

Matrices aléatoires et probabilités libres

Introduction au domaine de recherche

Alexis Imbert

Thèse sous la direction de Camille Male et Pierre Tarrago

Septembre 2024

1 Historique et motivations



Figure 1: Timbre-poste hongrois en l'honneur de Wigner.

d'intrication), en mathématiques financières, en théorie de l'apprentissage, pour les systèmes complexes en écologie, etc La théorie des probabilités libres a été créée par Voiculescu dans les années 1980 dans le contexte de la théorie des algèbres d'opérateurs, plus précisément afin des comprendre les algèbres de von Neumann des groupes libres. C'est dix ans plus tard, de manière accidentelle, qu'a été établi le premier lien avec les matrices aléatoires. Aujourd'hui, les probabilités libres sont l'outil le plus adapté pour comprendre les problèmes avec plusieurs matrices aléatoires indépendantes. On utilisera ici principalement deux sources importantes dans la théorie des probabilités libres : [5] et [6]. Des extensions ont très récemment été apportées à cette théorie par Camille Male via la théorie de trafics [4]. Nous n'en parlerons pas ici, par soucis de concisions mais cette théorie permet notamment d'étudier des matrices aléatoires invariantes en loi par conjugaison par des matrices de permutation, ce qui n'est généralement pas le cas pour les probabilités libres. En outre, elle introduit une analogie de l'espérance conditionnelle pour des matrices aléatoires là où les probabilités libres introduisent une analogie de l'espérance.

Cette introduction aux matrices aléatoires et aux probabilités libres a pour but de présenter les théorèmes principaux de ces domaines. De plus, les méthodes combinatoires utilisées ici pour démontrer ces théorèmes permettent de comprendre les liens, mais aussi les différences entre probabilités classiques et libres.

L'étude des matrices aléatoires date des années 1930 avec Wishart en statistiques, et des années 50 avec Wigner en physique nucléaire. Ces derniers travaux ont eu une influence très importante à la jonction entre mécanique statistique et quantique avec Dyson et Metha dans les années 60, puis dans diverses directions comme le chaos quantique, la gravité quantique et la théorie des jauge supersymétriques. Le sujet s'est développé en statistiques dans les années 70 avec Marchenko, Pastur, Girko, Bai, et Silverstein, mais aussi de nombreux disciplines: la théorie des nombres (hypothèse de Riemann), la topologie et la combinatoire (espace de module des courbes), les algèbres d'opérateurs (probabilités libres), ainsi que l'analyse (équations de Painlevé, polynômes orthogonaux, problèmes de Riemann-Hilbert, etc).

Les matrices aléatoires ont de nombreuses applications. Elles sont essentielles pour comprendre les statistiques inférentielles en grande dimension, avec des conséquences en télécommunication et traitement du signal (grands réseaux de capteurs), en théorie quantique de l'information (phénomène

2 Plusieurs modèles de matrices aléatoires et premier théorème limite

2.1 Premiers modèles de matrices aléatoires

Commençons par définir plusieurs ensembles (déterministes) de matrices réelles ou complexes.

1. *Matrices orthogonales* : soit $O_N \in \mathcal{M}_N(\mathbb{R})$, $O_N \in \mathcal{O}_N \iff O_N O_N^T = O_N^T O_N = I_N$.
2. *Matrices unitaires* : Soit $U_N \in \mathcal{M}_N(\mathbb{C})$, $U_N \in \mathcal{U}_N \iff U_N U_N^* = U_N^* U_N = I_N$.

On définit alors des modèles de matrices aléatoires à partir de ces ensembles déterministes. Notons dès maintenant que les matrices aléatoires que nous étudierons ici seront toujours hermitiennes. Elles auront alors toutes un spectre réel.

Définition 2.1 (Matrice du GOE : Gaussian orthogonal ensemble). *Soit G_N une matrice avec des entrées gaussiennes réelles standard i.i.d. alors*

$$H_N := \frac{(G_N + G_N^T)}{\sqrt{2N}}$$

est une matrice du GOE. De plus, elle est invariante en loi par conjugaison par des matrices orthogonales.

Définition 2.2 (Matrices du GUE : Gaussian unitary ensemble). *Soit G_N une matrice avec des entrées gaussiennes complexes standard iid (i.e. $(G_N)_{i,j} \stackrel{\Delta}{=} \mathcal{N}(0, \frac{1}{2}) + i\mathcal{N}(0, \frac{1}{2})$), alors*

$$H_N = \frac{G_N + G_N^*}{\sqrt{2N}}$$

est une matrice du GUE. De plus, elle est invariante en loi par conjugaison par des matrices unitaires.

Définition 2.3 (Matrices de Wigner). *Une matrice hermitienne $X_N = (\frac{x_{i,j}}{\sqrt{N}})_{i,j}$ est dite de Wigner si*

- les $(x_{i,j})_{i \leq j}$ sont indépendants;
- les $(x_{i,i})_{i \in [N]}$ sont i.i.d. de loi μ ;
- les $(x_{i,j})_{i < j}$ sont i.i.d. de loi ν ,

avec μ et ν indépendants de N . On suppose également que μ et ν admettent des moments de tout ordre.

Remarque 1. *L'hypothèse sur les moments est surtout là pour pouvoir donner une démonstration combinatoire du théorème de Wigner. En réalité, on peut énormément affaiblir cette hypothèse pour le théorème de Wigner mais on n'en parlera pas ici.*

On peut noter que les matrices du GOE et du GUE sont toutes des matrices de Wigner. On peut également généraliser les matrices de Wigner en les multipliant entrées par entrées par des quantités uniformément bornées et on peut alors obtenir des résultats similaires sur ces nouvelles matrices.

2.2 Théorème de Wigner

Une matrice aléatoire hermitienne $M \in \mathcal{M}_N(\mathbb{C})$ possède N valeurs propres réelles que l'on note $\lambda_1 \leq \dots \leq \lambda_N$. On définit alors une mesure (parfois aléatoire) contenant toute l'information des valeurs propres de la matrice.

Définition 2.4 (Mesure empirique des valeurs propres). *La mesure empirique des valeurs propres de la matrice M est la mesure de probabilité sur \mathbb{R} définie par*

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

Par le théorème spectral, celle-ci vérifie pour toute fonction f ,

$$\int f(\lambda) d\mu_M(\lambda) = \frac{1}{N} \text{Tr}(f(M)).$$

Convergence faible. On voudrait définir une notion de convergence pour des matrices aléatoires qui permette à des matrices de plus en plus grandes de converger quand même. Il est alors naturel de passer par la mesure empirique des valeurs propres et de définir une convergence sur l'ensemble des mesures de probabilités réelles. Soient $\nu, (\nu_N)_{N \geq 1}$ des mesures de probabilité sur \mathbb{R} . On dit que ν_N converge faiblement vers ν si pour toutes fonctions continues bornées f ,

$$\int f d\nu_N \rightarrow \int f d\nu$$

La topologie faible est métrisable avec la distance de Lévy-Prokhorov d_l et l'ensemble des mesures de probabilité sur \mathbb{R} muni de d_l est un espace polonais.

Si $(\nu_N)_{N \geq 1}$ sont des mesures aléatoires, on dit alors que ν_N converge faiblement vers ν presque sûrement lorsque p.s. pour toute fonction continue bornée f ,

$$\int f d\nu_N \rightarrow \int f d\nu$$

Nous pouvons alors désormais énoncer le résultat le plus essentiel de la théorie des matrices aléatoires.

Théorème 2.1 (Théorème de Wigner). Soit $(X_N)_{N \in \mathbb{N}}$ une suite de matrices de Wigner de taille $N \times N$. On suppose de plus que ses entrées sont centrées et on note σ^2 la variance commune des entrées au dessus de la diagonale. Alors p.s. faiblement,

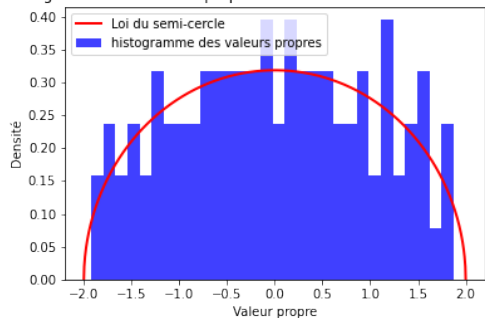
$$\mu_{X_N} \rightarrow \mu_{sc}$$

où μ_{sc} est la loi du demi-cercle sur l'intervalle $[-2\sigma, 2\sigma]$ de densité

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

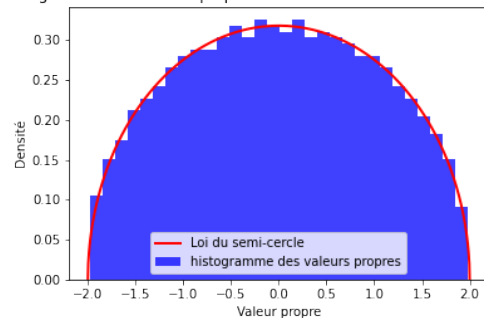
Tout comme le théorème central limite, le théorème de Wigner est très visuel, voici des histogrammes des valeurs propres pour des matrices du GUE (qui sont des matrices de Wigner) de taille 100×100 puis 1000×1000 .

Histogramme des valeurs propres d'une matrice GUE de taille 100x100



(a) Plot pour $N = 100$

Histogramme des valeurs propres d'une matrice GUE de taille 1000x100



(b) Plot pour $N = 1000$

2.3 Approche analytique

Une preuve d'une version affaiblie de ce théorème consiste à utiliser la transformée de Stieltjes des mesures.

Définition 2.5 (Transformée de Stieltjes). Soit ν une mesure réelle et X une variable aléatoire réelle de loi ν . La transformée de Stieltjes de ν est la fonction

$$g_\nu(z) = \int \frac{1}{t-z} d\nu(t) = \mathbb{E} \left(\frac{1}{X-z} \right) = \sum_{n \geq 0} z^{-(n+1)} \mathbb{E}(X^n). \quad (1)$$

La dernière écriture sous forme de série est valable notamment lorsque ν est à support compact, auquel cas, la série est absolument convergente sur son disque ouvert de convergence. De plus lorsque la mesure ν est absolument continue par rapport à la mesure de Lebesgue, on peut retrouver cette densité à partir de la transformée de Stieltjes via la formule d'inversion de Stieltjes. On écrit $d\nu(t) = \rho(t)dt$

$$\rho(x) = -\frac{1}{\pi} \lim_{\epsilon \rightarrow 0^+} \text{Im} g_\nu(x + i\epsilon). \quad (2)$$

Ainsi, lorsqu'on s'intéresse à une suite de mesures, la convergence de ces mesures sont étroitement liées à la convergence de leur transformées de Stieltjes. Selon le type de convergence on a parfois équivalence entre les deux, parfois seulement implication (voir [1]).

On peut alors calculer la transformée de Stieltjes de la loi du semi-cercle $g_{\mu_{sc}}(z) = \frac{-z + \sqrt{z^2 - 4}}{2}$ et remarquer que c'est l'unique solution à partie imaginaire positive de l'équation $s + \frac{1}{s+z} = 0$. Il faut alors montrer que la transformée de Stieltjes de $\mathbb{E}\mu_{X_N}$ vérifie asymptotiquement cette équation de point fixe, (voir [1] pour plus de précisions) puis montrer que les mesures μ_{X_N} et $\mathbb{E}\mu_{X_N}$ sont très proches avec grande probabilité lorsque N tend vers l'infini.

On ne rentre pas dans les détails ici car la preuve est un peu technique (on pourra regarder la partie 2.4.1 de [1] pour plus de détails).

2.4 Approche combinatoire

Nous allons désormais donner une ébauche de preuve du théorème de Wigner par une approche combinatoire très instructive, en utilisant la méthode des moments. Cette méthode est assez classique et est ici fortement inspirée de [2] et [1]. Pour cela, nous allons désormais supposer que :

- $\sigma^2 = 1$, quitte à dilater la loi limite, on peut toujours le supposer.
- Les lois μ et ν sont uniquement déterminées par leur moments. On pourrait par exemple supposer qu'elles sont à support compact. Se passer de cette hypothèse demande un peu de travail et nous n'entrerons pas dans les détails.

Lemme 2.2 (Méthode des moments). *Dans ce cadre, on a que si pour tout $k \geq 1$, $\int x^k d\mu_{X_N}(x) \xrightarrow{N \rightarrow \infty} \int x^k d\mu_{sc}(x)$ p.s., alors $\mu_{X_N} \rightarrow \mu_{sc}$ faiblement p.s..*

C'est un lemme classique de probabilités que nous ne prouverons pas ici.

Moments du semi-cercle. On peut d'une part facilement calculer les moments de la loi du semi-cercle,

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

Pour $k \geq 0$,

$$\int x^{2k+1} d\mu_{sc}(x) = 0, \text{ et } \int x^{2k} d\mu_{sc}(x) = C_k. \quad (3)$$

où $C_k = \frac{1}{k+1} \binom{2k}{k}$ est le k -ième nombre de Catalan.

Remarque 2. *L'apparition ici des nombres de Catalan jouera un rôle essentiel dans la suite.*

Convergence de l'espérance. On va en fait d'abord montrer que l'espérance de la mesure empirique des valeurs propres de X_N converge faiblement vers la loi du semi-cercle.

Lemme 2.3. *Pour tout $k \geq 0$,*

$$\mathbb{E} \int x^k d\mu_{X_N} \rightarrow \int x^k d\mu_{sc}$$

Démonstration. On a d'abord que

$$\mathbb{E} \int x^k d\mu_{X_N} = \frac{1}{N} \mathbb{E} \text{Tr} X_N^k$$

Puis on peut écrire :

$$\begin{aligned} \frac{1}{N} \mathbb{E} \text{Tr} X_N^k &= \frac{1}{N^{1+\frac{k}{2}}} \mathbb{E} \left[\sum_{(r(1), \dots, r(k))} \prod_{j=1}^k x_{r(j), r(j+1)} \right] \\ &= \frac{1}{N^{1+\frac{k}{2}}} \sum_{(r(1), \dots, r(k))} T(\mathbf{r}) \end{aligned}$$

où $\mathbf{r} = (r(1), \dots, r(k))$, $r(k+1) = r(1)$ et

$$T(\mathbf{r}) = \mathbb{E} \left[\prod_{j=1}^k x_{r(j), r(j+1)} \right]$$

On dit que deux chemins \mathbf{r} et \mathbf{r}' sont équivalents s'il existe $\sigma \in \mathfrak{S}_k$ tel que $\forall j, r(j) = r'(\sigma(j))$. Alors, $T(\mathbf{r})$ est constant sur chaque classe d'équivalence de cette relation d'équivalence puisque la loi de X_N est invariante par permutation des entrées. De plus, en notant $|\mathbf{r}| := \#\{r(1), \dots, r(k)\}$, le nombre de chemins dans la classe d'équivalence de \mathbf{r} est

$$N(N-1)\dots(N-|\mathbf{r}|+1) \sim N^{|\mathbf{r}|}$$

Notons alors \mathcal{C}_k l'ensemble des chemins canoniques, où l'on a choisi un chemin dans chaque classe d'équivalence en imposant $r(1) = 1$ et

$$r(j+1) \leq 1 + \max_{s \leq j} (r(s))$$

Alors,

$$\frac{1}{N} \mathbb{E} \text{Tr} X_N^k = \sum_{\mathbf{r} \in \mathcal{C}_k} \frac{N(N-1)\dots(N-|\mathbf{r}|+1)}{N^{1+\frac{k}{2}}} T(\mathbf{r})$$

Définissons désormais $G = (V, E)$ le graphe donné par $V = \{r(1), \dots, r(k)\}$ et $E = \{\{r(1), r(2)\}, \dots, \{r(k), r(1)\}\}$. Pour $e \in E$, on définit la multiplicité de l'arête e par

$$m_e := \sum_{j=1}^k \mathbf{1}_{\{r(j), r(j+1)\} = e}$$

Alors par indépendance,

$$T(\mathbf{r}) = \mathbb{E} \prod_{j=1}^k x_{r(j), r(j+1)} = \prod_{e \in E} \mathbb{E} [x_e^{m_e}]$$

Et comme les variables sont centrées, s'il existe $e \in E$ telle que $m_e = 1$, alors $T(\mathbf{r}) = 0$. Donc pour tout $e \in E$, $m_e \geq 2$ et ainsi :

$$k = \sum_{e \in E} m_e \geq 2 \times |E|$$

On en déduit que

$$|E| \leq \lfloor \frac{k}{2} \rfloor$$

et alors par la formule d'Euler sur les graphes :

$$|\mathbf{r}| = |V| \leq 1 + |E| \leq 1 + \lfloor \frac{k}{2} \rfloor$$

avec égalité si G est un arbre ($|V| = 1 + |E|$ dans ce cas) et que $\forall e \in E, m_e = 2$ (et en particulier k pair). On note $\tilde{\mathcal{C}}_k$ l'ensemble des chemins canoniques \mathbf{r} donnant ce cas d'égalité. De plus, les autres chemins

ne contribuent pas à l'ordre 1 puisque $|\mathbf{r}| \leq 1 + \frac{k}{2} - \frac{1}{2}$ et pour tout \mathbf{r} , $T(\mathbf{r})$ est borné uniformément en N . On obtient alors :

$$\begin{aligned} \frac{1}{N} \mathbb{E} \text{Tr} X_N^k &= \mathbf{1}_{k \text{ pair}} \sum_{\mathbf{r} \in \tilde{\mathcal{C}}_k} (\mathbb{E} x_{1,2}^2)^{\frac{k}{2}} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \\ &= \mathbf{1}_{k \text{ pair}} |\tilde{\mathcal{C}}_k| + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \end{aligned}$$

Enfin, nous ne détaillerons pas ici mais il vient très rapidement que les chemins de $\tilde{\mathcal{C}}_{2k}$ sont en bijection avec \mathcal{D}_k les mots de Dyck de longueur $2k$ (*i.e.* les parenthésages valides avec k parenthèses ouvertes) qui sont comptés par les nombres de Catalan C_k . On peut donc en conclure que

$$\frac{1}{N} \mathbb{E} \text{Tr} X_N^k = \mathbf{1}_{k \text{ pair}} C_{k/2} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$$

ce qui achève notre preuve. □

Convergence faible *p.s.* Il nous reste maintenant à montrer que la valeur du moment d'ordre k de la mesure empirique des valeurs propres de X_N est *assez proche* de celle du moment d'ordre k de la mesure empirique des valeurs propres de $\mathbb{E} X_N$. On ne le fait pas ici mais on peut montrer de manière similaire que pour tout $k \geq 1$,

$$\text{Var}\left(\frac{1}{N} \text{Tr}(X_N^k)\right) = \mathcal{O}\left(\frac{1}{N^2}\right)$$

Ainsi, par le théorème de convergence monotone,

$$\mathbb{E} \left[\sum_{n \geq 0} \left(\int x^k d\mu_{X_N} - \mathbb{E} \int x^k d\mu_{X_N} \right)^2 \right] < \infty$$

Donc *p.s.*, $\sum_{n \geq 0} \left(\int x^k d\mu_{X_N} - \mathbb{E} \int x^k d\mu_{X_N} \right)^2 < \infty$, et alors en particulier on a que *p.s.*,

$$\int x^k d\mu_{X_N} - \mathbb{E} \int x^k d\mu_{X_N} \rightarrow 0$$

Or par le lemme 2.3, $\mathbb{E} \int x^k d\mu_{X_N} \rightarrow \int x^k d\mu_{sc}$ et donc finalement :

$$\int x^k d\mu_{X_N} \rightarrow \int x^k d\mu_{sc}$$

On peut conclure par le lemme 2.2 que $\mu_{X_N} \rightarrow \mu_{sc}$ faiblement *p.s.*

3 Probabilités libres

La théorie des probabilités libres permet d'étudier les matrices aléatoires comme élément d'un espace de probabilité non-commutatif plutôt que comme un ensemble de variables aléatoires (ses entrées). Bien sûr, cette théorie a un impact plus large que les matrices aléatoires puisqu'elle formalise une notion de variable aléatoire non commutative dont les matrices en sont un cas particulier. On s'intéresse ici à cette théorie d'un point de vue formel mais on reviendra souvent à travers des exemples sur l'application de cette théorie aux matrices aléatoires.

3.1 Espace, loi et convergence

En probabilité libres, la notion équivalente à celle de l'espace de probabilité classique $(\Omega, \mathcal{F}, \mathbb{P})$ est celle d'**espace de probabilité non-commutatif** (\mathcal{A}, ϕ)

Définition 3.1. Un espace de probabilité non-commutatif (\mathcal{A}, ϕ) consiste en une $*$ -algèbre unitaire sur \mathbb{C} (pas forcément commutative) \mathcal{A} et un état ϕ : une forme linéaire unitaire ($\phi(1_{\mathcal{A}}) = 1$) positive ($\phi(a^*a) \geq 0$ pour tout $a \in \mathcal{A}$). On rappelle qu'une $*$ -algèbre unitaire est une algèbre munie d'une opération involutive $*$ et qui possède un élément unitaire $1_{\mathcal{A}}$.

La forme linéaire ϕ sera l'analogie non-commutatif de l'espérance. Les éléments $a \in \mathcal{A}$ sont appelés des variables aléatoires non-commutatives.

Remarque 3. En partant d'un espace de probabilité classique $(\Omega, \mathcal{F}, \mathbb{P})$, on peut construire un espace de probabilité non commutatif grâce à $L^{\infty-}(\Omega, \mathbb{P})$: l'espace des variables aléatoires ayant des moments de tout ordre. Dans ce cas $(L^{\infty-}(\Omega, \mathbb{P}), \mathbb{E})$ est un espace de probabilité non commutatif (même si ses éléments peuvent être commutatifs dans le cas de variables aléatoires réelles par exemple).

Dans ce cadre, on peut alors parler de variables aléatoires non-commutatives

- auto-adjointes, i.e. $a \in \mathcal{A}$ tels que $a = a^*$,
- unitaires, i.e. $u \in \mathcal{A}$ tels que $uu^* = u^*u = 1_{\mathcal{A}}$,
- normales, i.e. $a \in \mathcal{A}$ tels que $aa^* = a^*a$.

Comme on l'a vu dans la démonstration du théorème de Wigner, la notion de moments joue un rôle décisif dans l'étude de la mesure spectrale limite de matrices aléatoires. Plus encore, quand on se limite à l'étude de matrices aléatoires à entrées uniformément bornées, la convergence presque sûre des moments est suffisante à la convergence faible presque sûre de la mesure spectrale. On définit alors une notion de loi sur un espace de probabilité non commutatif (\mathcal{A}, ϕ) .

Définition 3.2. Soit J un ensemble d'indices. Pour $\mathbf{a} = (a_j)_{j \in J} \in \mathcal{A}$ la **loi non commutative** de \mathbf{a} est la donnée de la forme linéaire

$$\begin{aligned} \Phi_{\mathbf{a}} : \mathbb{C}\langle \mathbf{x}, \mathbf{x}^* \rangle &\longrightarrow \mathbb{C} \\ P &\longmapsto \phi(P(\mathbf{a})). \end{aligned}$$

Dans cette formule, $\mathbb{C}\langle \mathbf{x}, \mathbf{x}^* \rangle$ désigne l'ensemble des polynômes non-commutatifs à coefficients dans \mathbb{C} et à variables dans $\mathbf{x} \cup \mathbf{x}^* = (x_j)_{j \in J} \cup (x_j^*)_{j \in J}$.

Remarquons que pour un élément a auto-adjoint ($a = a^*$) la loi non commutative de a est simplement la donnée de $(\phi(a^n))_{n \in \mathbb{N}}$.

Grâce à cette notion de loi on peut définir une convergence en loi.

Définition 3.3. Soit $\mathbf{a}^{(N)} = (a_j^{(N)})_{j \in J} \in (\mathcal{A}_N, \phi_N)$ une suite de familles de variables aléatoires non commutatives (qui peuvent vivre dans des espaces de probabilités non-commutatifs différents). On dit que $\mathbf{a}^{(N)}$ **converge** vers $\mathbf{a} \in (\mathcal{A}, \phi)$ si pour tout polynôme non commutatif $P \in \mathbb{C}\langle \mathbf{x}, \mathbf{x}^* \rangle$, on a $\phi_N(P(\mathbf{a}^{(N)})) \rightarrow \phi(P(\mathbf{a}))$ (dans \mathbb{C}). La convergence de $\mathbf{a}^{(N)}$ vers \mathbf{a} est donc la convergence point par point de $\Phi_{\mathbf{a}^{(N)}}$ vers $\Phi_{\mathbf{a}}$.

Afin de mieux comprendre ces notions abstraites d'espaces de probabilités non-commutatifs, prenons l'exemple de matrices aléatoires. On note $L^{\infty-}(\Omega, \mathbb{P})$ l'espace des variables aléatoires ayant des moments de tout ordre. L'espace $(\mathcal{M}_N(L^{\infty-}(\Omega, \mathbb{P})), \mathbb{E} \frac{1}{N} \text{Tr}) =: (\mathcal{A}_N, \phi_N)$ est bien un espace de probabilité non-commutatif. Les matrices aléatoires hermitiennes sont donc des éléments auto-ajoints de cette algèbre. Soit $A_N \in \mathcal{A}_N$, en utilisant l'invariance par permutation circulaire de la trace, la loi non-commutative de A_N est la donnée pour tout $k, l \in \mathbb{Z}$ de

$$\mathbb{E} \left[\frac{1}{N} \text{Tr} A^k A^{*l} \right] = \mathbb{E} \left[\int z^k \bar{z}^l d\mu_{A_N}(z) \right].$$

Soit maintenant $(A_N)_{N \in \mathbb{N}} \in \mathcal{A}_N$ une suite de matrices aléatoires de taille N , la convergence de la suite $(A_N)_{N \in \mathbb{N}}$ en tant qu'éléments des espaces \mathcal{A}_N est alors définie comme étant la convergence de l'espérance des moments de leur mesures spectrales. Par exemple, si les A_N sont des matrices de Wigner dont les entrées sont de variance 1, le théorème de Wigner nous assure que ces moments convergent vers les moments de la loi du semi-cercle. En définissant alors un *élément semi-circulaire* $s \in (\mathcal{A}, \phi)$ d'un espace de probabilité non-commutatif comme étant un élément auto-adjoint dont les moments sont ceux d'une semi-circulaire, on obtient que la suite $(A_N)_{N \in \mathbb{N}}$ converge vers s . Il est important de comprendre que,

comme pour des variables aléatoires, l'important n'est pas tant ce que sont cet espace (\mathcal{A}, ϕ) et cet élément s mais c'est la loi de s qui nous donne le plus d'information. En fait, si l'on prend (\mathcal{A}, ϕ) et (\mathcal{B}, ψ) des algèbres engendrées respectivement par $\mathbf{a} = (a_1, \dots, a_s)$ et $\mathbf{b} = (b_1, \dots, b_s)$ où la distribution de \mathbf{a} dans (\mathcal{A}, ϕ) est la même que celle de \mathbf{b} dans (\mathcal{B}, ψ) , il existe un morphisme uniquement déterminé par \mathbf{a} et \mathbf{b} entre (\mathcal{A}, ϕ) et (\mathcal{B}, ψ) (voir [6]).

3.2 Liberté et cumulants libres

3.2.1 Liberté

Tout au long de cette sous-section, on se place dans un espace de probabilité non-commutatif (\mathcal{A}, ϕ) . Généralement, la notion de liberté est d'abord définie pour des sous-algèbres de \mathcal{A} puis pour des variables aléatoires, comme étant la liberté entre les sous-algèbres engendrées. Ici, on donne directement une définition pour les variables aléatoires.

Définition 3.4. Soient $(a_i)_{i \in I} \in \mathcal{A}$. On dit que les a_i sont libres (ou librement indépendants) si pour tout polynômes non-commutatifs $P_1, \dots, P_n \in \mathbb{C}\langle \mathbf{x}, \mathbf{x}^* \rangle$, on a $\phi(P_1(a_{i_1}) \cdots P_n(a_{i_n})) = 0$ dès que les deux conditions suivantes sont vérifiées :

1. $\forall j \in [n], \phi(P_j(a_{i_j})) = 0$,
2. $i_1 \neq i_2 \neq \dots \neq i_n$.

Autrement dit, un produit alterné d'éléments d'espérance nulle est d'espérance nulle si ces éléments sont libres.

Remarquons que la notion de liberté dépend de ϕ , tout comme l'indépendance classique est dépendante de la mesure de probabilité que l'on choisit sur l'espace de probabilité.

Tout comme l'indépendance de X et Y permet de calculer $\mathbb{E}XY$ grâce à $\mathbb{E}X$ et $\mathbb{E}Y$ dans le cas classique, la liberté permet de calculer des moments alternés dans le cas libre.

Proposition 3.1. Soient $\{a_1, a_2\}$ et b des éléments d'un espace de probabilité non commutatif (\mathcal{A}, ϕ) tels que $\{a_1, a_2\}$ et b soient libres. En développant le produit

$$0 = \phi((a_1 - \phi(a_1)1)(b - \phi(b)1)(a_2 - \phi(a_2)1),$$

on obtient l'égalité

$$\phi(a_1 b a_2) = \phi(b) \phi(a_1 a_2). \quad (4)$$

3.2.2 Cumulants

Afin de motiver la notion de cumulants, faisons un détour par le théorème central limite.

Prenons a_1, \dots, a_N des variables aléatoires auto-adjointes de même loi d'un espace de probabilité non commutatif (\mathcal{A}, ϕ) centrées de variance commune $\phi(a_i^2) = \sigma^2$. Pour l'instant, on s'intéresse au cas général donc on suppose que les variables sont indépendantes ou libre selon le cas. On veut donc trouver la limite de

$$\phi \left[\frac{(a_1 + \dots + a_N)^n}{N^{n/2}} \right].$$

Regardons alors à N et n fixés,

$$\phi[(a_1 + \dots + a_N)^n] = \sum_{1 \leq r(1), \dots, r(n) \leq N} \phi(a_{r(1)} \cdots a_{r(n)}).$$

Comme les a_i ont tous la même loi, la quantité dans la somme ne dépend pas de l'ordre des indices mais seulement du nombre d'indices ayant la même valeur. Pour encoder cette information, on va plutôt sommer sur les partitions $\pi = \{V_1, \dots, V_s\}$ de $\{1, \dots, n\}$ puis compter le nombre A_π^N de réindigage r tels que $r(i) = r(j)$ si et seulement si i et j sont dans un même bloc. On notera κ_π la valeur commune de $\phi(a_{r(1)} \cdots a_{r(n)})$ pour tous les r dans une même classe d'équivalence encodée par π . Maintenant, il nous reste à montrer que la plupart des partitions π donnent lieu à une contribution nulle à la limite $N \rightarrow \infty$.

Premièrement, les partition π contenant un singleton (c'est à dire qu'il existe un bloc $V_m = \{i\}$ pour un certain i dans $\{1, \dots, n\}$) amènent à une limite nulle. En effet, par indépendance ou liberté (voir la Proposition 3.1) et car les variables sont centrées,

$$\kappa_\pi = \phi(a_i)\phi(a_{r(1)} \cdots a_{r(n)}) = 0.$$

Ensuite, pour une partition $\pi = \{V_1, \dots, V_s\}$ fixée, il y a N possibilités pour choisir la valeur commune du premier bloc V_1 , puis $N - 1$ possibilités pour la valeur du second bloc V_2 etc. Ainsi, en notant $|\pi|$ le nombre de blocs de π , on a

$$A_\pi^N = N(N - 1) \cdots (N - |\pi| + 1)$$

qui croit comme $N^{|\pi|}$ lorsque N croit. Finalement, on obtient

$$\lim_{N \rightarrow \infty} \phi \left[\frac{(a_1 + \cdots + a_N)^n}{N^{n/2}} \right] = \lim_{N \rightarrow \infty} \sum_{\pi} N^{|\pi| - n/2} \kappa_\pi. \quad (5)$$

On conclut alors que seules les partitions π correspondant à des pairings (ou appariements), c'est-à-dire avec uniquement des blocs composés de deux éléments, comptent à la limite.

Cas classique Maintenant, dans le cas classique, c'est-à-dire lorsqu'on s'intéresse à des variables aléatoires réelles ayant des moments de tout ordre, il est facile de voir que pour de tels π , κ_π vaut σ^n et que le nombre de tels π vaut $(n - 1)(n - 3) \cdots 3 \times 1$. Ces nombres sont également les moments d'une loi normale centrée réduite. De plus, la loi normale est une loi déterminée par ses moments, c'est-à-dire que la donnée des moments suffit à caractériser la loi. C'est un exercice classique qu'on ne refait pas ici.

En conclusion, nous venons de prouver la version du théorème central limite suivante :

Théorème 3.2 (TCL classique). *Soient a_1, \dots, a_N des variables aléatoires auto-adjointes d'un espace (\mathcal{A}, ϕ) , centrées de même loi et indépendantes au sens classique. Supposons de plus qu'elles sont de variance σ^2 . Alors $\frac{a_1 + \cdots + a_N}{\sqrt{N}}$ converge (pour la notion de convergence définie précédemment) vers un élément normal x .*

Cas libre Repassons maintenant au cas non commutatif. Partons de l'équation (5) que l'on a établi précédemment. En fait, même parmi les pairings π , certains amènent à une contribution nulle. En effet, pour un pairing $\pi = \{V_1, \dots, V_k\}$ de $\{1, \dots, n = 2k\}$, prenons $(r(1), \dots, r(n))$ un multi-indice correspondant à la partition π . Il y a alors deux possibilités :

- Tous les indices consécutifs sont différents : $r(1) \neq r(2) \neq \cdots \neq r(n)$. Dans ce cas, par définition de la liberté, $\kappa_\pi = \phi(a_{r(1)} \cdots a_{r(n)}) = 0$.
- Deux indices consécutifs sont identiques : $r(m) = r(m + 1) = r$ pour un certain $m \in \{1, \dots, n\}$. Dans ce cas, comme cet indice n'apparaît pas ailleurs, $\{a_{r(1)}, \dots, a_{r(m-1)}, a_{r(m+2)}, \dots, a_{r(n)}\}$ et $\{a_{r(m)}, a_{r(m+1)}\}$ sont libres. En utilisant l'équation (3.1), on a

$$\kappa_\pi = \phi(a_{r(1)} \cdots a_{r(n)}) = \phi(a_r a_r) \phi(a_1 \cdots a_{r(m-1)} a_{r(m+2)} \cdots a_{r(n)}) = \sigma^2 \phi(a_1 \cdots a_{r(m-1)} a_{r(m+2)} \cdots a_{r(n)}).$$

Ainsi pour obtenir une contribution non nulle il faut pouvoir répéter cette opération jusqu'à réduire la taille du moment considéré à 0. La valeur de κ_π sera dans ce cas σ^n . On peut alors montrer (on ne le fera pas ici) que de telles partitions sont des partitions non-croisées, c'est-à-dire qu'il n'existe pas de $1 \leq p_1 \leq q_1 \leq p_2 \leq q_2 \leq n$ tels que p_1 soit apparié à p_2 et q_1 à q_2 . On note l'ensemble de pairings non croisés $NC_2(n)$. On peut également montrer que $\#NC_2(n) = C_{n/2}$ le nombre de Catalan. Enfin la loi semi-circulaire est une loi déterminée par ses moments, ses moments impairs sont nuls et ses moments pairs sont les nombres de Catalan. Ainsi, on vient de montrer une version libre du théorème central limite.

Théorème 3.3 (TCL libre). *Soient a_1, \dots, a_N des éléments auto-adjoints d'un espace (\mathcal{A}, ϕ) , centrées, de même loi et libres. Supposons de plus qu'ils sont de variance σ^2 . Alors $\frac{a_1 + \cdots + a_N}{\sqrt{N}}$ converge en loi vers un élément semi-circulaire s .*

Cette analogie que nous venons d'établir entre le monde libre est classique nous montre plusieurs choses. Premièrement la loi semi-circulaire joue le même rôle dans le cas libre que la loi normale dans le cas classique. Ensuite, même si l'on a considéré seulement des pairings ici, on a pu observer que la transition du cas classique au cas libre est équivalente, à un niveau combinatoire, à la transition entre toutes les partitions et les partitions non-croisées. Cette analogie nous permet de mieux comprendre la définition suivante des cumulants libres.

Définition 3.5 (Cumulants libres). *Les cumulants libres $(\kappa_n)_{n \in \mathbb{N}}$ sont des applications n -linéaires de \mathcal{A}^n dans \mathbb{C} définies implicitement par les formules suivantes. Soient $a_1, \dots, a_n \in \mathcal{A}$,*

$$\phi(a_1 \cdots a_n) = \sum_{\pi \in \text{NC}(n)} \kappa_\pi(a_1, \dots, a_n), \quad (6)$$

$$\kappa_\pi(a_1, \dots, a_n) = \prod_{B=\{a_{j_1} < \dots < a_{j_l(B)}\} \text{ blocs de } \pi} \kappa_{l(B)}(a_{j_1}, \dots, a_{j_l(B)}), \quad (7)$$

où $\text{NC}(n)$ désigne l'ensemble des partitions non-croisées de $\{1, \dots, n\}$ et $l(B)$ la taille du bloc B .

On peut vérifier que cette formule définit bien récursivement les cumulants libres. De plus, ces cumulants caractérisent la liberté de manière plus élémentaire que les moments puisque les cumulants mixtes d'éléments libres sont tous nuls. Plus précisément,

Théorème 3.4. *Soient $(a_i)_{i \in I} \in (\mathcal{A}, \phi)$, il y a équivalence entre*

- *Les $(a_i)_{i \in I}$ sont libres.*
- *Pour $n \geq 2$ et $i_1, \dots, i_n \in I$, $\kappa_n(a_{i_1}, \dots, a_{i_n}) = 0$ dès qu'il existe k, l tels que $i_k \neq i_l$.*

Par exemple pour deux variables aléatoires libres x et y dans (\mathcal{A}, ϕ) , on a $\kappa_n(x + y, \dots, x + y) = \kappa_n(x, \dots, x) + \kappa_n(y, \dots, y)$ en utilisant d'abord la multilinéarité des cumulants puis la propriété de liberté.

Tout comme on a pu établir une analogie combinatoire entre le cas classique et le cas libre, on peut également utiliser des outils analytiques adaptés au cas libre.

Définition 3.6. *Pour une variable non-commutative $a \in (\mathcal{A}, \phi)$, on peut également définir sa transformée de Stieltjes non commutative (que l'on nomme juste transformée de Stieltjes) de manière analogue au cas classique (Equation (1)).*

$$G_a(z) = \sum_{n \geq 0} z^{-(n+1)} \phi(a^n) \quad (8)$$

Dans le cas des matrices aléatoires, la transformée de Stieltjes est simplement l'espérance de la trace de la résolvante.

On définit maintenant la R -transformée d'une variable aléatoire non-commutative a qui va permettre de caractériser simplement la liberté entre deux variables aléatoires et qui, grâce à la relation entre moment et cumulants, va être liée à la transformée de Stieltjes.

Définition 3.7 (R -transformée). *Soit $a \in (\mathcal{A}, \phi)$ une variable aléatoire non-commutative.*

La R -transformée est la série

$$R_a(z) := \sum_{k \geq 1} \kappa_k(a, \dots, a) z^{k-1}.$$

On vérifie facilement grâce au théorème 3.4 que deux variables x et y sont libres si et seulement si on a $R_{x+y}(z) = R_x(z) + R_y(z)$.

Enfin, on peut montrer [6] que la transformée de Stieltjes et la R -transformée d'une même variable a sont reliées via l'équation au point fixe suivante. Pour z , avec une partie imaginaire assez grande,

$$G_a(z) = (z - R_a(G_a(z)))^{-1}. \quad (9)$$

Proposition 3.5 (Equations de subordinations). *Pour x et y , deux éléments libres d'un espace de probabilité non-commutatif, on a plusieurs équations reliant les différentes transformées.*

$$G_{x+y}(z) = (z - R_{x+y}(G_{x+y}(z)))^{-1} = (z - R_y(G_{x+y}(z)) - R_x(G_{x+y}(z)))^{-1} \quad (10)$$

$$= G_x(z - R_y(G_{x+y}(z))) = G_y(z - R_x(G_{x+y}(z))) \quad (11)$$

Les premières équations viennent de la caractérisation de la liberté par la R -transformée. Les deux dernières sont des manipulations de la définition de la R -transformée (voir [5] pour plus de détails).

Avant de donner des exemples, il faut évoquer la notion de liberté asymptotique, qui permet d'utiliser les caractéristiques de la liberté pour pouvoir en déduire des informations sur le spectre limite de la somme de matrices aléatoires. De manière similaire à la liberté, l'asymptotique liberté se définit sur la limites de sous-algèbres. On se contente ici de donner la définition pour deux familles de matrices, donc dans la suite d'espaces de probabilités $(\mathcal{M}_N(L^\infty(\Omega, \mathbb{P})), \mathbb{E}[\frac{1}{N}\text{Tr}\cdot])$.

Définition 3.8. Soient $\mathbf{A}^{(N)} = (A_j^{(N)})_{j \in J}$ et $\mathbf{B}^{(N)} = (B_j^{(N)})_{j \in J}$ deux familles de matrices qui admettent une limite en distribution non-commutative (respectivement \mathbf{a} et \mathbf{b} dans un espace non commutatif abstrait (\mathcal{A}, ϕ)). On dit que les familles $\mathbf{A}^{(N)}$ et $\mathbf{B}^{(N)}$ sont asymptotiquement libres si \mathbf{a} et \mathbf{b} le sont dans (\mathcal{A}, ϕ) .

Exemple 3.6. Des (suites de) matrices du GUE X_N, Y_N (disons d'entrées à variance commune $\sigma^2 = 1$ pour simplifier) indépendantes sont asymptotiquement libres (voir [5] Theorem 10). Si l'on note x, y respectivement leur limite en distribution non commutative dans un espace de probabilité non-commutatif (\mathcal{A}, ϕ) , par le théorème de Wigner, on sait que les moments de x et de y sont les moments de la loi semi-circulaire. Que peut-on en déduire sur la loi de $x + y$ (et donc sur la densité de valeur propre limite de $X_N + Y_N$) ? Utilisons la caractérisation par la R -transformée pour répondre à cette question. On peut montrer que seul le cumulatif d'ordre 2 d'une matrice du GUE est non nul et vaut $\sigma^2 = 1$. On obtient ainsi $R_x(z) = R_y(z) = z$, donc $R_{x+y}(z) = 2z$. Il suffit alors de résoudre

$$G_{x+y}(z) = (z - 2G_{x+y}(z))^{-1},$$

ce qui donne $G_{x+y}(z) = (z - \sqrt{z^2 - 8})/4$. Maintenant, grâce à la formule (2), on peut retrouver la limite de $\mu_{X_N + Y_N}$, la mesure empirique de valeur propre de la somme des deux matrices. Ainsi, cette suite de mesure converge faiblement p.s. vers une mesure μ_{x+y} absolument continue par rapport à la mesure de Lebesgue de densité

$$\rho(t) = \frac{\sqrt{8 - t^2}}{4\pi} \mathbb{1}_{|t| \leq 2\sqrt{2}}.$$

Remarquons que cette nouvelle mesure est simplement une dilatation de la loi semi-circulaire de la même manière que la loi de la somme de deux gaussiennes est encore une gaussienne mais de variance différente.

Visuellement, si l'on trace l'histogramme des valeurs propres de la somme de deux matrices du GUE renormalisées, celui-ci va s'approcher de la courbe d'une loi semi-circulaire de rayon $2\sqrt{2}$.

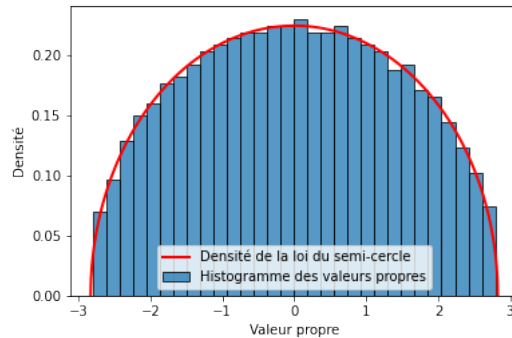


Figure 3: Densité de valeurs propres de $X_N + Y_N$ pour $N = 1000$

4 Conclusion et perspectives

Dans la première partie on a établi un théorème de convergence des mesures spectrales d'une grande classe de matrices aléatoires : les matrices de Wigner. Ce théorème étudie le comportement global au premier ordre du spectre mais on peut également s'intéresser à l'étude des valeurs propres sur les bords du domaine (comportement local) ou à la vitesse de convergence de la mesure spectrale vers la mesure

semi-circulaire (comportement global au second ordre). On peut par exemple consulter [3] pour plus de précisions à ce sujet. On a ensuite vu que les probabilités libres permettent d'étudier la limite de la mesure spectrale de la somme de deux matrices asymptotiquement libres. La question suivante serait alors de caractériser les matrices asymptotiquement libres ou au moins de trouver des conditions suffisantes. Par exemple, on sait que les matrices invariantes en loi par conjugaison par des matrices unitaires sont asymptotiquement libres (pourvu qu'elles convergent en distribution non-commutative). Camille Male [4] a récemment développé une théorie qui étend celle des probabilités libres et qui définit une nouvelle notion de loi, de convergence et d'indépendance : les trafics. Ainsi, dans cette théorie, des matrices invariantes en loi par conjugaison par des matrices de permutation uniforme sont asymptotiquement trafic-libres (pourvu qu'elles convergent en trafic-distribution). On pourrait alors également se poser la question de la vitesse de convergence de telles matrices vers leur limite. C'est en partie cette question qui constitue l'objet de ma thèse.

References

- [1] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*. English. Vol. 118. Camb. Stud. Adv. Math. Cambridge: Cambridge University Press, 2010. ISBN: 978-0-521-19452-5. DOI: 10.1017/CB09780511801334.
- [2] Bordenave Charles. "Lecture notes on random matrix". English. In: (2019).
- [3] Sandrine Dallaporta. "Quelques aspects de l'étude quantitative de la fonction de comptage et des valeurs propres de matrices aléatoires". 2012TOU30181. PhD thesis. 2012, 1 vol. (133 p.) URL: <http://www.theses.fr/2012TOU30181>.
- [4] Camille Male. *Traffic distributions and independence: permutation invariant random matrices and the three notions of independence*. English. Vol. 1300. Mem. Am. Math. Soc. Providence, RI: American Mathematical Society (AMS), 2020. ISBN: 978-1-4704-4298-9; 978-1-4704-6399-1. DOI: 10.1090/memo/1300.
- [5] James A. Mingo and Roland Speicher. *Free probability and random matrices*. English. Vol. 35. Fields Inst. Monogr. Toronto: The Fields Institute for Research in the Mathematical Sciences; New York, NY: Springer, 2017. ISBN: 978-1-4939-6941-8; 978-1-4939-6942-5. DOI: 10.1007/978-1-4939-6942-5.
- [6] Alexandru Nica and Roland Speicher. *Lectures on the combinatorics of free probability*. English. Vol. 335. Lond. Math. Soc. Lect. Note Ser. Cambridge: Cambridge University Press, 2006. ISBN: 0-521-85852-6. DOI: 10.1017/CB09780511735127.