\mu\right)<\delta} I_N^{(\beta)}\left(D(\lambda), D_N\right) \exp\left[-\frac{N\beta}{4}\sum_{i=1}^N \lambda_i\right] \Delta(\lambda)^\beta \prod \lambda_i
\end{aligned}$$

où on a noté $D(\lambda)$ la matrice diagonale des $\lambda_i$.

En utilisant le résultat du point 1, on a :

$$\begin{aligned}
& \mathbb{P}\left(d\left(\hat{\mu}_Y^N, \mu\right) < \delta\right) \\
\leq\ & 1/Z \exp\left[\log I^{(\beta)}\left(\mu_D, \mu\right) + g\left(\delta, L\right) + N^2 \frac{\beta}{2} \int x^2 d\hat{\mu}_{D_N}^N\left(x\right)\right] \times \\
& \int_{d\left(\frac{1}{N}\sum_i \delta_{\lambda_i}, \mu\right) < \delta} \exp\left[-\frac{N\beta}{4} \sum_i \lambda_i^2\right] \Delta\left(\lambda\right)^\beta \prod_i d\lambda_i \\
=\ & \exp\left[-\frac{N^2\beta}{4} \int x^2 d\hat{\mu}_{D_N}^N\left(x\right) + \log I^{(\beta)}\left(\mu_D, \mu\right) + g\left(\delta, L\right)\right] \mathbb{P}\left(d\left(\hat{\mu}_X^N, \mu\right) < \delta\right)
\end{aligned}$$

On obtient de même une minoration, avec un $-g\left(\delta, L\right)$ cette fois.
D'où le résultat lorsqu'on fait $N \to \infty$ puis $\delta \to 0$.

□

# 5 Grandes déviations pour des matrices de Wigner non centrées

On s'intéresse dans cette partie à des matrices de Wigner non centrées $Y_N$ :

$$Y_N := D_N + X_N$$

où $X_N$ est une matrice de Wigner standard, telle que nous l'avons rencontrée au début de ce travail, et $D_N$ une matrice diagonale déterministe.
En fait, les techniques que nous utiliserons s'appuyeront sur le processus suivant :

$$Y_N\left(t\right) := D_N + H_N\left(t\right)$$

ùo $\left(H_N\left(t\right)\right)_t$ est un processus à valeurs dans l'espace des matrices hermitiennes (ou symétrique ou symplectique) dont les entrées sont des mouvements brownien, de telle sorte que $H_N\left(1\right) = X_N$.
Pour illustrer les techniques utilisées, nous allons d'abord étudier un cas plus simple qui nous permettra d'illustrer la méthode et de bien la comprendre.

## 5.1 Exemple d'application : Mesure empirique du mouvement brownien

Nous nous intéressons aux grandes déviations de la quantité suivante :

$$\hat{\mu}_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{B_t^i}$$

où les $B^i$ sont des mouvements browniens indépendants. Nous allons en fait démontrer un principe de grandes déviations faibles, ce qui revient à supposer que $\hat{\mu}_t^N$ est exponentiellement tendu. On peut le montrer mais on ne s'y attardera pas ici.

### 5.1.1 Définitions utiles et martingales

On notera $C$ l'espace dans lequel nous raisonnerons, c'est-à-dire $\mathcal{C}\left([0,1],\mathcal{P}\left(\mathbb{R}\right)\right)$ et $D$ la distance suivante sur $C$ :

$$D\left(\mu,\nu\right) := sup_t d\left(\mu_t,\nu_t\right)$$

De plus, pour tout $f,g$ dans $\mathcal{C}_b^{2,1}\left(\mathbb{R}\times[0,1]\right)$ et $\mu$ dans $C$ :

$$T^{0,t}\left(f,\mu\right) := \int f_t d\mu_t - \int f_0 d\mu_0 - \int_0^t \int \partial_2 f_s d\mu_s ds - \int_0^t \int \frac{1}{2}\partial_1^2 f_s d\mu_s ds$$

et

$$< f,g >_{\mu}^{0,t} := \int_0^t \int \partial_1 f_s \partial_1 g_s d\mu_s ds$$

Alors, on a :

**Théorème 5.1.** $\left(T^{0,t}\left(f,\hat{\mu}^N\right)\right)_t$ *est une martingale dont le crochet est exactement* $\frac{1}{N}\left(< f,g >_{\hat{\mu}^N}^{0,t}\right)_t$

*Démonstration.* Posons

$$F\left(x_1,\ldots,x_N,t\right) := \frac{1}{N}\sum_{i=1}^N f\left(x_i,t\right)$$

Alors, $F\left(B_t^1,\ldots,B_t^N,t\right) = \int f\left(x,t\right) d\hat{\mu}_t^N\left(x\right)$. D'où, en appliquant la formule d'Itô à $F$ :

$$\begin{aligned}\sum_i \int_0^t \partial_{x_i} F\left(B_s^1,\ldots,B_s^N,s\right) dB_s^i &= F\left(B_t^1,\ldots,B_t^N,t\right) - F\left(0,\ldots,0,0\right) - \\ &\int_0^t \partial_s F\left(B_s^1,\ldots,B_s^N,s\right) ds - \\ &\frac{1}{2}\sum_{i,j}\int_0^t \partial_{x_i}\partial_{x_j} F\left(B_s^1,\ldots,B_s^N,s\right) ds\end{aligned}$$

Le membre de gauche est une martingale et donc le membre de droite aussi. Or, si on se rappelle de l'expression de $F$, on voit que le membre de droite correspond en fait à $T^{0,t}\left(f,\hat{\mu}^N\right)$.
Le crochet est :

$$\begin{aligned}\sum_i \int_0^t \left[\partial_{x_i} F\left(B_s^1,\ldots,B_s^N,s\right)\right]^2 ds &= \sum_i \int_0^t \left[\frac{f\left(B_s^i,s\right)}{N}\right]^2 ds \\ &= \frac{1}{N} < f,f >_{\hat{\mu}^N}^{0,t}\end{aligned}$$

D'où le résultat voulu ! □

On remarque en particulier que l'on a le résultat suivant, dû aux martingales exponentielles (en considérant la martingale $NT^{0,t}\left(f,\hat{\mu}^N\right)$) :

$$\mathbb{E}\left[\exp\left(NT^{0,t}\left(f,\hat{\mu}^N\right)-\frac{N}{2}<f,f>_{\hat{\mu}^N}^{0,t}\right)\right]=1$$

On posera, pour raisons de commodité,

$$T\left(f,\mu\right):=T^{0,1}\left(f,\mu\right)-\frac{1}{2}<f,f>_{\mu}^{0,1}$$

### 5.1.2 Borne supérieure

Nous allons à présent démontrer une borne supérieure pour le principe faible de grande déviations, c'est-à-dire que nous allons montrer quelque chose du type :

$$\limsup_{\delta\to 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\left(D\left(\hat{\mu}^N,p\right)\leq\delta\right)\leq -S\left(p\right)$$

Soit donc $p\in C$. On peut supposer $p_0=\delta_0$ car, p.s., $\hat{\mu}_0^N=\delta_0$. On a alors :

$$\begin{aligned}
\mathbb{P}\left(D\left(\hat{\mu}^N,p\right)\leq\delta\right) &= \mathbb{E}\left(\mathbb{1}_{D(\hat{\mu}^N,p)\leq\delta}\frac{\exp\left[NT\left(f,\mu\right)\right]}{\exp\left[NT\left(f,\mu\right)\right]}\right)\\
&\leq \exp\left[-N\inf_{B(p,\delta)}T\left(f,.\right)\right]\mathbb{E}\left(\mathbb{1}_{D(\hat{\mu}^N,p)\leq\delta}\exp\left[NT\left(f,\mu\right)\right]\right)\\
&\leq \exp\left[-N\inf_{B(p,\delta)}T\left(f,.\right)\right]\mathbb{E}\left(\exp\left[NT\left(f,\mu\right)\right]\right)\\
&= \exp\left[-N\inf_{B(p,\delta)}T\left(f,.\right)\right]
\end{aligned}$$

La dernière ligne provenant de la remarque faite au paragraphe précédant concernant la martingale exponentielle. On obtient finalement :

$$\limsup_{\delta\to 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\left(D\left(\hat{\mu}^N,p\right)\leq\delta\right)\leq -T\left(f,p\right)$$

Cela est valable pour tout $f\in\mathcal{C}_b^{2,1}\left(\mathbb{R}\times[0,1]\right)$. On peut donc prendre le sup sur ces fonctions et cela nous donne un candidat pour une bonne fonction de taux :

$$\begin{aligned}
S\left(p\right) &:= \sup_{f\in\mathcal{C}_b^{2,1}(\mathbb{R}\times[0,1])}\left[T^{0,1}\left(f,p\right)-\frac{1}{2}<f,f>_p^{0,1}\right]\\
&= \sup_{f\in\mathcal{C}_b^{2,1}(\mathbb{R}\times[0,1])}\sup_{\lambda}\left[\lambda T^{0,1}\left(f,p\right)-\frac{\lambda^2}{2}<f,f>_p^{0,1}\right]\\
&= \frac{1}{2}\sup_{f\in\mathcal{C}_b^{2,1}(\mathbb{R}\times[0,1])}\frac{T^{0,1}\left(f,p\right)^2}{<f,f>_p^{0,1}}
\end{aligned}$$

où l'on a écrit $f$ sous la forme $\lambda f$ et on a maximisé un polynôme de degré 2
L'obtention de la borne supérieure a été relativement aisé. On remarquera qu'il s'agit d'un argument de type Girsanov, c'est-à-dire que l'on a en fait réalisé un changement de probabilité à l'aide des martingales exponentielles. La borne inférieure va être un peu plus retors.

### 5.1.3 Borne inférieure

On va reprendre l'idée du changement de probabilité à la Girsanov et on va donc poser $\mathbb{P}^{N,k}$ la probabilité de densité

$$\exp\left[NT\left(k,\hat{\mu}^N\right)\right]$$

Sous $\mathbb{P}^{N,k}$, la suite de mesure $\hat{\mu}^N$ est presque sûrement tendue. En effet, $\hat{\mu}^N$ est exponentiellement tendue sous $\mathbb{P}$ :

$$\text{Pour tout } M>0, \text{ il existe } K \text{ compact tq } \limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\hat{\mu}^n\in K^c\right)\leq -M$$

Or, la densité de $\mathbb{P}^{N,k}$ est bornée par un $\exp\left(C\left(k\right)\right)$. D'où :

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}^{N,k}\left(\hat{\mu}^n\in K^c\right) &\leq \limsup_{n\to\infty}\frac{\log\mathbb{P}\left(\hat{\mu}^n\in K^c\right)+C\left(k\right)}{n}\\ &\leq -M+C\left(k\right)\end{aligned}$$

Donc, sous $\mathbb{P}^{N,k}$, cela reste exponentiellement tendu !
Donc, par Borel-Cantelli, p.s. les $\hat{\mu}^N$ sont dans $K$ au bout d'un certain rang. Par conséquent, p.s. les $\hat{\mu}^N$ sont tendues.
Or, Girsanov nous dit que la quantité suivante est une martingale (en t) sous la probabilité modifiée (pour $f$ fixée) :

$$M_t^N := T^{0,t}\left(\hat{\mu}^N,f\right)-\int_0^t\int\partial_x f_u\left(x\right)\partial_x k_u\left(x\right)d\hat{\mu}^N\left(x\right)du$$

De plus, le crochet de cette martingale vaut :

$$\frac{1}{N}\int_0^t\int\left[\partial_x f_u\left(x\right)\right]^2 d\hat{\mu}^N\left(x\right)du\leq\frac{\sup f}{N}$$

ce qui tend vers 0 à $t$ fixé et l'inégalité de Burkholder-Davis-Gundy permet alors d'affirmer que la martingale $M_t^N$ converge vers 0 uniformément pour $t\in[0,1]$.
Soit alors $\mu$ une valeur d'adhérence de $\hat{\mu}^N$ (qui existe p.s. sous $\mathbb{P}^{N,k}$ car la suite de mesures est p.s. tendue).$\mu$ vérifie alors :

$$T^{0,1}\left(\mu,f\right)=\int_0^1\int\partial_x f_u\left(x\right)\partial_x k_u\left(x\right)d\mu\left(x\right)du\forall f\in\mathcal{C}_b^{2,1}\left(\mathbb{R}\times[0,1]\right)\qquad(8)$$

Nus aurons alors besoin du lemme suivant :

**Lemme 5.1.** *L'équation 8 admet, pour k fixé, une et une seule solution $\nu$.*

*Démonstration.* Admis □

On en déduit alors que $\mu^N$ converge sous la probabilité modifiée vers cette unique solution, que l'on peut par exemple appeler $\nu_t^k$. On a alors :

$$\begin{aligned}
& \liminf_{\delta\to 0}\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}\left(\sup_{t\in[0,1]} d\left(\hat{\mu}_t^N,\nu_t^k\right)<\delta\right)\\
= & \liminf_{\delta\to 0}\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{E}^{N,k}\left[\mathbb{1}_{\sup d\left(\hat{\mu}_t^N,\nu_t^k\right)<\delta}\exp\left(-NT\left(k,\hat{\mu}^N\right)\right)\right]\\
\geq & \liminf_{\delta\to 0}\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{E}^{N,k}\left[\mathbb{1}_{\sup d\left(\hat{\mu}_t^N,\nu_t^k\right)<\delta}\exp\left(-NT\left(k,\nu^k\right)-N\epsilon\right)\right]\\
\geq & -T\left(k,\nu^k\right)-\epsilon+\liminf_{\delta\to 0}\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}^{N,k}\left(\sup_{t\in[0,1]} d\left(\hat{\mu}_t^N,\nu_t^k\right)<\delta\right)
\end{aligned}$$

Où l'on a prix $\epsilon$ strictement positif quelquonque. La troisiùme ligne provient de la continuité de $\mu\mapsto T\left(k,\mu\right)$ et est donc vraie dès que $\delta$ est assez petit. Comme on sait par ailleurs que sous $\mathbb{P}^{N,k}$, on a la convergence étroite de $\hat{\mu}^{N,k}$ vers $\nu^k$, on a que le dernier terme de la dernière ligne vaut 0. D'où :

$$\liminf_{\delta\to 0}\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}\left(\sup_{t\in[0,1]} d\left(\hat{\mu}_t^N,\nu_t^k\right)<\delta\right)\geq -T\left(k,\nu^k\right)$$

Comme ceci est valable pour tout $k$, on a finalement :

$$\liminf_{\delta\to 0}\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}\left(\sup_{t\in[0,1]} d\left(\hat{\mu}_t^N,\nu_t^k\right)<\delta\right)\geq -S\left(\nu^k\right)$$

Ce qui prouve la borne inférieure du principe de grandes déviation. On voit donc quelle est la méthode suivie : on utilise ici encore un changement de probabilité et, moyennant l'étude d'une équation et de l'unicité de ses solutions, on utilise la convergence (étroite) sous cette nouvelle probabilité.

## 5.2 Principe de grande déviation pour une matrice de Wigner non-centrée

Passons donc maintenant aux "choses sérieuses" : les matrices de Wigner non-centrées telles que nous les avons définies précédemment. Pour des raisons de commodités, nous nous limiterons dans la suite au cas $\beta = 2$.

### 5.2.1 Quelques définitions

Comme dans l'exemple qui vient d'être traité, il va nous falloir recourir à des martingales. On pose, pour $f$ et $g$ dans $\mathcal{C}_b^{2,1}\left(\mathbb{R}\times[0,1]\right)$, $0\leq s\leq t\leq 0$ et $\nu$

dans $\mathcal{C}\left([0,1],\mathcal{P}\left(\mathbb{R}\right)\right)$ :

$$\begin{aligned} S^{s,t}\left(\nu,f\right) &= \int f\left(x,t\right)d\nu_t\left(x\right)-\int f\left(x,s\right)d\nu_s\left(x\right)-\int_s^t\int \partial_2 f\left(x,u\right)d\nu_u\left(x\right)du \\ &- \frac{1}{2}\int_s^t\int\int \frac{\partial_1 f\left(x,u\right)-\partial_1 f\left(y,u\right)}{x-y}d\nu_u\left(x\right)d\nu_u\left(y\right)du \\ <f,g>_\nu^{s,t} &= \int_s^t\int \partial_1 f\left(x,u\right)\partial_1 g\left(x,u\right)d\nu_u\left(x\right)du \\ \bar{S}^{s,t}\left(\nu,f\right) &= S^{s,t}\left(\nu,f\right)-\frac{1}{2}<f,f>_\nu^{s,t} \end{aligned}$$

Evidemment, on s'y attend, on a le résultat suivant :

**Théorème 5.2.** *Pour tout $N$ entier naturel, pour tout $f$ dans $\mathcal{C}_b^{2,1}\left(\mathbb{R}\times[0,1]\right)$ et pour tout $s\in[0,1[$, on a que :*

$$\left(S^{s,t}\left(\hat{\mu}^N,f\right)\right)_{s\leq t\leq 1}$$

*est une martingale borée dont le crochet vaut :*

$$\frac{1}{N^2}<f,f>_{\hat{\mu}^N}^{s,t}$$

On pose en outre, pour toute mesure $\mu$ de probabilité :

$$S_\mu\left(\nu\right):=\begin{cases} \infty & \text{si } \nu_0\neq\mu \\ \sup_{f\in\mathcal{C}_b^{2,1}(\mathbb{R}\times[0,1])}\sup_{0\leq s\leq t\leq 1}\bar{S}^{s,t}\left(\nu,f\right) & \text{sinon}\end{cases}$$

### 5.2.2 Prérequis

Nous allons admettre que $S_{\mu_D}$ est une bonne fonction de taux.
Nous allons aussi admettre le résultat suivant : $\hat{\mu}^N$ est exponentiellement tendu. De la sorte, nous pouvons nous restreindre à la démonstration d'un principe de grandes déviations faible.

### 5.2.3 Borne supérieure

Du fait de la tension exponentielle, il suffit de montrer une borne supérieure pour un principe de grandes déviations faible. C'est le même raisonnement que précédemment. Soit donc $p\in C$. On peut supposer $p_0=\mu_D$ car, p.s., $\hat{\mu}_0^N=\mu_D$. On a alors :

$$\begin{aligned} \mathbb{P}\left(D\left(\hat{\mu}^N,p\right)\leq\delta\right) &= \mathbb{E}\left(\mathbb{1}_{D(\hat{\mu}^N,p)\leq\delta}\frac{\exp\left[N^2\bar{S}^{s,t}\left(f,\mu\right)\right]}{\exp\left[N^2\bar{S}^{s,t}\left(f,\mu\right)\right]}\right) \\ &\leq \exp\left[-N^2\inf_{B(p,\delta)}\left\{\bar{S}^{s,t}\left(f,.\right)u\right\}\right]\mathbb{E}\left(\exp\left[N^2\bar{S}^{s,t}\left(f,\mu\right)\right]\right) \\ &= \exp\left[-N^2\inf_{B(p,\delta)}\bar{S}^{s,t}\left(f,.\right)\right] \end{aligned}$$

On remarquera en particulier que, puisque le crochet de $S^{s,t}$ est $\frac{1}{N^2} < ., . >^{s,t}$, il faut consid{erer cette fois la martingale $N^2 S^{s,t}$ dans l'exponentielle, car son crochet sera alors $N^2 < ., . >^{s,t}$ et on voit donc bien apparaitre $N^2 \bar{s}^{s,t}$.
D'où u

$$\lim_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{N^2} \log \mathbb{P}\left(D\left(\hat{\mu}^N, p\right) < \delta\right) \leq -\bar{S}^{s,t}\left(p, f\right)$$

On obtient alors le résultat en optimisant sur $f$ et $(s,t)$ et en remarquant que l'on a $-\infty$ si $p_0 \neq \mu_D$.

### 5.2.4 Borne inférieure : le cas de $\mathcal{MC}$

Nous emploirerons à nouveau une méthode fondée sur l'unicité des solutions à une équation. Plus précisément, on s'intéresse à l'équation suivante :

$$S^{s,t}\left(f, \nu\right) = < f, k >_{\nu}^{s,t} \quad \forall f \in \mathcal{C}_b^{2,1}\left(\mathbb{R} \times [0,1]\right) \tag{9}$$

On remarquera, en raisonnant comme dans l'exemple de la mesure empirique du mouvement brownien, que l'unicité d'une solution $\nu$ à cette équation implique que $\hat{\mu}^N$ converge étroitement vers $\nu$ sous la probabilité modifiée, absolument continue par rapport à $\mathbb{P}$ et de densité :

$$\exp\left[N^2 \bar{S}^{0,1}\left(k, .\right)\right]$$

Nous appelerons cette probabilité à nouveau $\mathbb{P}^{N,k}$.
Or, on a le théorème suivant concernant l'unicité :

**Théorème 5.3.** *Posons : $\mathcal{MF}$ l'ensemble des fonctions $h$ de $\mathcal{C}_b^{\infty,1}\left(\mathbb{R} \times [0,1] \cap \mathcal{C}\left([0,1], L^2\left(\mathbb{R}\right)\right)\right.$ telles qu'il existe $C$ et $\epsilon$ tous les deux strictement positifs et vérifiant :*

$$\sup_{t \in [0,1]} |\hat{h}_t\left(\lambda\right)| \leq C e^{-\epsilon |\lambda|}$$

*où $\hat{h}$ désigne la transformée de Fourier.*
*Alors, pour tout $h$ dans $\mathcal{MF}$ il existe une et une seule solution à l'équation et on note $\mathcal{MC}$ l'ensemble de telles solutions.*

On admettra ce théorème.
Supposons donc à présent que $\mathcal{O}$ est un ouvert de $\mathcal{C}\left([0,1], \mathcal{P}\left(\mathbb{R}\right)\right)$ et $\nu \in \mathcal{O} \cap \mathcal{MC}$ et $\delta$ strictement positif suffisament petit, tel que l'on puisse écrire :

$$\begin{aligned}
\mathbb{P}\left(\hat{\mu}^N \in \mathcal{O}\right) &\geq \mathbb{P}\left(d\left(\hat{\mu}^N, \nu\right) < \delta\right) \\
&= \mathbb{E}^{N,k}\left[\mathbb{1}_{d(\hat{\mu}^N, \nu) < \delta} \exp\left(-N^2 S^{0,1}\left(\hat{\mu}^N, k\right) - \frac{N^2}{2} < k, k >_{\hat{\mu}^N}^{0,1}\right)\right] \\
&\geq \exp\left(-N^2 S^{0,1}\left(\nu, k\right) - \frac{N^2}{2} < k, k >_{\nu}^{0,1} - N^2 g\left(\delta\right)\right) \mathbb{P}^{N,k}\left(d\left(\hat{\mu}^N, \nu\right) < \delta\right)
\end{aligned}$$

où $g\left(\delta\right)$ tend vers 0 quand $\delta$ tend vers 0.
Et par conséquent, en utilisant de manière analogue à l'exemple d'application

précédemment traité la convergence étroite sous $\mathbb{P}^{N,k}$, on trouve :

$$\liminf_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}\left(\hat{\mu}^N \in \mathcal{O}\right) \geq -\left(S^{0,1}\left(\nu, k\right) - \frac{1}{2} < k, k >_{\nu}^{0,1}\right)$$

On identifie dans le membre de droite cette quantité comme étant $-S_{\mu_D}(\nu)$. On prendra bien note cependant du fait que cette formule est vrai, d'après ce que nous avons montré, *a priori* seulement pour $\nu \in \mathcal{O} \cap \mathcal{MC}$. En prenant le sup dans le membre de droite sur cet ensemble, on obtient alors :

$$\liminf_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}\left(\hat{\mu}^N \in \mathcal{O}\right) \geq -\inf_{\mathcal{O}\cap\mathcal{MC}} S_{\mu_D}$$

Nous devons maintenant chercher à nous affranchir de cette restriction à $\mathcal{MC}$ pour obtenir une vraie borne inférieure de principe de grande déviation. Pour cela, nous allons procéder par étapes.

### 5.2.5 Borne inférieure : le cas de $\mathcal{A}$

Nous allons d'abord nous appuyer sur le résultat suivant :

**Théorème 5.4** (Approximation de chemins de mesures)**.** *Soit $\nu$ dans*

$$\mathcal{A} := \left\{ \mu \in \mathcal{C}\left([0,1], \mathcal{P}(\mathbb{R})\right) \text{ tel que } \exists \epsilon > 0 \sup_t \mu_t\left(|x|^{5+\epsilon}\right) < \infty \right\}$$

*Posons en outre :*

$$\nu_t^{\epsilon,\Delta} := P^\epsilon * \nu_{t_k} + \frac{t - t_k}{\Delta}\left(P^\epsilon * \nu_{t_{k+1}} - P^\epsilon * \nu_{t_k}\right)$$

*où $P^\epsilon$ est une famille de lois de Cauchy et $0 = t_1 < t_2 < \ldots < t_n = 1$ une partition de pas $\Delta$. On a en outre choisi $k$ tel que $t \in [t_k, t_{k+1}]$.*
*Alors :*

1. *$\nu^{\epsilon,\Delta}$ définit bien un chemin dans l'espace des mesures de probabilités.*
2. *$\nu^{\epsilon,\Delta}$ est dans $\mathcal{MC}$.*
3. *$\lim_{\epsilon\to 0} \lim_{\Delta\to 0} S^{0,1}\left(\nu^{\epsilon,\Delta}\right) = S^{0,1}(\nu)$*
4. *$\lim_{\epsilon\to 0} \lim_{\Delta\to 0} \nu^{\epsilon,\Delta} = \nu$*

*Démonstration.* On ne prouvera pas tout le théorème. On verra cependant dans la partie consacrée aux probabilités libres comment montrer la diminution d'entropie par convolution libre (la convolution avec une loi de Cauchy est la même chose que la convolution libre par cette loi, d'où l'intérêt de considérer ce type de lois). □

Le résultat précédent nous permet d'élargir le résultat par approximation, mais il reste un problème : $\nu_0^{\epsilon,\Delta} = \mu_D * P_\epsilon$. En d'autres termes, la condition

initiale a été changée !
En fait, on peut montrer que

$$X_t^{N,\epsilon} := U_N C_\epsilon^N U_N^* + D_N + H_t^N$$

où $U_N$ est choisie selon la mesure de Haar et $C_\epsilon^N$ est une matrice diagonale dont la mesure spectrale converge vers une loi de Cauchy, a bien pour condition initiale $P_\epsilon \boxplus \mu_D = P_\epsilon * \mu_D$. (Voir la partie sur les probabilités libres, au sujet de $\boxplus$.) On a donc, en notant $\hat{\mu}_t^{N,\epsilon}$ la mesure spectrale de $X_t^{N,\epsilon}$, le résultat suivant :
Pour tout ouvert $\mathcal{O}$ :

$$\liminf_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}\left(\hat{\mu}^{N,\epsilon} \in \mathcal{O}\right) \geq -\inf\left\{S_{P_\epsilon * \mu_D}(\nu), \nu \in \mathcal{O}, \nu = P_\epsilon * \mu, \mu \in \mathcal{A} \cap \{S_{\mu_D} < \infty\}\right\}$$

Ce n'est bien sûr qu'une reformulation du résultat précédent ! Notre but est maintenant de l'étendre au cas $\epsilon = 0$.

**Théorème 5.5.** *Soit $\mathcal{O}$ un ouvert de $\mathcal{C}([0,1], \mathcal{P}(\mathbb{R}))$. Alors :*

$$\liminf_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}\left(\hat{\mu}^N \in \mathcal{O}\right) \geq -\inf_{\mathcal{O}\cap\mathcal{A}} S_{\mu_D}$$

*Démonstration.* On pose :

$$K_L := \left\{\mu \in \mathcal{P}(\mathbb{R}) \text{ tel que } \int \log\left(x^2+1\right) d\mu(x) \leq L\right\}$$

et

$$\mathcal{K}_L^{N,\epsilon} := \bigcap_{t\in[0,1]} \left\{\hat{\mu}_t^{N,\epsilon} \in K_L\right\} \cap \bigcap_{t\in[0,1]} \left\{\hat{\mu}_t^N \in K_L\right\}$$

Notons au passage que cette dernière quantité est mesurable, car les $\hat{\mu}_t$ sont des chemins **continus** et il suffit donc de prendre les intersections sur les rationnels par exemple.
On peut alors montrer qu'il existe une fonction $f$ de deux variables tel que :

$$D\left(\hat{\mu}^{N,\epsilon}, \hat{\mu}^N\right) \leq f(N,\epsilon)$$

$$\limsup_{\epsilon\to 0} \limsup_{N\to\infty} f(N,\epsilon) = 0$$

On peut de plus (nous l'admettrons ici), pour tout $M$ positif se donner $L_M$ tel que pour tout $L \geq L_M$ :

$$\sup_{0\leq\epsilon\leq 1} \mathbb{P}\left(\mathcal{K}_L^{N,\epsilon}\right) \leq \exp\left(-MN^2\right)$$

Prenons alors $M > S_{\mu_D}(\mu) + 1$ et $L = L_M$ associé. Pour $\delta$ strictement positif, on se donne $\epsilon > 0$ assez petit tel que $D(P_\epsilon * \mu, \mu) < \delta/3$ et on a alors :

$$\begin{aligned}
\mathbb{P}\left(\hat{\mu}^N \in B(\mu,\delta)\right) &\geq \mathbb{P}\left(D\left(\hat{\mu}^N, \mu\right) < \delta/3, \hat{\mu}^{N,\epsilon} \in B(P_\epsilon * \mu, \delta/3), \mathcal{K}_L^{N,\epsilon}\right) \\
&\geq \mathbb{P}\left(\hat{\mu}^{N,\epsilon} \in B(P_\epsilon * \mu, \delta/3)\right) - \mathbb{P}\left(\mathcal{K}_L^{N,\epsilon c}\right) - \mathbb{P}\left(D\left(\hat{\mu}^N, \mu\right) \geq \delta/3, \mathcal{K}_L^{N,\epsilon}\right)
\end{aligned}$$

Le second terme est majoré par $\exp\left[-N^2\left(S_{\mu_D}+1\right)\right]$.
Le troisième terme est nul pour $N$ assez grand et $\epsilon$ assez petit.
Le résultat sur $\hat{\mu}^{n,\epsilon}$ donne en outre que, pour tout $\eta > 0$ et pour $N$ assez grand et $\epsilon$ assez petit, le premier terme est minoré par :

$$\exp\left[-N^2 S_{P_\epsilon * \mu_D}\left(P_\epsilon * \mu\right) - N^2\eta\right] \geq \exp\left[-N^2 S_{\mu_D}\left(\mu\right) - N^2\eta\right]$$

Cette dernière inégalité provient du fait que pour la loi de Cauchy, convolution et convolution libre sont identique et que la convolution libre par une loi de probabilité fait diminuer l'entropie, comme on le montrera dans la partie consacrée aux probabilités libres.
Le théorème est donc démontré. □

#### 5.2.6 Borne inférieure : conclusion

Pour conclure la preuve, on doit montrer qu'il existe pour tout $\nu$ des $\nu_n$ dans $\mathcal{A}$ tel que les $\nu_n$ convergent étroitement vers $\nu$ et $S_{\mu_D}(nu_n)$ converge vers $S_{\mu_D}(\nu)$. Cette preuve ne présente pas un intérêt particulier et nous admettrons donc ce résultat.

# 6 Introduction : le concept de théorie des probabilités de variables non-commutatives

## 6.1 Idée générale

Nous allons à présent introduire quelques éléments de le théorie des probabilités libres. Ce cadre fournit des outils puissants pour l'étude des matrices aléatoires. Néanmoins, nous ne pourrons qu'esquisser des idées générales, car il s'agit là d'une théorie vaste qui ne saurait être étudiée en détails dans le cadre de ce mémoire !
L'idée fondamentale à avoir à l'esprit pour comprendre la suite est l'idée du passage du commutatif au non-commutatif. En gros, on peut associer à un espace l'algèbre des fonctions sur cet espace, de sorte que des proprieté sur l'espace se traduisent par des proprietés sur l'algèbre. Cette algèbre est bien sûr commutative. En prenant à présent des algèbre qui ne sont plus forcément commutative, l'étude de telles proprietés peut être vue comme l'étude des proprieté d'espaces "non commutatif" (qui n'ont bien sûr pas d'existence intrinsèque).
C'est ainsi que l'on peut créer une géométrie non-commutative, une topologie non-commutative et...une théorie des probabilités non-commutatives.
Dans notre cas, nous allons remplacer l'espace de probabilité $\Omega$ par l'algèbre de ses variables aléatoires. En fait, une définition "naïve" d'une telle généralisation non-commutative serait :

**Définition 6.1.** *Soit $\mathcal{A}$ une algèbre unitaire munie d'une application $\phi : \mathcal{A} \to \mathbb{C}$ telle que $\phi(1) = 1$. On dit alors que $\mathcal{A}$ est un espace de probabilité non-commutatif.*

On voit bien que $\phi$ joue le rôle de l'espérance dans le cas classique. Cette définition est certes cohérente mais nous aurons en fait besoin d'une structure un peu plus riche.

## 6.2 Définitions

Il nous faut tout d'abord rappeler un certain nombre d'objets provenant de la théorie des opérateurs :

**Définition 6.2.** *Une $C^*$-algèbre (ou algèbre stellaire) $(\mathcal{A}, *)$ est une algèbre normée munie d'une involution $*$ telle que :*

$$\forall X, Y \in \mathcal{A}, ||XY|| \leq ||X||||Y||, ||X^*|| = ||X||, ||XX^*|| = ||X||^2$$

*L'algèbre est dite unitaire si elle contient un élément neutre.*
*Un élément $X$ de $\mathcal{A}$ est dit autoadjoint si $X = X^*$*

L'exemple typique que l'on a en tête lorsqu'on introduit cette notion est bien sûr le cas des algèbre d'opérateurs lin''eaires bornés sur un espace de Hilbert. En fait, toute algèbre stellaire peut être vue comme une sous-algèbre stellaire d'une telle algèbre, par l'intermédiaire de la construction de Gelfand-Naimark-Segal. Fixons $H$ un espace de Hilbert et $B(H)$ son algèbre d'opérateurs associée. On définit alos :

**Définition 6.3.** *Soit $\mathcal{A}$ une sous-algèbre stellaire de $B$. On dit que $\mathcal{A}$ est une algèbre de von Neumann si elle est fermée pour la topologie engendrée par les semi-normes suivantes :*

$$X \mapsto < X\zeta, \eta >$$

*pour tout $\zeta, \eta$ dans $H$.*

Il est clair que cela ajoute de la structure : si, par exemple, $H = \mathbb{R}^2$, on pourrait prendre $\mathcal{A} = \left\{ \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix}, a \in \mathbb{R}^*, b \in \mathbb{R} \right\}$. Alors, c'est bien une sous-algèbre stellaire mais pas une algèbre de von Neumann.
Raisonnons par analogie avec les probabilités classiques. Si $(\Omega, \mathcal{B}, \mathbb{P})$ est un espace probabilisé, alors on peut lui associer canoniquement l'algèbre de von Neumann suivante :

$$\mathcal{A} := L^\infty(\Omega, \mathcal{B}, \mathbb{P})$$

munie de l'involution identité.
Il est alors naturel de vouloir considérer l'espérance d'éléments de $\mathcal{A}$. Nous allons à présent construire une "espérance" pour des variables non commutatives :

**Définition 6.4.** *Soit $\tau$ une forme linéaire sur $\mathcal{A}$ qui est une algèbre de von Neumann. On dit que $\tau$ est un état si :*

1. *$\tau(XX^*) \in \mathbb{R}^+$ pour $X \in \mathcal{A}$ (positivité)*
2. *$\tau(Id) = 1$ (masse totale unitaire)*

*Si, en plus, on a que pour tout $X, Y \in \mathcal{A}$, $\tau(XY) = \tau(YX)$ (tracialité), on dit que $\tau$ est un état tracial (par analogie avec la proprieté de la trace d'une matrice).*
*On dit alors que le couple $(\mathcal{A}, \tau)$ est un $W^*$-espace de probabilité.*

Donnons deux exemples de tels espaces : le cas sus-mentionné d'un espace de probabilité standard muni de l'espérance et le cas où $\mathcal{A} = M_n(\mathbb{C})$ avec pour état $\tau(M) = \frac{1}{n} Tr(M)$. Il s'agit en fait du seul état tracial possible sur cet $\mathcal{A}$, comme on peut s'en rendre compte en considérant les matrices élémentaires $\epsilon_{i,j}$. Dans le cas classique, la loi d'une variable aléatoire est entiérement déterminée par la donnée des espérances de cette variable sur, par exemple, les fonctions continues à support compact. Nous pouvons définir par analogie la loi de variables non-commutatives :

**Définition 6.5.** *Soient $A_1, \ldots, A_m$ $m$ éléments d'un $W^*$-espace de probabilité. On dit que $\tau_{A_1,\ldots,A_m}$ est leur loi, où :*

$$\forall P \in \mathbb{C} < X_1, \ldots, X_m >, \tau_{A_1,\ldots,A_m}(P) := \tau(P(A_1, \ldots, A_m))$$

*où $\mathbb{C} < X_1, \ldots, X_m >$ désigne l'espace des polynômes à $m$ variables qui ne commutent pas.*

On voit donc que les lois sont des formes linéaires sur l'espace des polynômes non commutatifs à $m$ variables, avec des proprietés spécifiques (pasitivité, tracialité, masse totale unitaire). Notons $\mathcal{M}^m$ l'ensemble de ces formes linéaires.
Donnons au passage une proprieté qui nous aidera à faire le lien avec le cas classique :

**Théorème 6.1.** *Soit $X$ un élément autoadjoint d'un $W^*$-espace de probabilité. Alors, sa loi $\tau_f$ est en fait une mesure $\mu$ à support compact sur la droite réelle au sens où*

$$\int P(t)\, d\mu(t) = \tau_f(P)$$

*pour tout polynôme $P$.*

*Démonstration.* Il s'agit d'une simple application du théorème de représentation de Riesz pour la forme linéaire $P\tau_f(P)$. La mesure est à support réel car $\tau_f$ ne prend que des valeurs réelles. En effet :

$$\tau\left(X^2\right) = \tau(XX*) \in \mathbb{R}$$

et donc cela reste vrai pour toute les puissances paires. pour les puissances impaires, par exemple 1 :

$$\tau\left((X+1)^2\right) = \tau\left(X^2\right) + 2\tau(X) + 1$$

C'est donc aussi réel. Riesz s'applique donc dans le cas réel. □

Existe-il toujours des opérateurs de loi donnée ?

**Théorème 6.2.** *Soit $\tau$ dans $\mathcal{M}^m$. Alors il existe un $W^*$-espace de probabilité$\mathcal{A}$ et $A_1, \ldots, A_m$ des éléments de $\mathcal{A}$, tels que $\tau$ soit la loi de ces éléments.*

*Démonstration.* On construit d'abord une forme bilinéaire sur $\mathbb{C} < X_1, \ldots, X_m >$ de la manière suivante :

$$< P, Q >:= \tau\left(PQ^*\right)$$

On désigne alors par $L^2(\tau)$ l'espace $\mathbb{C} < X_1, \ldots, X_m >$ vu avec cette forme bilinéaire.. Quitte à quotienter par les éléments de norme nulle, on peut supposer que c'est un vrai produit scalaire. On prend alors pour $H$ le completé de cet espace vectoriel, ce qui nous donne bien un espace de Hilbert.
De plus, $\mathbb{C} < X_1, \ldots, X_m >$ agit sur $L^2(\tau)$, et donc sur $H$, par multiplication à gauche. On considère alors $\mathcal{A}$ comme étant $\mathbb{C} < X_1, \ldots, X_m >$, vu comme agissant sur $H$, completé pour la topologie engendrée par les semi-normes $X \mapsto < X\zeta, \eta >$. On a alors bien une algèbre de von Neumann. Si on la munit de l'état tracial $\tau$, on a alors bien le résultat. □

Continuons notre raisonnement par analogies. Dans le cas classique, l'ensemble des mesures de probabilités était muni d'une topologie, celle de la convergence en loi. Pouvons-nous faire de même ici ?

**Définition 6.6.** *Soient $\tau_n$ des lois de $m$ variables non-commutatives. Soit $\tau$ une autre telle loi. On dit que $\tau_n$ converge vers $\tau$ (ou encore que les variables sous-jacentes des $\tau_n$ convergent en loi vers celles de $\tau$) si :*

$$\tau_n(P) \to \tau(P)$$

*pour tout $P$ de $\mathbb{C} < X_1, \ldots, X_m >$ .*

Dans le cas classique, nous disposions d'une métrique (par exemple celle de Skorokhod). Ici, quitte à faire une hypothèse supplémentaire, nous pouvons aussi nous y ramener.

**Théorème 6.3.** *Soit $\mathcal{M}_R^m := \{\mu \in \mathcal{M}^m, |\mu(X_i^p)| \leq R^p, \forall p \in \mathbb{N}, 1 \leq i \leq m\}$. Alors, $\mathcal{M}_R^m$, muni de la topologie définie par la convergence au sens précédant, est un espace polonais.*

*Démonstration.* Nous devons d'abord définir une **distance** compatible avec la notion de convergence définie précédemment.
On définit d'abord un sous-ensemble $B$ de polynômes. Soit

$$N(\mu) := \sum_{p \geq 0} \frac{|\mu(X_i^p)|}{(2R)^p}$$

qui est une norme sur $\mathcal{M}_R^m$. On peut alors voir un polynôme $P$ comme une forme linéaire sur $\mathcal{M}_R^m$, de la manière suivante :

$$\mu \mapsto \mu(P)$$

On appelle $B$ l'ensemble des polynômes tels que leur norme d'opérateur par rapport à $N$ soit inférieure à 1.
Soit $(P_n)_n$ une suite dense de polynômes dans $B$
Soient $\mu$ et $\nu$ dans $\mathcal{M}_R^m$. On définit la distance comme suit :

$$d(\mu,\nu) := \sum_{n\geq 0} \frac{1}{2^n} |\mu(P_n) - \nu(P_n)|$$

C'est bien défini car :

$$|\mu(P_n) - \nu(P_n)| \leq N(\mu - \nu) \leq 4$$

car on se trouve sur $\mathcal{M}_R^m$.
Remarquons que cette distance métrise bien la topologie voulue. En effet, supposons que $\mu_n$ converge vers $\mu$ au sens précédemment défini. Soit $\epsilon$ strictement positif et $M$ entier naturel tel que $\sum_{i\geq M} \frac{2R}{2^i} < \frac{\epsilon}{2}$. On a a par ailleurs, $\mu_n(P_k) \to \mu(P_k)$ pour tout $k \in \{0,\dots,M\}$. En prenant $N$ assez grand, on a alors pour toit $n \geq N$ :

$$\sum_{i=0}^{M} \frac{1}{2^i} |\mu_n(P_i) - \mu(P_i)| < \frac{\epsilon}{2}$$

et donc $d(\mu_n,\mu) < \epsilon$.
Réciproquement, supposons que $\mu_n$ converge vers $\mu$ au sens de la distance $d$. Soit $\epsilon > 0$. Soit $N$ tel que pour tout $n > N$, on ait :

$$d(\mu_n,\mu) < \frac{\epsilon}{3}$$

Soit $P$ un polynôme. On se donne alors $P_k$ un polynôme de la suite-dense tel que $|||P_k - P||| \leq \frac{\epsilon}{6R}$. On a :

$$|\mu_n(P) - \mu(P)| \leq |\mu_n(P) - \mu_n(P_k)| + |\mu_n(P_k) - \mu(P_k)| + |\mu(P_k) - \mu(P)|$$

Le tout est alors clairement majoré par $\epsilon$. D'où la convergence en loi.
C'est **complet**. En effet, soit $(\mu_n)_n$ une suite de Cauchy. Soit $P_k$ un des polynômes de la suite dense dans $B$. On a en particulier que $(\mu_n(P_k))_n$ est aussi une suite de Cauchy (et ce même uniformément en $k$), dans $\mathbb{R}$ cette fois, et est donc convergente vers un certain $a_k$.
On se donne $\mu$ forme linéaire sur $\mathbb{C} < X_1,\dots,X_m >$ telle que $\mu(P_k) = a_k$ pour tout entier $k$. Cela définit entiérement $\mu$ car pour tout polynôme $P$, on peut supposer qu'il est dans $B$ quitte à le diviser par un scalaire et il y a alors une suite $P_{k_i}$ qui converge vers $P$. On a, pour $\epsilon > 0$ et $n$ assez grand :

$$|a_{k_{i+j}} - a_{k_i}| \leq \epsilon + |\mu_n\left(P_{k_{i+j}}\right) - \mu_n(P_{k_i})|$$

Comme la convergence vers les $a_k$ a lieu uniforément en $k$, quitte à prendre $i$ assez grand, on voit bien que la suite $(a_{k_i})_i$ est de Cauchy et converge vers un $a$.

Ce qui est implicite dans cette affirmation, c'est que $\mu_n$ est continu par rapport aux polynômes, mais cela provient directement du fait que les monômes sont bornés par $R^d$, où $d$ est le degré du monôme.
On pose alors

$$\mu(P) = a$$

On note au passage que cette définition est bien indépendante de la suite $P_{k_i}$ choisie car si une autre suite $P_{l_i}$ converge vers $P$, alors en considérant l'union des deux suites, toujours de Cauchy, on constate que les limites des $\mu(P_{k_i})$ et des $\mu(P_{l_i})$ sont identique.
On vérifie aussi sans peine que $\mu$ appartient bien à $\mathcal{M}_R^m$, d'où la complétude.
C'est aussi **séparable**. Il suffit pour le voir de considérer les polynômes à coefficients rationnels. □

Pour essayer de comprendre ce que représente $\mathcal{M}_R^m$, on donne le résultat suivant :

**Théorème 6.4.** *$\mathcal{M}_R^1$ peut être identifié avec l'ensemble des lois de probabilités sur $[-R, R]$.*

*Démonstration.* Soit $\mu$ une telle loi de probabilité, alors :

$$\mu(X^p) = \int x^p d\mu(x) \leq R^p$$

De plus, l'espérance vérifie trivialement les conditions pour être un état tracial. Réciproquement, soit $\mu$ dans $\mathcal{M}_R^1$. On a que $P \mapsto \mu(P)$ s'étend de manière unique en une forme linéaire sur les fonctions réelles à supports compacts. par le théorème de représentation de Riesz, il existe une unique mesure, notons la $\mathbb{P}$, telle que $\mu(f) = \int f d\mathbb{P}$
Comme $\mu(1) = 1$, il est clair que $\mathbb{P}$ est en fait une probabilité.
De plus,

$$R^{2p} \geq \mu\left(X^{2p}\right) \geq (R+\delta)^{2p} \mathbb{P}([R+\delta, \infty[)$$

Ce qui implique, en faisant tendre $p$ vers l'infini, que $\mathbb{P}([R+\delta, \infty[) = 0$ pour tout $\delta$ positif et donc, finalement, que $\mathbb{P}$ est bien à support dans $[-R, R]$. □

## 6.3 Liberté

En probabilités classiques, on a la notion d'indépendance. On peut l'étendre au contexte libre.

**Définition 6.7.** *Soient $X_1, \ldots, X_n \in \mathcal{A}$. on dit que $X_1, \ldots, X_n$ sont indépdendantes si*

$$\tau(P_1(X_{i_1}) \ldots P_k(X_{i_k})) = \tau(P_1(X_{i_1})) \ldots P_k(X_{i_k})$$

*pour tout $k$ entier naturel, $P_1, \ldots, P_k$ polynômes et $i_{12} \neq \ldots \neq i_k$.*

Cependant, la notion la plus utile en théorie des probabilités libres n'est pas tant celle de l'indépdendance que celle de la liberté.

**Définition 6.8.** *On dit que* $(X_1, \ldots, X_m)$ *et* $(Y_1, \ldots, Y_n)$ *sont libres si pour tout* $N$ *entier et tous* $(P_i, Q_i)_{1 \leq i \leq N} \in (\mathbb{C} < X_1, \ldots, X_m > \times \mathbb{C} < X_1, \ldots, X_n >)^N$ *on a :*
*Si pour tout* $i$, $\tau(P_i(X_1, \ldots, X_m)) = 0$ *et* $\tau(Q_i(Y_1, \ldots, Y_n)) = 0$, *alors on a forcément :*

$$\tau\left(\prod_{1 \leq i \leq n}^{\rightarrow} P_i(X_1, \ldots, X_m) Q_i(Y_1, \ldots, Y_n)\right) = 0$$

On remarque que c'est analogue au cas de l'indépendance en probabilités classiques, dans la mesure où il suffit de connaitre les lois marginales pour connaitre la loi d'ensemble. En effet, supposons, pour simplifier, que $X$ et $Y$ soient libres. Alors :

$$\begin{aligned} \tau(XY) &= \tau((X - \tau(X))Y) + \tau(X)\tau(Y) \\ &= \tau(X - \tau(X))(Y - \tau(Y)) + \tau(X - \tau(X))\tau(Y) + \tau(X)\tau(Y) \end{aligned} \tag{11}$$

Le premier terme s'annule du fait de la liberté et les autres termes ne dépendent que des lois marginales et sont donc connus.
On peut cependant remarquer que, si cette notion est analogue à l'indépendance dans le cas classique, elle n'en est pas une transposition exacte, contrairement à la notion d'indépendance libre introduite plus haut. Pour souligner les différences, faisons deux remarques :

1. Si $X$ et $Y$ sont libres, $P, Q$ des polynômes et si $\tau(P(X)) = 0$ et $\tau(Q(Y)) = 0$, alors $\tau(P(X)Q(Y)P(X)Q(Y)) = 0$, alors que si $X$ et $Y$ étaient des variables aléatoires (au sens classique) indépendantes, alors

   $$\mathbb{E}(P(X)Q(Y)P(X)Q(Y)) = \mathbb{V}ar(P(X)Q(Y))$$

   qui n'a *a priori* aucune raison d'être nul.

## 6.4 Transformée-R

Dans le cas classique, on pouvait recourir à la transformée de Fourier (fonction caractŕistique) pour caractériser une loi. On a ici un équivalent sous la forme de la transformée-R. Nous devons d'abord introduire quelques notions :

**Définition 6.9.** *On définit*

1. *Soit* $H$ *un espace de Hilbert. On appelle espace de Fock complet de* $H$ *la chose suivante (*$\Omega$ *est un vecteur unitaire arbitraire) :*

   $$F(H) := \mathbb{C}\Omega \oplus \bigoplus_{i=1}^{\infty} H^{\otimes n}$$

2. *Soit $h$ dans $H$. On définit l'opérateur création $l_h^*$ :*

$$l_h^* : h_1 \otimes \ldots \otimes h_n \mapsto h \otimes h_1 \otimes \ldots \otimes h_n$$

*et l'opérateur annihilation $l_h$ :*

$$l_h : h_1 \otimes \ldots \otimes h_n \mapsto < h, h_1 > h_2 \otimes \ldots \otimes h_n$$

*et envoie $\Omega$ sur $0$.*
*En particulier, si les $E_i$ forment une base, on pose $l_i = l_{e_i}$.*

3. *On appelle $\mathcal{E}$ l'espace des sommes formelles du type suivant, avec $Q$ entier naturel et les coefficients dans $\mathbb{C}$ :*

$$\sum_{p \geq 0, O \leq q \leq Q} \sum_{i_1 \ldots i_p, j_1 \ldots j_q \in \{1 \ldots n\}} c_{i_1,\ldots,i_p;j_1,\ldots,j_q} l_{i_1} \ldots l_{i_p} l_{j_1}^* \ldots l_{j_q}^*$$

4. *Cet espace est muni d'une structure de $W^*$-espace de probabilité en définissant $\omega$ comme étant la forme linéaire envoyant un élément du type ci-dessus sur le coefficient constant $c_{\emptyset,\emptyset}$*

La proprieté suivante va nous permettre de définir la R-transformée :

**Théorème 6.5.** *Soit $\tau \in \mathcal{M}^1$, alors il existe une unique variable aléatoire $T$ (non-commutative) de la forme*

$$l_1^* + \sum_{k=0}^{\infty} \alpha_{k+1} l_1^k$$

*dont la loi est $\tau$.*

*Démonstration.* Prenons $T$ de la forme ci-dessus. On remarque que $l_k l_k^* = Id$ et par conséquent :

$$\omega\left(T^m\right) = \alpha_m + U\left(\alpha_1, \ldots, \alpha_{m-1}\right)$$

où $U$ est un polynôme qui ne nous intéresse pas. On voit donc que les moments de $\tau$ définissent de manière unique les coefficients $\alpha_i$. □

On définit alors la R-transformée de $\tau$ comme étant la série formelle $R_\tau := \sum_{i=0}^{\infty} \alpha_{i+1} x^i$. Le théorème prcédent montre que la R-transformée caractérise bien la loi. On admettra la résultat suivant :

**Théorème 6.6.** *Soit $\tau$ dans $\mathcal{M}^1$. On note $\tau_k$ son kième moment et on pose :*

$$G\left(\zeta\right) := \frac{1}{\zeta} + \sum_{i=1}^{\infty} \tau_i \zeta^{-i-1}$$

*et*

$$K\left(z\right) := \frac{1}{z} + R_\tau\left(z\right)$$

*Alors, $G$ et $K$ sont inverses l'une de l'autre au sens de la composition.*

## 6.5 Calcul stochastique libre

De même que dans le cas classique, on pouvait définir une notion de calcul stochastique, cela s'étend au cas libre.

**Définition 6.10.** *Soit $X$ un élément de notre $W^*$-espace de probabilité. On dit que $X$ suit une loi semi-circulaire de covariance $V$ si :*

$$\tau\left(P\left(X\right)\right) = \frac{1}{\pi V}\int_{-2\sqrt{V}}^{2\sqrt{V}} \sqrt{4V - x^2}dx$$

Les lois semi-circulaires jouent en probabilité libre un rôle analogue à celui joué dans le cas classique par les lois gaussiennes. En particulier, on a un analogue du théorème central limite : Si $X_n$ sont des éléments libres de $\mathcal{A}$ tels que $\tau\left(X_i\right) = 0$ et $\tau\left(X_i^2\right) = 1$, alors

$$\frac{X_1 + \ldots + X_n}{\sqrt{n}} \xrightarrow{loi} S$$

où $S$ est une loi semi-circulaire de variance 1.

**Définition 6.11.** *Soit $(\mathcal{A}, \tau)$ un $W^*$-espace de probabilité, et $S := (S_t)_{t\geq 0}$ une famille d'opérateurs dans $\mathcal{A}$. On dit que $S$ est un mouvement brownien libre si :*

1. *$S_0 = 0$*
2. *Pour tout $0 \leq s \leq t$, $S_t - S_s$ et $\sigma\left(S_u, u \leq s\right)$ sont libres, où $\sigma\left(S_u, u \leq s\right)$ désigne l'algèbre de von Neumann engendrée par les $S_u$*
3. *Pour tout $0 \leq s \leq t$, $S_t - S_s$, la loi de $S_t - S_s$ suit une loi semi-circulaire de covariance $s - t$*

On remarque qu'il s'agit d'une généralisation tout-à-fait naturelle du mouvement brownien habituel.
Nous allons à présent, à titre d'information, donner l'idée de la construction d'un tel mouvement brownien libre. Ceci est donné à titre informatif et ne joue pas de rôle dans la suite. Pour en comprendre l'idée, nous donnerons d'abord une construction d'un mouvement brownien classique très semblable à ce que nous ferons dans le cas libre.

### 6.5.1 Construction du mouvement brownien classique

Soit $H := L^2\left(\mathbb{R}^+, dx\right)$. C'est un espace séparable, qui admet donc une base de Hilbert $(e_i)_i$. Pour $f \in H$, nous noterons $f_i$ ses coordonnées dans cette base. On pose $\Omega := \prod_{i\in\mathbb{N}} \mathbb{R}$ muni de sa tribu borélienne pour la topologie produit usuelle. On muni cet ensemble d'une mesure gaussienne $\mathbb{P}$ (c'est-à-dire que les éléments de cet espace peuvent être vus comme une suite iid de gaussiennes centrées réduites).
On a alors une injection

$$\begin{aligned} \Psi : \quad & H \to L^2\left(\Omega, \mathbb{P}\right) \\ & h \mapsto \left\{ k \in \Omega \mapsto \sum_i k_i h_i \right\} \end{aligned}$$

On vérifie que :

$$\begin{aligned} ||\Psi(h)||^2 &= \mathbb{E}\left(\left(\sum_i X_i h_i\right)^2\right) \\ &= \mathbb{E}\left(\sum_i X_i^2 h_i^2 + \sum_{i\neq j} X_i X_j h_i h_j\right) \\ &= \sum_i h_i^2 \\ &= ||h||^2 \end{aligned}$$

Où l'on a utilisé le fait que les $X_i$ sont des gaussiennes centrées réduites iid. On vient donc à la fois de montrer que $\Psi(h)$ est bien dans $L^2(\Omega, \mathbb{P})$ et que $\Psi$ est un plongement isométrique.
On pose alors : $B_t := \Psi\left(\mathbb{1}_{[0,t]}\right)$ pour tout $t$ réel positif. Alors, $(B_t)_{t\geq 0}$ est un mouvement brownien standard.
En effet, $B_0 = 0$ est clair et de plus, pour $0 \leq s \leq t$ :

$$\begin{aligned} \mathbb{E}(B_t B_s) &= <\Psi\left(\mathbb{1}_{[0,t]}\right), \Psi\left(\mathbb{1}_{[0,s]}\right)> \\ &= <\mathbb{1}_{[0,t]}, \mathbb{1}_{[0,s]}> \\ &= s \end{aligned}$$

### 6.5.2 Construction du mouvement brownien libre

Reprenons notre espace de Fock complet :

$$F := \mathbb{C}\Omega \oplus \bigoplus_{n=1}^{\infty} L^2(\mathbb{R}_+)^{\otimes n}$$

On équippe de plus chacun des $L^2(\mathbb{R}_+)^{\otimes n}$ du produit scalaire suivant :

$$< f_1 \otimes \ldots \otimes f_n, g_1 \otimes \ldots \otimes g_n >_n := < f_1, g_1 > \ldots < f_n, g_n >$$

et on définit à partir de là le produit scalaire sur $F$. Si $f = \sum f_n$ et $g = \sum g_n$ sont des éléments de $F$ avec leurs décompositions canoniques, on pose :

$$< f, g > := \sum_{i=0}^{n} < f_i, g_i >_i$$

Cela définit bien un produit scalaire !
On pose à partir des opérateurs annihilation et création, pour $t \geq 0$ :

$$\begin{aligned} l_t &:= l_{\mathbb{1}_{[0,t]}} \\ l_t^* &:= l^*_{\mathbb{1}_{[0,t]}} \\ X_t &:= l_t + l_t^* \end{aligned}$$

Soit $\mathcal{W}$ l'algèbre engendrée par les $X_t, t \geq 0$, et $\tau$ l'état suivant :

$$\tau(T) := < T\Omega, \Omega > \text{ pour tout opérateur } T \text{ sur } F$$

On vérifie sans peine que $\tau$ est bien un état tracial.
On a en fait que $X_t$ est un mouvement brownien libre. Nous n'allons pas le démontrer mais remarquons que $X_0 = 0$.
Nous pouvons aussi montrer, par exemple, que $X_1$ est distribué selon une loi semi-circulaire standard. Il est clair que cette variable est auto-adjointe car $l_1$ est l'adjoint de $l_1^*$. Donc, par le théorème 6.1, $X_1$ a une loi qui correspond à une mesure $\mu$ à support compact sur $\mathbb{R}$. Or, la R-transformée de $X_1$ est clairement $R(z) = z$ et donc le théorème 6.6 nous dit que la transformée de Cauchy $G$ de $\mu$ est racine de $1/G + G = \zeta$. C'est-à-dire :

$$G(\zeta) = \frac{\zeta - \sqrt{\zeta^2 - 4}}{2}$$

Où l'on a utilisé le fait que $G$ tend vers 0 quand $\zeta$ tend vers l'infini pour choisir le bon signe
On peut alors vérifier que c'est là la transformée de Cauchy d'une loi semi-circulaire.

## 6.6 Probabilités libres et matrices aléatoires

### 6.6.1 Intérêt

Considérer le cadre des probabilités libres pour étudier les matrices aléatoires est assez pratique. On se souviendra par exemple que la loi décrivant la répartition asymptotique des valeurs propres, quand la taille tend vers l'infini, est la distribution semi-circulaire, cette même distribution qui joue un rôle central en probabilité libre, un rôle analogue aux gaussiennes dans le cas standard.
En fait, il y a là quelque chose de profond. Il existe par exemple un résultat de liberté asymptotique :
Soit, pour tout $n \in \mathbb{N}$, $\left\{X_{i,1}^{(n)}\right\}_i, \ldots \left\{X_{i,k}^{(n)}\right\}$ des familles de variables aléatoires (libres) dont la loi jointe est $\mu_n$. On dit que les familles de variables aléatoires sont asymptotiquement libres (quand $n$ tend vers l'infini) si :

- Il existe une loi $\mu$ qui soit limite des $\mu_n$.
- Des variables aléatoires $\{X_{i,1}\}_i, \ldots \{X_{i,k}\}_i$ de loi jointe $\mu$ sont libres.

Alors, sous ces définitions, si pour tout $n$, $M_{1,n}, M_{2,n}, \ldots$ sont des exemplaires indépendants de matrices du GUE, elles sont asymptotiquement libres.

### 6.6.2 Démonstration des résultats

Dans cette partie, on cherche à démontrer les résultats dont nous avions besoin dans le cas matrices de Wigner non-centrées et que nous n'avions pas encore pu démontrer.

**Théorème 6.7.** *Soient $p \in \mathcal{P}(\mathbb{R})$ et $\nu \in \mathcal{C}([0,1], \mathcal{P}(\mathbb{R}))$. On a :*

$$S^{0,1}(\nu \boxplus p) \leq S^{0,1}(\nu)$$

*Démonstration.* On va en fait donner l'idée de la preuve. Il faudrait en effet sinon introduire tout une théorie du calcul stochastique dans le cas libre, ce qui dépasserait le cadre de ce mémoire. On rappelle que l'on a :

$$S^{0,1}(\nu) = \frac{1}{2} < k, k >_{0,1}^{\nu} \tag{12}$$

On va considérer l'équation :

$$dX_t = dS_t + \partial_x k_t(X_t)\, dt \tag{13}$$

La loi de $X_t$ vérifie l'équation 12. On peut s'en convaincre par analogie avec le cas standard :

$$dZ_t = dB_t + k(Z_t)\, dt$$

donne, si on note $\kappa_t$ la loi de $Z_t$ :

$$\begin{aligned} f(X_1, 1) - f(X_0, 0) &= \int_0^1 \partial_1 f(X_s, s)\, dX_s + \\ &\quad \int_0^1 \partial_2 f(X_s, s)\, ds + \frac{1}{2} \int_0^1 \partial_1^2 f(X_S, s d < X >_s \end{aligned} \tag{15}$$

On peut alors en prendre l'espérance, en observant que :

$$\int_0^1 \partial_1 f(X_s, s)\, dX_s = \int_0^1 \partial_1 f(X_s, s)\, dB_s + \int_0^1 \partial_1 f(X_s, s)\, \partial_x k_s(X_s)\, ds$$

Le premier terme du membre de droite étant d'espérance nulle, on obtient bien finalement :

$$T^{0,1}(f, \kappa) = < f, k >_{\kappa}^{0,1}$$

La démonstration dans le cas de l'équation différentielle stochastique libre 13 se fait de le même façon mais en l'occurence avec une formule d'Itô libre.
On va supposer en outre que la solution de 12 est unique, de sorte que $X_t$ suit exactement $\nu_t$.
On prend alors $C$ de loi $p$ et libre avec les $X_t$. On pose $Y_t := X_t + C$, qui est donc de loi $\mu_t := \nu_t \boxplus p$ et qui vérifie l'équation :

$$dY_t = dS_t + \partial_x k_t(X_t)\, dt \tag{16}$$

À partir de là, on va pouvoir faire le calcul. On remarquera cependant que l'on fait appel à un équivalent libre de l'espérance conditionelle, $\tau(.|A)$ étant la

projection orthogonale dans $L^2(\tau)$ sur la sous-algèbre engendrée par $A$.

$$\begin{aligned} S^{0,1}(\mu, f) &= \int_0^1 \tau\left(\partial_x f_t(X_t + C)\, \partial_x k_t(X_t)\right) dt \\ &= \int_0^1 \tau\left(\partial_x f_t(X_t + C)\, \tau\left(\partial_x k_t(X_t) \,|X_t + C\right)\right) dt \end{aligned}$$

On obtient la première ligne en raisonnant comme avant. Là encore, l'analogie avec le cas classique permet de s'en convaincre.
bref, on en arrive à :

$$\begin{aligned} S^{0,1}(\mu) &= \sup_{f \in \mathcal{C}_b^{2,1}(\mathbb{R}\times[0,1])} \left[ S^{0,1}(\mu, f) - \frac{1}{2} < f, f >_{0,1}^{\mu} \right] \\ &= \sup_f \left[ \int_0^1 \tau\left(\partial_x f_t(X_t + C)\, \tau\left(\partial_x k_t(X_t) \,|X_t + C\right)\right) dt - \frac{1}{2} \int_0^1 \tau\left(\partial_x f_t(X_t + C)^2\right) dt \right] \end{aligned}$$

Or, le polynôme $\alpha X - \frac{1}{2}X^2$ a un maximum pour $X = \alpha$. Donc l'intégrande est maximal lorsque $\partial_x f_t(X_t + C) = \tau\left(\partial_x k_t(X_t) \,|X_t + C\right)$. D'où :

$$\begin{aligned} S^{0,1}(\mu) &\leq \frac{1}{2} \int_0^1 \tau\left(\tau\left(\partial_x k_t(X_t) \,|X_t + C\right)^2\right) dt \\ &\leq \frac{1}{2} \int_0^1 \tau\left(\partial_x k_t(X_t)^2\right) dt \\ &= S^{0,1}(\nu) \end{aligned}$$

La dernière inégalité provenant de l'inégalité de Jensen appliquée à l'espérance conditionelle. □

# 7 Conclusion

Nous venons de finir notre tour d'horizon de quelques proprietés liées aux grandes déviations de la mesure spectrale de matrices aléatoires. Nous avons en particulier constaté que l'on peut obtenir des résulats de grandes déviations sans trop de difficultés dans le cas de matrices de Wigner centrées. Le cas "non-centrées", quant à lui, nous a amené à introduire des concepts plus riches, en particulier à faire un tour du coté des probabilités libres.
De manière générale, la théorie des matrices aléatoires est un domaine d'étude qui suscite encore beaucoup d'intérêt à l'heure actuelle. Il existe des liens étonnants avec d'autres domaines des mathématiques : les probabilités libres, comme nous l'avons constaté, mais également, à titre d'exemple, les fonctions zeta de Riemann et la combinatoire.

# Références

[1] Philippe Biane and Roland Speicher. Stochastic calculus with respect to free brownian motion and wigner space, 1997.

[2] Mac Duffee. *Vectors and Matrices.*

[3] Peter Forrester. *Log-gases and Random matrices.*

[4] Alice Guionnet. Large deviations and stochastic calculus for large random matrices. *Probability Surveys*, 1, 2004.

[5] C. Landim. *Scaling limits of interacting particles systems.*

[6] Alexandru Nica and Roland Speicher. *Lectures on the Combinatorics of Free Probability.*

[7] Voiculescu, Dykema, and Nica. *Free Random Variables.* American Mathematical Society, 1992.