

# Matrices aléatoires et perturbations de rang fini.

Jean Rochet

Octobre 2012

## Table des matières

<b>1</b>	<b>Présentation des matrices aléatoires.</b>	<b>2</b>
1.1	Petit historique. . . . .	2
1.2	Modèles de matrices aléatoires. . . . .	2
1.3	Résultats importants. . . . .	2
1.4	Perturbation de rang fini . . . . .	4
<b>2</b>	<b>Perturbation de rang fini pour des matrices iid</b>	<b>4</b>
2.1	Résultat sur les valeurs propres extrêmes . . . . .	4
2.2	Idée de la démonstration . . . . .	4
<b>3</b>	<b>Perturbation de rang fini pour des matrices hermitiennes</b>	<b>5</b>
3.1	Inégalité d'entrelacement de Weyl. . . . .	7
3.2	Etude de la plus grand valeur propre . . . . .	8
<b>4</b>	<b>Et maintenant ?</b>	<b>12</b>
<b>5</b>	<b>Annexe</b>	<b>13</b>
5.1	Résultat sur les matrices . . . . .	13
5.2	Inégalité de concentration . . . . .	13

# 1 Présentation des matrices aléatoires.

## 1.1 Petit historique.

C'est dans les années 1930 qu'a été introduite pour la première fois la notion de matrice aléatoire par des statisticiens comme Wishart dans le cadre d'étude de matrices de covariances empiriques. Cependant, cette théorie a véritablement pris son envol dans les années 1950 lorsque des physiciens tels que Wigner ou Dyson s'y intéressèrent pour l'étude du spectre d'émission d'atomes lourds.

Aujourd'hui, les matrices aléatoires trouvent de nombreuses applications dans divers domaines tels que la théorie des nombres (notamment la conjecture de Riemann), la mécanique quantique, la mécanique statistique, l'étude des réseaux de télécommunication, la finance ...

## 1.2 Modèles de matrices aléatoires.

Il existe plusieurs manières de définir de l'aléa dans une matrice. On présente ici quelques modèles connus.

- **Matrices de Wigner :**

$M = (m_{i,j})_{1 \leq i,j \leq n}$  avec pour tout  $i \neq j$   $m_{i,j}$  est une variable gaussienne complexe centrée réduite et pour tout  $i$   $m_{i,i}$  est une gaussienne réelle centrée réduite. On suppose de plus que pour tout  $i$  et  $j$ ,  $m_{i,j} = \overline{m_{j,i}}$  et que les  $m_{i,j}$  sont indépendants pour tout  $i \leq j$ .

- **Matrices iid :**

$M = (m_{i,j})_{1 \leq i,j \leq n}$  tel que les  $m_{i,j}$  soient tous identiquement distribués et indépendants, centrés et de variance 1. On appellera **distribution atomique** la loi commune des entrées de la matrice.

- **Matrices aléatoires dont on connaît les valeurs singulières :**

Soit  $s_1, s_2, \dots, s_n$   $n$  réels positifs (déterministes ou non). On considère la matrice aléatoire

$$M = U \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix} V$$

avec  $U$  et  $V$  deux matrices indépendantes (entre elles et de  $(s_i)$ ), distribuées selon la mesure de Haar sur l'ensemble des matrices unitaires. La matrice obtenue a pour valeurs singulières les  $(s_i)$  (on rappelle que les valeurs singulières d'une matrice  $M$  peuvent être définies comme étant les valeurs propres que  $\sqrt{MM^*}$ ).

## 1.3 Résultats importants.

Dans l'étude des matrices aléatoires, on s'intéresse plus particulièrement aux valeurs propres de la matrice, ou plutôt, à leur distribution spectrale empirique (DSE)

$$\mu_n := \frac{1}{n} \sum_{k=1}^n \delta_{\lambda_k}.$$

Ainsi, en considérant une suite de matrices aléatoires (dont la taille augmente), les questions naturelles que l'on se pose sont : est-ce que la suite des DSE converge ? vers quelle mesure ? (Il faut bien entendu renormaliser par une puissance de  $n$  adaptée pour ne pas obtenir une limite infini ou nulle).

On donne ici des résultats sur ces questions.

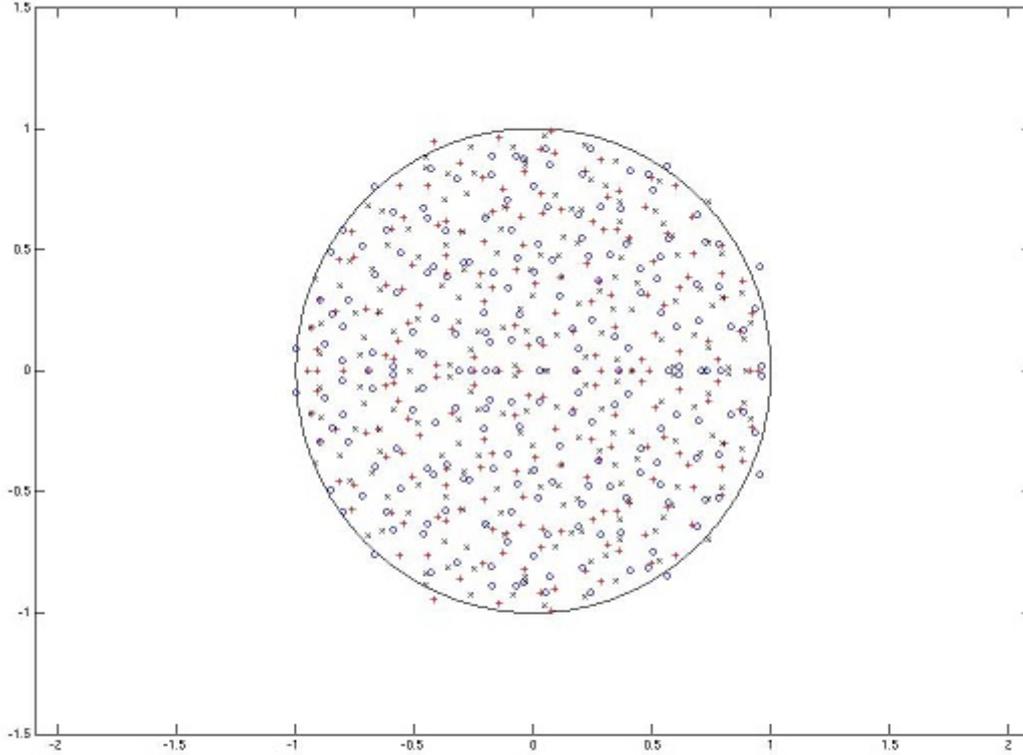


FIGURE 1 – Représentation de  $\mu_n$  pour trois matrices iid de taille  $200 \times 200$

**Théorème 1** (Wigner 1955). Soit  $M_n$  une suite de matrices de Wigner et on considère  $\mu_n$  la DSE de  $\frac{M_n}{\sqrt{n}}$ . Alors, la suite  $\mu_n$  converge faiblement presque sûrement vers la loi de  $[-2, 2]$  et de densité

$$\frac{1}{2\pi} \sqrt{4 - x^2} \mathbf{1}_{x \in [-2, 2]} dx,$$

appelée loi du demi-cercle (de rayon 2).

On donne équivalent de ce théorème pour les matrices iid.

**Théorème 2** ([6]). Soit  $M_n$  une suite de matrices iid où et on considère  $\mu_n$  la DSE de  $\frac{M_n}{\sqrt{n}}$ . Alors, la suite  $\mu_n$  converge faiblement presque sûrement vers la mesure complexe à support dans le disque unité de densité

$$\frac{1}{\pi} \mathbf{1}_{|z| \leq 1} dz$$

appelée loi circulaire.

La figure 1 est une illustration de ce dernier théorème : on a représenté les spectres de trois matrices iid dont la distribution atomique est respectivement une loi de Bernoulli ( $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$ ), une loi normale centrée réduite et une loi uniforme sur le segment  $[-\sqrt{3}, \sqrt{3}]$ . On observe que leurs valeurs propres sont presque toutes déjà localisées dans le disque unité.

**Remarque 1.1.** On pourra se référer au *Single Ring Theorem* de [3] pour obtenir des résultats dans le cas de matrices "dont on connaît les valeurs singulières". Ce théorème stipule entre autre que le support de la mesure limite de la suite des DSE est un tore (et non plus un disque).

## 1.4 Perturbation de rang fini

Nous allons maintenant nous intéresser aux problèmes de matrices perturbées.

On considère toujours une suite de matrices aléatoires  $X_n$  ainsi qu'une suite de matrices  $P_n$  (déterministe ou non) dont le rang est borné (indépendamment de  $n$ ). On s'intéresse alors au spectre de  $X_n + P_n$  (ce qu'on appelle une *perturbation de rang fini* de  $X_n$ ).

Dans le cas trivial où  $X_n$  et  $P_n$  commutent, on sait alors (en numérotant convenablement leurs valeurs propres respectives) que

$$\lambda_i(X_n + P_n) = \lambda_i(X_n) + \lambda_i(P_n).$$

Dans le cas général, cette question est évidemment plus complexe. En effet, comme nous verrons dans plusieurs situations, la distribution des valeurs propres continue à tendre vers la même limite comme on peut s'y attendre, l'étude de celle-ci ne nous apporte donc pas d'informations. Ainsi, ce qui va nous intéresser est plutôt le comportement asymptotique des valeurs propres dites *extrêmes* (le sup ou l'inf par exemple), ce qui est plus délicat.

## 2 Perturbation de rang fini pour des matrices iid

### 2.1 Résultat sur les valeurs propres extrêmes

On donne ici un résultat sur les valeurs propres extrêmes de certaines matrices iid pour des perturbations de rang fini.

**Théorème 2.1.** *Soit  $X_n$  une suite de matrices aléatoires iid dont la distribution atomique a des moments d'ordre quatre finies, et pour tout  $n$ , soit  $P_n$  une matrice déterministe dont le rang et la norme d'opérateur sont des  $O(1)$ . Soit  $\varepsilon > 0$ , et l'on suppose que pour  $n$  suffisamment grand,  $P_n$  ne possède pas de valeurs propres dans le tore  $\{z \in \mathbb{C}, 1 + \varepsilon < |z| < 1 + 3\varepsilon\}$  et qu'il y a  $j$  valeurs propres  $\lambda_1(P_n), \lambda_2(P_n), \dots, \lambda_j(P_n)$  (avec  $j = O(1)$ ) dont le module est plus grand ou égal à  $1 + 3\varepsilon$ . Alors, presque sûrement, pour  $n$  suffisamment grand, il y a exactement  $j$  valeurs propres de  $\frac{1}{\sqrt{n}}X_n + P_n$   $\lambda_1(\frac{1}{\sqrt{n}}X_n + P_n), \dots, \lambda_j(\frac{1}{\sqrt{n}}X_n + P_n)$  dont le module est plus grand ou égale à  $1 + 2\varepsilon$ , et en numérotant de manière appropriée les valeurs propres, on a même*

$$\lambda_i(\frac{1}{\sqrt{n}}X_n + P_n) = \lambda_i(P_n) + o(1), \quad \text{pour tout } i \in \{1, \dots, j\}.$$

Le comportement asymptotique des valeurs propres extrêmes s'illustrent très bien. On peut observer sur la figure 2 les valeurs propres d'une matrices  $500 \times 500$  iid (de distribution atomique  $x$  tel que  $\mathbb{P}(x = 1) = \mathbb{P}(x = -1) = \frac{1}{2}$ ) perturbée par une matrice diagonale  $P_n = \text{diag}(2, 2 + i, 2 - i, 0 \dots, 0)$ .

### 2.2 Idée de la démonstration

Une première idée essentielle est de remarquer la chose suivante : Si on considère deux matrices  $A \in \mathcal{M}_{n,k}(\mathbb{C})$  et  $B \in \mathcal{M}_{k,n}(\mathbb{C})$ , alors

$$\det(I_n + AB) = \det(I_k + BA).$$

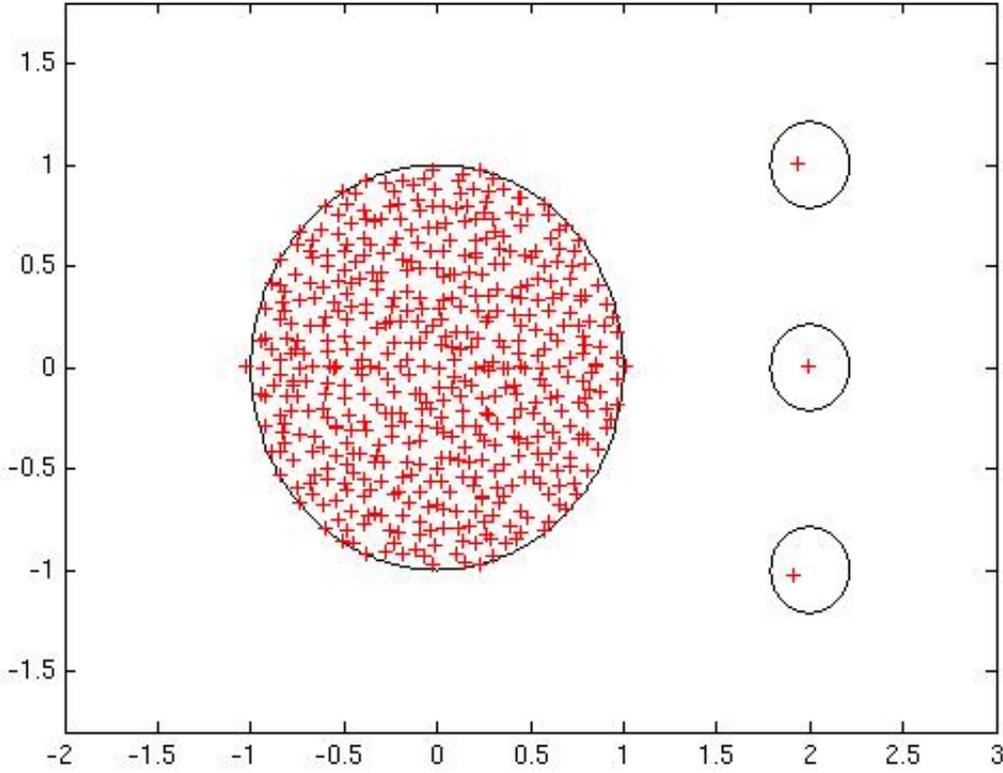


FIGURE 2 – Représentation des valeurs propres de  $X_n + P_n$  avec  $n = 500$  et  $P_n = \text{diag}(2, 2 + i, 2 - i, 0, \dots, 0)$

(ceci est démontré en annexe).

Ainsi, comme  $P_n$  est de rang fini (disons de rang  $k$ ), on peut l'écrire comme produit de deux matrices  $P_n = A_n \cdot B_n$  avec  $A_n \in \mathcal{M}_{n,k}(\mathbb{C})$  et  $B_n \in \mathcal{M}_{k,n}(\mathbb{C})$ , et donc

$$\det\left(\frac{1}{\sqrt{n}}X_n + P_n - zI\right) = \det\left(\frac{1}{\sqrt{n}}X_n - zI\right) \det\left(I + \left(\frac{1}{\sqrt{n}}X_n - zI\right)^{-1}A_n B_n\right) \quad (1)$$

$$= \det\left(\frac{1}{\sqrt{n}}X_n - zI\right) \det\left(I + B_n\left(\frac{1}{\sqrt{n}}X_n - zI\right)^{-1}A_n\right). \quad (2)$$

Ainsi, pour étudier le spectre de  $1/\sqrt{n} \cdot X_n + P_n$  (de taille  $n \times n$ ), on se ramène à étudier un déterminant d'une matrice  $k \times k$ , que nous noterons

$$f(z) := \det\left(I + B_n\left(\frac{1}{\sqrt{n}}X_n - zI\right)^{-1}A_n\right)$$

qui est une fonction méromorphe dont il n'est pas difficile de voir que ses zéros sont les valeurs propres de  $1/\sqrt{n} \cdot X_n + P_n$  et ses pôles sont les valeurs propres de  $1/\sqrt{n}X_n$ .

Une autre idée essentielle est alors de poser une autre fonction méromorphe

$$g(z) := \det\left(I + \frac{P_n}{z}\right) = \det\left(I + B_n(-zI)^{-1}A_n\right),$$

et de montrer que

$$\sup_{|z| \geq 1+2\varepsilon} |f(z) - g(z)| \xrightarrow{\text{p.s.}} 0$$

pour pouvoir d'une part appliquer le théorème de Rouché et conclure sur le nombre de valeurs propres de  $1/\sqrt{n}X_n + P_n$  sur  $\{|z| \geq 1 + 2\varepsilon\}$  et pour pouvoir d'autre part montrer que ces valeurs propres-là tendent vers la même limite que celles de  $P_n$ . On pourra se référer à [6] pour une preuve plus détaillée.

### 3 Perturbation de rang fini pour des matrices hermitiennes

On se donne dans cette section une suite de matrices  $n \times n$  aléatoires  $X_n$  hermitiennes, dont on note les valeurs propres par  $\lambda_1(X_n) \geq \lambda_2(X_n) \geq \dots \geq \lambda_n(X_n)$ . On notera  $\mu_{X_n}$  la distribution spectrale empirique de  $X_n$ .

On fait l'hypothèse que  $\mu_{X_n}$  converge presque sûrement faiblement vers une mesure de probabilité déterministe  $\mu_X$  à support compact (on notera le support  $[a, b]$ ). On suppose aussi que les valeurs propres extrêmes de  $X_n$  convergent respectivement vers les extrémités du compact, i.e.

$$\begin{cases} \lambda_1(X_n) & \xrightarrow{\text{p.s.}} b \\ \lambda_n(X_n) & \xrightarrow{\text{p.s.}} a \end{cases}$$

Pour un entier  $r \geq 1$  fixé, on se donne  $\theta_1 \geq \theta_2 \geq \dots \geq \theta_r$   $r$  réels déterministes non nuls qui ne dépendent pas de  $n$ . Pour tout  $n$  ( $\geq r$ ), on se donne une matrice aléatoire hermitienne  $P_n$  de rang  $r$  et dont les  $r$  valeurs propres non nulles sont les  $(\theta_i)_{1 \leq i \leq r}$ . On notera  $s$  l'indice tel que  $\theta_1 \geq \dots \geq \theta_s > 0 > \theta_{s+1} \geq \dots \geq \theta_r$ .

On suppose enfin que soit  $X_n$ , soit  $P_n$  est invariante en loi par conjugaison par une matrice unitaire.

On notera pour la suite pour toute fonction  $f$

$$f(c^+) := \lim_{x \downarrow c} f(x); \quad f(c^-) := \lim_{x \uparrow c} f(x)$$

On considère pour la suite la matrice aléatoire  $\tilde{X}_n = X_n + P_n$  qui est une perturbation additive de rang  $r$  de  $X_n$ .

**Théorème 3.1.** *Les valeurs propres de  $\tilde{X}_n$  ont le comportement à l'infini suivant :*

$$\bullet \forall 1 \leq i \leq s, \quad \lambda_i(\tilde{X}_n) \xrightarrow{\text{p.s.}} \begin{cases} G_{\mu_X}^{-1}(1/\theta_i) & \text{si } \theta_i > 1/G_{\mu_X}(b^+), \\ b & \text{sinon,} \end{cases}$$

alors que pour tout  $i > s$ ,  $\lambda_i(\tilde{X}_n) \xrightarrow{\text{p.s.}} b$ .

$$\bullet \forall 0 \leq j < r - s, \quad \lambda_j(\tilde{X}_n) \xrightarrow{\text{p.s.}} \begin{cases} G_{\mu_X}^{-1}(1/\theta_{r-j}) & \text{si } \theta_j < 1/G_{\mu_X}(a^-), \\ a & \text{sinon,} \end{cases}$$

alors que pour tout  $j \geq r - s$ ,  $\lambda_{n-j}(\tilde{X}_n) \xrightarrow{\text{p.s.}} a$ .

On rappelle que

$$G_{\mu_X}(z) = \int \frac{d\mu_X(t)}{z - t} \quad \text{pour } z \notin \text{supp } \mu_X,$$

désigne la transformée de Stieltjes de  $\mu_X$ , et  $G^{-1}(\cdot)$  est sa fonction réciproque.

**Remarque 3.1.** Il est important de remarquer que  $G_{\mu_X}$  vérifie :

$$G_{\mu_X}(z) \in \mathbb{R} \Leftrightarrow z \in \mathbb{R},$$

En effet,

$$\Im(G_{\mu_X}(z)) = \Im(z) \cdot \int_a^b \frac{1}{(\Re(z) - t)^2 + \Im(z)^2} d\mu_X(t).$$

On peut donc voir  $G_{\mu_X}$  comme une fonction réelle. De plus, on remarque que pour tout  $x \in \mathbb{R} - [a, b]$ ,  $G'_{\mu_X}(x) < 0$ , en effet

$$G'_{\mu_X}(x) = \int_a^b \frac{-1}{(x-t)^2} d\mu_X(t) < 0.$$

On peut donc dire que  $G_{\mu_X}$  décroît de 0 à  $G_{\mu_X}(a^-)$  sur  $] -\infty, a[$  et décroît de  $G_{\mu_X}(b^+)$  à 0 sur  $]b, \infty[$ . Ainsi, si  $x \notin ]G_{\mu_X}(a^-), G_{\mu_X}(b^+)[$ , alors  $G^{-1}(y) = x$  n'a pas de solution.

Nous allons faire la démonstration dans le cas  $r = 1$  qui est bien entendu plus simple mais qui permet de bien comprendre l'idée générale de la preuve.

### 3.1 Inégalité d'entrelacement de Weyl.

Dans le cas hermitien, nous pouvons utiliser le lemme suivant :

**Lemme 3.2** (Entrelacement de Weyl [4]). *Soit  $A$  et  $B$  deux matrices hermitiennes. Alors, pour tout  $k \in \{1, \dots, n\}$ , on a*

$$\lambda_k(A) + \lambda_n(B) \leq \lambda_k(A+B) \leq \lambda_k(A) + \lambda_1(B) \quad (3)$$

De plus, si  $B$  est de rang 1 et est positive, on a

$$\lambda_k(A) \leq \lambda_k(A+B) \leq \lambda_{k-1}(A) \quad (4)$$

si  $B$  est négative, on a plutôt

$$\lambda_{k+1}(A) \leq \lambda_k(A+B) \leq \lambda_k(A).$$

où l'on utilise la convention  $\lambda_0(M) = +\infty$  et  $\lambda_{n+1}(M) = -\infty$  pour toute matrice  $M \in \mathcal{M}_n(\mathbb{C})$ .

**Remarque 3.2.** Il existe une version plus générale de ce lemme (qu'on peut aussi trouver dans [4]) dans le cas où  $B$  est de rang  $p \leq n$  avec  $s$  valeurs propres strictement positives et  $p-s$  valeurs propres strictement négatives. On a alors

$$\lambda_{k+(p-s)}(A) \leq \lambda_k(A+B) \leq \lambda_{k-s}(A).$$

Ce lemme nous permet alors facilement de conclure par encadrement que  $\mu_{X_n+P_n} \rightarrow \mu_X$ .

On peut alors commencer la preuve du théorème 3.1.

**Preuve :** (du théorème 3.1 (début))

On traite le cas où  $\text{tr}(P_n) = \theta > 0$ , i.e.  $P_n$  positive (dans le cas contraire, le raisonnement est identique). Soit  $i > 1$ , on a par le lemme 3.2 que

$$\lambda_i(X_n) \leq \lambda_i(\tilde{X}_n) \leq \lambda_{i-1}(X_n) \leq \lambda_1(X_n) \quad (5)$$

ce qui implique par hypothèse que

$$\limsup_{n \rightarrow \infty} \lambda_i(\tilde{X}_n) \leq \lim_{n \rightarrow \infty} \lambda_1(X_n) = b.$$

De plus, montrons que pour tout  $\varepsilon > 0$ ,

$$\liminf_{n \rightarrow \infty} \lambda_i(X_n) \geq b - \varepsilon. \quad (6)$$

En effet, pour  $\varepsilon > 0$  et pour tout  $n$  suffisamment grand, on sait qu'on peut trouver plus de  $i+1$  valeurs propres de  $X_n$  dans l'intervalle  $]b - \varepsilon, b]$ , ce qui implique forcément que la  $i$ -ième plus grande

des valeurs propres est supérieur à  $b - \varepsilon$ , d'où (6).

On en déduit alors, grâce à (5) que

$$\lambda_i(\tilde{X}_n) \xrightarrow{\text{p.s.}} b.$$

**N.B. :** En fait, on a montré que pour tout  $i$ , on a

$$\liminf \lambda_i(\tilde{X}_n) \geq b \quad \text{et} \quad \limsup \lambda_{n-i}(\tilde{X}_n) \leq a$$

□

### 3.2 Etude de la plus grand valeur propre

Comme on a supposé  $r = 1$ , on peut écrire  $P_n = \theta u_n u_n^*$  avec  $u_n$  un vecteur colonne unitaire. Comme on a supposé que soit  $X_n$  soit  $P_n$  était invariant en loi par conjugaison par une matrice unitaire, on peut supposer sans perte de généralité que  $X_n = \text{diag}(\lambda_1, \dots, \lambda_n)$  et que  $u_n$  est distribué uniformément sur la sphère de dimension  $n$ . En effet, on sait que,  $X_n$  étant hermitienne, on peut trouver  $V_n$  matrice unitaire telle que

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

$X_n + P_n$  a les mêmes valeurs propres que  $\text{diag}(\lambda_1, \dots, \lambda_n) + \theta V_n u_n (V_n u_n)^*$ . Or  $V_n u_n$  est uniformément distribué sur la sphère de dimension  $n$  que ce soit parce que  $V_n$  est distribué selon la mesure de Haar (sur l'ensemble des matrices unitaires) ou parce que  $u_n$  est distribué uniformément sur la sphère. On s'est donc ramené à étudier le cas

$$X_n + P_n = \text{diag}(\lambda_1, \dots, \lambda_n) + \theta u_n u_n^*$$

où  $u_n$  est uniformément distribué sur la sphère de dimension  $n$ .

On écrit (c.f. (1))

$$\det((X_n + P_n) - zI) = \det(X_n - zI) \cdot \det(I + (X_n - zI)^{-1} P_n). \quad (7)$$

On a alors que pour tout  $z$  qui n'est pas une valeur propre de  $X_n$ ,

$$z \in \text{Sp}(X_n + P_n) \Leftrightarrow \det(I + (X_n - zI)^{-1} P_n) = 0,$$

or

$$\begin{aligned} \det(I + (X_n - zI)^{-1} P_n) &= \det(I + (X_n - zI)^{-1} \theta u_n^* u_n) \\ &= \det(1 + \theta u_n (\text{diag}(\lambda_1, \dots, \lambda_n) - zI)^{-1} u_n^*) \\ &= 1 + \theta u_n (\text{diag}(\frac{1}{\lambda_1 - z}, \dots, \frac{1}{\lambda_n - z})) u_n^*. \end{aligned}$$

On se retrouve donc avec l'égalité suivante

$$\begin{aligned} 1 + \theta u_n (\text{diag}(\frac{1}{\lambda_1 - z}, \dots, \frac{1}{\lambda_n - z})) u_n^* &= 0 \\ \Leftrightarrow \sum_{k=1}^n \frac{1}{z - \lambda_k} |u_{n,k}|^2 &= \frac{1}{\theta} \\ \Leftrightarrow G_{\mu_n}(z) &= \frac{1}{\theta} \end{aligned}$$

où  $\mu_n$  est la mesure spectrale de  $X_n$  "pondéré" par  $u_n$

$$\mu_n := \sum_{k=1}^n |u_{n,k}|^2 \delta_{\lambda_k}.$$

Ainsi, pour conclure, il nous faut montrer que  $\mu_n$  converge faiblement presque sûrement vers  $\mu_X$ . Nous commençons par montrer le lemme suivant.

**Lemme 3.3.** Soit  $u_n = (u_{1,n}, \dots, u_{n,n})$  une suite de vecteurs unitaires distribués uniformément sur la sphère unité de  $\mathbb{C}^n$ . Soit  $(x_n)_n$  une suite de variables aléatoires réelles indépendantes de  $(u_n)_n$ . On suppose que  $\sup_n |x_n| < \infty$  et que

$$\frac{1}{n}(x_1 + x_2 + \dots + x_n) \xrightarrow{\text{p.s.}} \alpha \in \mathbb{R}.$$

Alors,

$$S_n := \sum_{k=1}^n |u_{k,n}|^2 x_n \xrightarrow{\text{p.s.}} \alpha.$$

**Preuve :**

On définit  $\mathcal{F} := \sigma(x_n; n \geq 1)$  (la tribu engendrée par la famille de variables aléatoires  $(x_n)_n$ ). Comme  $u_n$  est indépendante de  $\mathcal{F}$ , on a

$$\mathbb{P}(S_n \rightarrow \alpha) = \mathbb{E} \left[ \mathbb{E} [\mathbf{1}_{S_n \rightarrow \alpha} | \mathcal{F}] \right].$$

On peut donc supposer que la suite  $(x_n)_n$  est déterministe. On conclut en utilisant le lemme 5.2 et la remarque 5.1.  $\square$

**Lemme 3.4.** Soit  $u_n = (u_{1,n}, \dots, u_{n,n})$  une suite de vecteurs unitaires distribués uniformément sur la sphère unité de  $\mathbb{C}^n$ . Soit  $\lambda^{(n)} = (\lambda_1^{(n)}, \dots, \lambda_n^{(n)})$  une famille de réels indépendants de  $(u_n)$  et tel que presque surement  $\sup_{n,k} |\lambda_k^{(n)}| < \infty$ . On suppose aussi qu'il existe une mesure de probabilité déterministe  $\mu$  tel que

$$\frac{1}{n} \sum_{k=1}^n \delta_{\lambda_k^{(n)}} \text{ converge faiblement vers } \mu \text{ p.s.}$$

Alors

$$\mu_n := \sum_{k=1}^n |u_{k,n}|^2 \delta_{\lambda_k^{(n)}} \text{ converge faiblement vers } \mu \text{ p.s.}$$

**Preuve :**

Tout d'abord, comme  $\sup_{n,k} |\lambda_k^{(n)}| < \infty$ , la suite  $\mu_n$  est tendue.

Soit  $f$  une fonction continue et qui tend vers 0 en  $\pm\infty$ , alors par hypothèse on sait que (comme  $f$  est bornée)

$$\mu_n(f) \xrightarrow{\text{p.s.}} \mu(f).$$

Or l'ensemble des fonctions continues qui tendent vers 0 en  $\pm\infty$  est séparable pour la topologie de la convergence uniforme (ce qui n'est pas le cas des fonctions continues simplement bornées), on peut donc considérer une famille de fonctions  $(f_k)_{k \in \mathbb{N}}$  dense. Ainsi, pour tout  $k$ , il existe  $\Omega_k$ , un espace de mesure 1 tel que

$$\forall \omega \in \Omega_k, \quad \mu_n(\omega)(f_k) \rightarrow \mu(f_k).$$

ainsi, comme  $\bigcap_{k \geq 0} \Omega_k$  est aussi de mesure 1, on a

$$\text{p.s. } \forall k \geq 0, \quad \mu_n(f_k) \rightarrow \mu(f_k).$$

Par densité des  $(f_k)_k$ , on montre facilement que pour tout  $f$  continue qui tend vers 0 en  $\pm\infty$

$$\text{p.s. } \forall k \geq 0, \quad \mu_n(f) \rightarrow \mu(f).$$

En effet, pour tout  $\varepsilon > 0$ , il existe  $k \geq 1$  tel que  $\|f - f_k\|_\infty \leq \varepsilon$ , et  $N > 0$  tel que pour  $n \geq N$ ,  $|\mu_n(f_k) - \mu(f_k)| < \varepsilon$  donc

$$\begin{aligned} |\mu_n(f) - \mu(f)| &\leq |\mu_n(f - f_k)| + |\mu_n(f_k) - \mu(f_k)| + |\mu(f - f_k)| \\ &\leq 3\varepsilon. \end{aligned}$$

Soit maintenant  $f$  une fonction continue bornée. On se donne aussi une fonction  $g$  continue et qui tend vers 0 en  $\pm\infty$ , alors on a

$$\limsup_{n \rightarrow \infty} |\mu(f) - \mu_n(f)| \leq \limsup_{n \rightarrow \infty} (|\mu|(f(1-g)) + |\mu_n|(f(1-g)))$$

car  $fg$  est une fonction continue dont les limites sont nulles. Comme la suite  $\mu_n$  est tendue, il existe un compact  $K$  tel que  $\mu(K^c) + \mu_n(K^c) \leq \varepsilon$  donc en prenant  $g$  égale à 1 sur  $K$ , on a

$$\limsup_{n \rightarrow \infty} |\mu(f) - \mu_n(f)| \leq \varepsilon.$$

On en déduit que  $\mu_n$  converge faiblement vers  $\mu$  presque sûrement.  $\square$

Ainsi, comme on vient de voir que  $\mu_n$  converge faiblement vers  $\mu$  presque sûrement, on a

$$G_{\mu_n}(z) \rightarrow G_\mu(z) \quad \forall z \in \mathbb{C} - [a, b],$$

on a en fait la convergence uniforme en  $z$  sur  $C_\eta := \{z \in \mathbb{C}, d(z, [a, b]) \geq \eta\}$  pour tout  $\eta > 0$ . En effet, la famille de fonctions

$$\left( f_z := \begin{cases} [a, b] & \rightarrow \mathbb{C} \\ x & \mapsto \frac{1}{z-x} \end{cases} \right)_{z \in C_\eta}$$

est équi-Lipschitzienne. On peut donc appliquer le théorème d'Ascoli. On peut faire le même raisonnement pour la famille de fonctions

$$\left( g_z := \begin{cases} [a, b] & \rightarrow \mathbb{C} \\ x & \mapsto \frac{1}{(z-x)^2} \end{cases} \right)_{z \in C_\eta}$$

ce qui permet de dire que

$$G'_{\mu_n}(z) \rightarrow G'_\mu(z) \text{ uniformément pour tout } z \in \mathbb{C} - [a, b].$$

On peut maintenant terminer la démonstration du théorème dans le cas  $r = 1$ .

**Preuve :**(du théorème 3.1 (suite et fin))

On traitera le cas  $P_n > 0$  (le cas  $< 0$  étant similaire). On a déjà vu que

$$\liminf \lambda_1(\tilde{X}_n) \geq b.$$

• Maintenant, supposons que  $\limsup \lambda_1(\tilde{X}_n) = c > b$ , alors on considère  $\phi$  une injection croissante de  $\mathbb{N}$  telle que  $\lim \lambda_1(\tilde{X}_{\phi(n)}) = c$ , alors comme la convergence est uniforme sur  $C_{\frac{c-b}{2}}$  et que l'on sait que pour  $n$  suffisamment grand  $d(\lambda_1(\tilde{X}_{\phi(n)}), b) > \frac{c-b}{2}$ , on a

$$\frac{1}{\theta} = G_{\mu_{\phi(n)}}(\lambda_1(\tilde{X}_{\phi(n)})) \rightarrow G_{\mu_X}(c) < G_{\mu_X}(b^+)$$

cela implique donc que  $1/\theta < G_{\mu_X}(b^+)$ . Pour contraposé, on a donc montré que

$$\frac{1}{\theta} \geq G_{\mu_X}(b^+) \Rightarrow \limsup \lambda_1(\tilde{X}_n) \leq b.$$

• Maintenant, supposons que  $\theta > 1/G_{\mu_X}(b^+)$ . On a donc

$$G_{\mu_X}^{-1}\left(\frac{1}{\theta}\right) > b.$$

Soit  $\varepsilon > 0$  arbitraire tel que  $b + \varepsilon < G_{\mu_X}^{-1}\left(\frac{1}{\theta}\right)$ . On considère  $\gamma$  le contour du cercle du plan complexe de diamètre  $[G_{\mu_X}^{-1}\left(\frac{1}{\theta}\right) - \varepsilon, G_{\mu_X}^{-1}\left(\frac{1}{\theta}\right) + \varepsilon]$ , on a alors

$$1 = \frac{1}{2i\pi} \int_{\gamma} \frac{-\theta G'_{\mu_X}(z)}{1 - \theta G_{\mu_X}(z)} dz = \lim_{n \rightarrow \infty} \frac{1}{2i\pi} \int_{\gamma} \frac{-\theta G'_{\mu_n}(z)}{1 - \theta G_{\mu_n}(z)} dz$$

par la convergence uniforme de  $G'_{\mu_n}$  et de  $G_{\mu_X}$  sur  $B_{\mathbb{C}}(G_{\mu_X}^{-1}\left(\frac{1}{\theta}\right), \varepsilon)$ . Or

$$\frac{1}{2i\pi} \int_{\gamma} \frac{-\theta G'_{\mu_n}(z)}{1 - \theta G_{\mu_n}(z)} dz \in \mathbb{Z}$$

donc il existe un rang  $N$  à partir duquel pour tout  $n \geq N$ , on a

$$\frac{1}{2i\pi} \int_{\gamma} \frac{-\theta G'_{\mu_n}(z)}{1 - \theta G_{\mu_n}(z)} dz = 1$$

autrement dit, comme on sait que la seule valeur hors  $[\lambda_n(\tilde{X}_n), \lambda_1(\tilde{X}_n)]$  qui annule  $1 - \theta G_{\mu_n}(z)$  est  $\lambda_1(\tilde{X}_n)$ , par conséquent,

$$\left| \lambda_1(\tilde{X}_n) - G_{\mu_X}^{-1}\left(\frac{1}{\theta}\right) \right| < \varepsilon.$$

On en conclut que

$$\theta > 1/G_{\mu_X}(b^+) \Rightarrow \lambda_1(\tilde{X}_n) \xrightarrow{\text{p.s.}} G_{\mu_X}^{-1}\left(\frac{1}{\theta}\right).$$

□

## 4 Et maintenant ?

Mon travail de recherche va maintenant consister à étudier les valeurs propres extrêmes de matrices "dont on connaît les valeurs singulières" et de trouver des résultats analogues.

## 5 Annexe

### 5.1 Résultat sur les matrices

**Lemme 5.1.** *Soit  $A \in \mathcal{M}_{n,k}(\mathbb{C})$  et  $B \in \mathcal{M}_{k,n}(\mathbb{C})$ . Alors*

$$\det(I_d + AB) = \det(I_d + BA).$$

**Preuve :**

On introduit les matrices de  $\mathcal{M}_{n+k}(\mathbb{C})$  suivantes

$$\begin{pmatrix} I & A \\ -B & I \end{pmatrix} \quad \begin{pmatrix} I & 0 \\ B & I \end{pmatrix}$$

On a alors

$$\begin{pmatrix} I & 0 \\ B & I \end{pmatrix} \begin{pmatrix} I & A \\ -B & I \end{pmatrix} = \begin{pmatrix} I & A \\ 0 & BA + I \end{pmatrix}$$

et

$$\begin{pmatrix} I & A \\ -B & I \end{pmatrix} \begin{pmatrix} I & 0 \\ B & I \end{pmatrix} = \begin{pmatrix} I + AB & A \\ 0 & I \end{pmatrix}$$

le déterminant de ces deux produits étant le même, on peut conclure.  $\square$

### 5.2 Inégalité de concentration

Nous donnons ici une inégalité de concentration dont on s'est servi (on se pourra se référer au théorème 2.3 et à la proposition 1.8 de [5]).

**Lemme 5.2.** *Pour tout  $n \geq 1$ , on se donne  $(u_1, \dots, u_n)$ , la première colonne d'une matrice orthogonale distribuée selon la mesure de Haar. Alors, il existe deux constantes  $C$  et  $c$  telles que pour tout  $n$  suffisamment grand et pour tout fonction  $f_n$  1-Lipschitzienne sur la sphère unité de  $\mathbb{R}^n$ , et pour tout  $\varepsilon > 0$ ,*

$$\mathbb{P} \left( |f_n(u_1, \dots, u_n) - \mathbb{E} [f_n(u_1, \dots, u_n)]| \geq \varepsilon \right) \leq C e^{-c n \varepsilon^2}.$$

**Remarque 5.1.** Le lemme 5.2 implique que si  $\mathbb{E} [f_n(u_1, \dots, u_n)]$  converge vers une limite finie, alors  $f_n(u_1, \dots, u_n)$  converge presque sûrement vers cette même limite (c'est une application directe du lemme du Borel-Cantelli).

De plus, en identifiant  $\mathbb{C}^n$  avec  $\mathbb{R}^{2n}$ , on peut généraliser le lemme 5.2 au cas complexe.

## Références

- [1] Z.D.Bai, Y.Q.Yin, *Necessary and sufficient conditions for almost sure convergences of the largest eigenvalue of a Wigner matrix*. Ann. Probab. 16(4), 1729-1741 (1988).
- [2] F.Benaych-Georges, R.N.Rao *The eigenvalues and eigenvectors of finite, low rank perturbations of large radom matrices*. Adv. Math. (2011), Vol. 227, no. 1, 494-521.
- [3] A.Guionnet, M.Krishnapur, and O.Zeitouni, *The single ring theorem*, arXiv :0909.2214v2 (2009).
- [4] R.A.Horn, C.R.Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
- [5] M. Ledoux. *The concentration of measure phenomenon*. Providence, RI, AMS, 2001.
- [6] T.Tao, *Outliers in the spectrum of iid matrices with bounded rank perturbations*. Probab. Theory and Related fields.