

Le théorème de la loi circulaire

Nathan Noiry *

Mémoire de M2

ENS Lyon, Parcours Probabilités

Encadré par Laurent Ménard et Nathanaël Enriquez

*nathan.noiry@ens-lyon.fr

Remerciements

Je tiens à remercier Laurent Ménard qui m'a initié aux matrices aléatoires pendant mon stage de troisième année de Licence, et donné envie de poursuivre dans cette voie. Ses conseils et explications ont été précieux dans la rédaction de ce document. Merci aussi à Nathanaël Enriquez qui, malgré un agenda chargé, a toujours trouvé le temps de répondre à mes questions. C'est une chance d'être encadré par ces deux chercheurs qui sont un exemple pour moi. Merci de m'avoir fait confiance et proposé ce sujet de mémoire.

Merci à tous les membres du laboratoire Modal'X, pour leur accueil et leur bonne humeur.

Merci aux doctorants, postdoctorants et ATER : Julien, Marie, Paul, Gabriela, Zaid, Emilien, Yan et Moustapha.

Table des matières

1	Énoncé et heuristique	4
1.1	La loi circulaire	4
1.2	Stratégie de preuve : hermitisation	5
1.3	Résultats principaux et organisation du mémoire	6
2	Identification de la limite : le cas gaussien	6
2.1	Loi jointe du spectre	7
2.2	Fonctions de corrélations	10
2.3	Convergence de $\mu_{n^{-1/2}X}$	11
3	Uniforme intégrabilité du logarithme	13
3.1	Réduction du problème	13
3.2	Distance d'un vecteur aléatoire à un sous-espace	15
3.3	Conclusion	16
4	Contrôle presque sûr de $s_n(X + M)$	17
4.1	Problème de Littlewood-Offord	17
4.2	Vecteurs pauvres	18
4.3	Vecteurs riches	19
5	Contrôle en probabilité de $s_n(X + M)$	20
5.1	Une découpe astucieuse de la sphère unité	20
5.2	Vecteurs compressibles	20
5.3	Vecteurs incompressibles	22
6	Convergence des mesures singulières	25
6.1	Réduction du problème	25
6.2	Matrice hermitisée	26
6.3	Résolvante quaternionique	27
7	Preuve alternative : méthode des moments	29
7.1	Quelques notations	29
7.2	Le cas $z = 0$: le quart de cercle	30
7.3	Le cas général	32
7.4	Série génératrice	35
	Annexe A Analyse matricielle	39
	Annexe B Concentration de la mesure	41

1 Énoncé et heuristique

1.1 La loi circulaire

Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice complexe de taille n . On note $\lambda_1(A), \dots, \lambda_n(A)$ les valeurs propres de A , numérotées de sorte que $|\lambda_1(A)| \geq \dots \geq |\lambda_n(A)|$, avec des arguments croissants en cas d'égalité. La mesure spectrale μ_A de A est définie par

$$\mu_A := \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i(A)},$$

où δ_z désigne la mesure de Dirac en z . Par ailleurs, notons $A^* = \bar{A}^t$ la matrice conjuguée puis transposée de A . Comme AA^* est hermitienne, ses valeurs propres sont réelles positives, ce qui permet de définir les valeurs singulières de A comme étant les racines des valeurs propres de $AA^* : s_i(A) := \sqrt{\lambda_i(AA^*)}$ pour tout $1 \leq i \leq n$. La mesure singulière ν_A de A est alors définie par

$$\nu_A := \frac{1}{n} \sum_{i=1}^n \delta_{s_i(A)}.$$

Soit $\{X_{ij}\}_{i,j \geq 1}$ une famille de variables aléatoires complexes i.i.d. de variance égale à 1 : $\text{Var}(X_{11}) = \mathbf{E}|X_{11}|^2 - |\mathbf{E}X_{11}|^2 = 1$. Pour tout $n \geq 1$, on introduit la matrice aléatoire complexe $X_n = (X_{ij})_{1 \leq i,j \leq n}$. Par souci de clarté, on omettra l'indice n . Rappelons qu'une suite de mesures complexes $\{m_n\}_{n \geq 1}$ converge étroitement vers une mesure m lorsque pour toute fonction $f : \mathbb{C} \rightarrow \mathbb{R}$ continue bornée,

$$\int_{\mathbb{C}} f(z) dm_n(z) \xrightarrow{n \rightarrow +\infty} \int_{\mathbb{C}} f(z) dm(z).$$

On note parfois $m_n \xrightarrow{(e)} m$. La loi circulaire désigne la loi uniforme sur le disque unité complexe. Elle est universelle au sens du théorème suivant :

Théorème 1.1 (de la loi circulaire). *Presque sûrement, la suite des mesures spectrales $\{\mu_{n^{-1/2}X}\}_{n \geq 1}$ converge étroitement vers la loi circulaire.*

Avant de donner une idée de la preuve de ce résultat, remarquons que la suite des seconds moments des $\mu_{n^{-1/2}X}$ est bornée. Ceci est dû à la généralisation des inégalités de Weyl donnée en annexe (théorème A.3), qui entraîne que

$$\int_{\mathbb{C}} |\lambda|^2 d\mu_A = \frac{1}{n} \sum_{i=1}^n |\lambda_i(A)|^2 \leq \frac{1}{n} \sum_{i=1}^n s_i(A)^2 = \int_{\mathbb{R}_+} s^2 d\nu_A. \quad (1)$$

En effet, la somme des carrés des valeurs singulières de X étant égale à la trace de XX^* , la loi des grands nombres assure que

$$\int_{\mathbb{R}_+} s^2 d\nu_{n^{-1/2}X} = \frac{1}{n^2} \sum_{1 \leq i,j \leq n} |X_{ij}|^2$$

possède une limite finie. L'inégalité de Tchebychev montre alors que la suite $\{\mu_{n^{-1/2}X}\}_{n \geq 1}$ est tendue.

1.2 Stratégie de preuve : hermitisation

Le terme hermitisation désigne un procédé permettant de se ramener à un cadre hermitien, en remplaçant l'étude des mesures spectrales $\mu_{n-1/2X}$ par l'étude de la famille des mesures singulières $\{\nu_{n-1/2X-zI}, z \in \mathbb{C}\}$. Cette idée remonte aux travaux de Girko [10] ; on la donne ici sous forme de lemme.

Lemme 1.1. *Supposons qu'il existe une famille de lois $\{\nu_z\}_{z \in \mathbb{C}}$, chacune ne dépendant que de $z \in \mathbb{C}$, telle que presque sûrement, pour presque tout $z \in \mathbb{C}^1$, on ait*

1. $\nu_{n-1/2X-zI} \xrightarrow{(e)} \nu_z$;
2. \log est uniformément intégrable pour $\{\nu_{n-1/2X-zI}\}_{n \geq 1}$.

Alors la suite $\{\mu_{n-1/2X}\}_{n \geq 1}$ converge p.s. étroitement vers la loi circulaire.

L'objet sous-jacent au lemme 1.1 est le potentiel logarithmique d'une mesure complexe m , qui est la fonction :

$$U_m : z \in \mathbb{C} \mapsto - \int_{\mathbb{C}} \log |\lambda - z| dm(\lambda).$$

Comme $(2\pi)^{-1} \log |\cdot|$ est une solution fondamentale du laplacien en dimension deux, on a la formule $\Delta U_m = -2\pi m$ au sens des distributions, donc U_m caractérise la mesure m .

Le point de départ dans la preuve du lemme 1.1 est la formule suivante :

$$U_{\mu_{n-1/2X}}(z) = - \int_{\mathbb{R}_+} \log(s) d\nu_{n-1/2X-zI}(s), \quad (2)$$

qui est une conséquence de l'égalité

$$\int_{\mathbb{C}} \log |\lambda| d\mu_A = \log \left(\prod_{i=1}^n |\lambda_i(A)| \right)^{1/n} = \log |\det A|^{1/n} = \int_{\mathbb{R}_+} \log(s) d\nu_A. \quad (3)$$

On a vu précédemment que la suite $\{\mu_{n-1/2X}\}_{n \geq 1}$ est tendue. Notons μ un point d'accumulation. Si les hypothèses 1 et 2 du lemme 1.1 sont vérifiées, le lemme 4.5 de [4] assure que le logarithme est uniformément intégrable contre la suite de mesures $\{\mu_{n-1/2X}\}_{n \geq 1}$. On peut donc passer à la limite dans (2). Ceci fournit une caractérisation du potentiel logarithmique U_μ :

$$U_\mu(z) = - \int_{\mathbb{R}_+} \log(s) d\nu_z(s).$$

Ainsi la suite des mesures spectrales admet un unique point d'accumulation, et ce dernier est entièrement caractérisé par les mesures $\{\nu_z\}_{z \in \mathbb{C}}$. Comme ces mesures *ne dépendent pas de la loi des coefficients X_{ij}* , la mesure limite est universelle : il suffit de considérer une loi particulière de variance égale à 1 sur les X_{ij} pour l'identifier. La loi privilégiée est sans surprise la gaussienne complexe standard, pour laquelle la limite circulaire a été identifiée par Mehta (c.f. par exemple [14, Chapitre 15]), grâce à la connaissance de la loi jointe du spectre. La partie 2 de ce mémoire s'intéresse à cette preuve. On y rend en particulier rigoureuse l'obtention de la loi jointe du spectre.

1. au sens de la mesure de Lebesgue.

1.3 Résultats principaux et organisation du mémoire

Pour montrer le théorème de la loi circulaire, il faut vérifier les hypothèses 1 et 2 du lemme 1.1. Le premier point est un corollaire du théorème 1.1 de [7], stipulant que si $\{R_n\}_{n \geq 1}$ est une suite de matrices déterministes telles que la suite des mesures singulières ν_{R_n} converge étroitement vers une mesure ν_∞ , alors $\nu_{n^{-1/2}X - R_n}$ converge p.s. étroitement vers une loi limite ne dépendant que de ν_∞ . Le cas $R_n = zI_n$ (I_n désignant la matrice identité) constitue le

Lemme 1.2. *Pour tout $z \in \mathbb{C}$, il existe une loi ν_z ne dépendant que de z telle que presque sûrement :*

$$\nu_{n^{-1/2}X - zI} \xrightarrow{(e)} \nu_z.$$

La preuve proposée dans [4] met en place une méthode de résolvante innovante, baptisée résolvante quaternionique par les auteurs. Après en avoir présenté les idées dans la partie 6, on donnera une nouvelle manière d'obtenir ce résultat par la méthode des moments, dans la dernière partie 7.

Le second point est plus délicat. L'uniforme intégrabilité du logarithme requiert en effet un contrôle non trivial de la plus petite valeur singulière des matrices $n^{-1/2}X - zI$. Le résultat le plus général dans cette direction a été obtenu par Tao et Vu dans [16].

Lemme 1.3. *Pour tout $A, C_1 > 0$, il existe $B, C_2 > 0$ tels que pour toute matrice complexe M de taille n vérifiant $s_1(M) \leq n^{C_1}$, on ait la majoration :*

$$\mathbf{P}(s_n(X + M) \leq n^{-B}) \leq C_2 n^{-A}.$$

Une autre version du théorème de la loi circulaire, où la convergence presque sûre est remplacée par une convergence en probabilité, est obtenue grâce à une estimation moins fine, possédant un intérêt propre :

Lemme 1.4. *Il existe deux constantes $c, r > 0$ telles que pour tout $n \geq c$ et toute matrice M de taille n :*

$$\mathbf{P}\left(s_n(X + M) \leq \frac{1}{n^{2r+1}}, s_1(X + M) \leq n^r\right) \leq 2c \sqrt{\frac{\log(cn^r)}{n}}.$$

Avant de donner une preuve de ces lemmes dans les parties 4 et 5, nous verrons comment ces derniers permettent effectivement d'obtenir l'uniforme intégrabilité du logarithme contre la suite de mesures $\{\nu_{n^{-1/2}X - zI}\}_{n \geq 1}$ dans la partie 3.

La preuve complète du théorème de la loi circulaire nécessiterait un document beaucoup plus conséquent. Plutôt que d'effectuer tous les détails des preuves, on en donne souvent les idées principales, et l'on admettra parfois des résultats techniques.

2 Identification de la limite : le cas gaussien

Comme nous l'avons expliqué dans la partie 1.2, le lemme 1.1 est valable sous l'unique hypothèse $\text{Var}(X_{11}) = 1$. Ceci implique que la limite μ des mesures spectrales $\mu_{n^{-1/2}X}$ est universelle. Dans cette partie nous considérons le

cas particulier des lois gaussiennes afin d'identifier μ qui n'est autre que la loi uniforme sur le disque unité.

L'ensemble de Ginibre est défini comme l'ensemble des matrices aléatoires complexes $X = (X_{ij})_{1 \leq i, j \leq n}$ possédant des coefficients i.i.d. de loi gaussienne standard sur \mathbb{C} . On identifiera une matrice de taille n et un vecteur de \mathbb{R}^{2n^2} , de sorte que la densité de la loi induite par rapport à la mesure de Lebesgue est

$$A \in \mathcal{M}_n(\mathbb{C}) \mapsto \frac{1}{\pi^{n^2}} \exp\left(-\operatorname{Tr}(AA^*)\right). \quad (4)$$

Fixons X une matrice de taille n appartenant à l'ensemble de Ginibre.

2.1 Loi jointe du spectre

La loi jointe des valeurs propres de X est connue depuis Ginibre [9]. Pour tout $z = (z_1, \dots, z_n) \in \mathbb{C}^n$, on note $\Delta(z) = \prod_{1 \leq i < j \leq n} (z_i - z_j)$ le déterminant de Vandermonde associé. Alors :

Théorème 2.1. *La loi du n -uplet $(\lambda_1(X), \dots, \lambda_n(X))$ est absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^{2n} et sa densité est donnée par*

$$(z_1, \dots, z_n) \mapsto \left(\pi^{n^2} \prod_{j=1}^{n-1} j!\right)^{-1} |\Delta(z)|^2 \exp\left(-\sum_{1 \leq k \leq n} |z_k|^2\right) \mathbf{1}_{|z_1| > \dots > |z_n|}.$$

Bien que ce résultat soit classique, je n'ai trouvé aucune référence donnant une preuve complète. Les arguments présentés ici sont inspirés de [6] qui traite le cas hermitien et de [14]. Pour obtenir la loi jointe du spectre, l'idée est d'isoler les variables spectrales, et d'intégrer la densité (4) contre les variables non spectrales. Comme l'ensemble des matrices complexes possédant des valeurs propres ayant le même module est négligeable pour la mesure de Lebesgue, on peut raisonner sur l'ensemble des matrices

$$\operatorname{GL}_n^>(\mathbb{C}) = \{A \in \operatorname{GL}_n(\mathbb{C}) : |\lambda_1(A)| > \dots > |\lambda_n(A)|\}.$$

Notons par ailleurs $\mathcal{T}_n^>$ l'ensemble des matrices triangulaires supérieures de taille n dont la diagonale est formée de complexes de module strictement décroissant, et \mathcal{U}_n l'ensemble des matrices unitaires de taille n . La décomposition de Schur est l'application surjective

$$\begin{aligned} \mathcal{S} : \mathcal{U}_n \times \mathcal{T}_n^> &\longrightarrow \operatorname{GL}_n^>(\mathbb{C}) \\ (U, T) &\longmapsto UTU^*. \end{aligned}$$

Malheureusement, \mathcal{S} n'est pas injective. Plus précisément, (U_1, T_1) et (U_2, T_2) ont la même image par \mathcal{S} si et seulement s'il existe une matrice diagonale $R \in \mathcal{U}_n$ telle que $U_2 = U_1 R$ et $T_2 = R^* T_1 R$.

En effet, le sens réciproque est immédiat et pour l'implication, remarquons que les vecteurs propres de $\mathcal{S}(U_1, T_1)$ sont les images des vecteurs propres de T_1 par U_1 . Comme le i -ième vecteur propre de T_1 s'écrit sous la forme ${}^t(a_1, \dots, a_{i-1}, 1, 0, \dots, 0)$, pour des complexes a_j bien choisis, on en déduit que si u_i désigne la i -ème colonne de U_1 , alors u_i appartient à l'espace vectoriel engendré par les i premiers vecteurs propres de $\mathcal{S}(U_1, T_1)$. Mais comme U_1 est

unitaire, la famille $(u_i)_{1 \leq i \leq n}$ est obtenue en appliquant le procédé d'orthogonalisation de Gram-Schmidt aux vecteurs propres de $\mathcal{S}(U_1, T_1)$. Notons que l'ordre des vecteurs propres est fixé puisqu'on raisonne sur $\mathrm{GL}_n^>(\mathbb{C})$. Ainsi, si $\mathcal{S}(U_1, T_1) = \mathcal{S}(U_2, T_2)$, les colonnes respectives de U_1 et U_2 ne peuvent différer que d'une phase, autrement dit il existe une matrice diagonale unitaire R telle que $U_2 = U_1 R$, et partant, $T_2 = R^* T_1 R$.

Pour obtenir une bijection, il faut considérer l'espace des matrices unitaires quotienté par le sous-groupe des matrices unitaires diagonales \mathcal{U}_1^n . Le prix à payer est que l'on va devoir raisonner localement. Notons $p : \mathcal{U}_n \rightarrow \mathcal{U}_n/\mathcal{U}_1^n$ la projection canonique, qui est lisse pour la topologie quotient. On admettra le résultat suivant, dont une preuve est donnée dans [18, Théorème 3.58].

Lemme 2.1. *La projection p admet des sections locales lisses. Autrement dit pour tout $[U] \in \mathcal{U}_n/\mathcal{U}_1^n$, il existe un voisinage \mathcal{V} de $[U]$ et une application lisse $\tau : \mathcal{V} \rightarrow \mathcal{U}_n$ telle que $p \circ \tau = \mathrm{id}$.*

Soit $[U_0] \in \mathcal{U}_n/\mathcal{U}_1^n$. Alors le lemme 2.1 assure l'existence d'un voisinage ouvert \mathcal{V}_0 de $[U_0]$ et d'une section lisse $\tau_0 : \mathcal{V}_0 \rightarrow \mathcal{U}_n$. Remarquons que l'application

$$\begin{aligned} \mathcal{V}_0 \times \mathcal{T}_n^> &\longrightarrow \mathrm{GL}_n^>(\mathbb{C}) \\ ([U], T) &\longmapsto \mathcal{S}(\tau_0([U]), T) \end{aligned}$$

est injective. En effet, pour tout $[U] \in \mathcal{U}_n/\mathcal{U}_1^n$, l'image $\tau_0([U])$ est entièrement déterminée par le choix de la section. Par exemple, on peut choisir τ_0 telle que la première coordonnée non nulle de chaque colonne de $\tau_0([U])$ est réelle positive. Ainsi, comme $\tau_0([U])T\tau_0([U])^* = \tau_0([U]')T'\tau_0([U]')^*$ implique que $\tau_0([U]') = \tau_0([U])R$ pour une matrice R unitaire diagonale, on en déduit que R ne peut être que l'identité.

L'espace $\mathcal{U}_n/\mathcal{U}_1^n$ est compact, et c'est une variété (réelle) de dimension $n^2 - n$. Soit $\{(\mathcal{V}_i, c_i)\}_{1 \leq i \leq L}$ une famille finie de cartes locales telles que les ouverts \mathcal{V}_i recouvrent $\mathcal{U}_n/\mathcal{U}_1^n$. Notons τ_i la section locale associée à l'ouvert \mathcal{V}_i , qui existe quitte à choisir des ouverts plus petits. Alors il existe des paramétrisations locales de $\mathrm{GL}_n^>(\mathbb{C})$:

$$\begin{aligned} \Phi_i : \mathbb{C}_>^n \times \mathbb{C}^{(n^2-n)/2} \times \mathcal{O}_i &\longrightarrow \mathrm{GL}_n^>(\mathbb{C}) \\ (z, t, u) &\longmapsto \mathcal{S}(\tau_i \circ c_i^{-1}(u), T(z, t)), \end{aligned}$$

où les $\mathcal{O}_i = c_i(\mathcal{V}_i)$ sont des ouverts de $\mathbb{C}^{(n^2-n)/2}$, et $\mathbb{C}_>^n$ l'ensemble des n -uplet de complexes (z_1, \dots, z_n) ayant des modules strictement décroissants. Notons que l'on a fait apparaître les variables spectrales (c'est la composante $\mathbb{C}_>^n$), en paramétrant l'ensemble $\mathcal{T}_n^>$ de la manière suivante :

$$T : (z, t) \longmapsto \begin{pmatrix} z_1 & t_{12} & \cdots & t_{1n} \\ 0 & z_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{n-1,n} \\ 0 & \cdots & 0 & z_n \end{pmatrix}.$$

Le théorème 2.18 de [13] assure qu'il existe une partition de l'unité $\{\rho_i\}_{1 \leq i \leq L}$ subordonnée au recouvrement $\{\mathrm{Im}\Phi_i\}_{1 \leq i \leq L}$. Soit $F : \mathrm{GL}_n^>(\mathbb{C}) \rightarrow \mathbb{R}$ une fonction mesurable positive, ne dépendant que des valeurs spectrales. On notera toujours

F l'application induite sur les n -uplets de complexes. La formule du changement de variables entraîne que

$$\int F(A) e^{-\text{Tr}(AA^*)} dA = \int F(A) e^{-\text{Tr}(A^*A)} \sum_{1 \leq i \leq L} \rho_i(A) dA$$

se réécrit

$$\sum_{1 \leq i \leq L} \int F(z) e^{-\sum_{1 \leq i \leq n} |z_i|^2 - \sum_{i < j} |t_{ij}|^2} |\text{Jac}_{\Phi_i}(z, t, u)| \rho_i(\Phi_i(z, t, u)) dz dt du. \quad (5)$$

On est donc ramené à un problème local. On fixe un indice i et on omet les indices sur les applications dans ce qui suit. Notre but est de calculer le jacobien de Φ , en espérant pouvoir l'écrire en séparant la variable spectrale z des autres variables t et u . Pour faire apparaître un tel résultat, on va se donner une base particulière de $\mathcal{M}_n(\mathbb{C})$. Pour toute paire d'indices (k, l) , on note E_{kl} la matrice possédant des zéros partout, et un 1 en (k, l) : $E_{kl}(i, j) = \delta_{(k,l)}(i, j)$. Voici l'ordre particulier selon lequel on ordonne cette base :

$$(E_{11}; \dots; E_{nn}; E_{12}; E_{13}; \dots; E_{1n}; E_{23}; \dots; E_{2n}; \dots; E_{n-2, n-1}; E_{n-2, n}; E_{n-1, n}; \\ E_{n1}; E_{n2}; \dots; E_{n, n-1}; E_{n-1, 1}; \dots; E_{n-1, n-2}; \dots; E_{31}; E_{32}; E_{21}). \quad (6)$$

Notons $U = \tau \circ c^{-1}$ pour simplifier les calculs. Alors les dérivées partielles en z et en t sont particulièrement simples :

$$\frac{\partial}{\partial z_i}(UTU^*) = U \frac{\partial T}{\partial z_i} U^* \quad \text{et} \quad \frac{\partial}{\partial t_{ij}}(UTU^*) = U \frac{\partial T}{\partial t_{ij}} U^*. \quad (7)$$

Par ailleurs, si $u = (u_1, \dots, u_{(n^2-n)/2})$, alors pour tout $i \in \{1, \dots, (n^2-n)/2\}$:

$$\frac{\partial}{\partial u_i}(UTU^*) = \frac{\partial U}{\partial u_i} T U^* + U T \frac{\partial U^*}{\partial u_i}.$$

Le fait que $UU^* = I_n$ mène alors à la formule suivante :

$$\frac{\partial}{\partial u_i}(UTU^*) = U \left(U^* \frac{\partial U}{\partial u_i} T - T U^* \frac{\partial U}{\partial u_i} \right) U^*. \quad (8)$$

Comme $\det(UU^*) = 1$, le jacobien de Φ est le déterminant de la matrice par blocs suivante, écrite dans le choix de base (6) :

$$\begin{pmatrix} I_n & 0 & * \\ 0 & I_{(n^2-n)/2} & * \\ 0 & 0 & M \end{pmatrix},$$

où la i -ème colonne de la matrice M est formée des

$$\left(U^* \frac{\partial}{\partial u_i}(UTU^*) U \right)_{kl},$$

pour des indices $k > l$ ordonnés selon notre choix de base.

On notera $W^{(i)} = U^* \partial U / \partial u_i$ dans ce qui suit. La relation (8) entraîne que pour tout $i \in \{1, \dots, (n^2 - n)/2\}$ et toute paire d'indices $k > l$:

$$\left(U^* \frac{\partial}{\partial u_i} (UTU^*) U \right)_{kl} = \sum_{j < l} W_{kj}^{(i)} T_{jl} + \sum_{j > k} T_{kj} W_{jl}^{(i)} + (T_{ll} - T_{kk}) W_{kl}^{(i)}.$$

Soient $k > l$. Un moment de réflexion montre que la "(k, l)-ème" ligne de la matrice M – l'expression fait sens par notre choix de base (6) – est le produit d'une matrice ligne de dépendant que de T , dont les coefficients sont nuls à partir de la "(k, l)-ième +1" coordonnée, par la matrice dont la i -ème colonne est $(W_{pq}^{(i)})$, avec $p > q$ ordonnées selon (6). On en déduit que M est le produit d'une matrice triangulaire inférieure de diagonale $(T_{ll} - T_{kk}, k > l)$ et d'une matrice ne dépendant que des $W^{(i)}$, donc que de la variable non spectrale u . Ainsi, le jacobien *réel* de Φ s'écrit :

$$\text{Jac}_{\Phi}(z, t, u) = |\Delta(z)|^2 J(u),$$

où J est une fonction ne dépendant que de u .

Les calculs précédents ne dépendaient pas de la paramétrisation locale Φ_i . En reprenant notre calcul (5), on peut séparer nos trois variables et intégrer contre (t, u) . Il existe alors une constante strictement positive C_n telle que

$$\int F(A) e^{-\text{Tr}(AA^*)} dA = C_n \int_{\mathbb{C}^n} F(z) |\Delta(z)|^2 e^{-\sum_{1 \leq i \leq n} |z_i|^2} dz.$$

Pour l'évaluation de C_n , on pourra se référer à [14, Chapitre 15] pour se convaincre que

$$C_n = \left(\pi^{n^2} \prod_{j=1}^{n-1} j! \right)^{-1}.$$

2.2 Fonctions de corrélations

Notons ϕ_n la densité de la loi jointe des valeurs propres de X (non ordonnées). Par le théorème 2.1, on a la formule explicite

$$\phi_n(z_1, \dots, z_n) = \frac{1}{n!} C_n |\Delta(z)|^2 \exp \left(- \sum_{1 \leq k \leq n} |z_k|^2 \right).$$

En intégrant cette dernière contre les $n - k$ dernières variables, on obtient :

$$\phi_{n,k}(z_1, \dots, z_k) := \int_{\mathbb{C}^{n-k}} \phi_n(z_1, \dots, z_n) dz_{k+1} \cdots dz_n.$$

On parle des fonctions de corrélations des valeurs propres de X . Pour les calculer, on réécrit ϕ_n en faisant apparaître un déterminant :

$$\phi_n(z_1, \dots, z_n) = \frac{1}{n!} C_n \exp \left(- \sum_{k=1}^n |z_k|^2 \right) \det (z_k^{i-1})_{1 \leq i, k \leq n} \det (\bar{z}_k^{i-1})_{1 \leq i, k \leq n}.$$

Par combinaisons linéaires entre les lignes de $(z_k^{i-1})_{1 \leq i, k \leq n}$, pour toute famille de polynômes $\{P_i\}_{1 \leq i \leq n}$ tels que P_i est de degré i , il existe une constante C_P telle que

$$\det (P_{i-1}(z_k))_{1 \leq i, k \leq n} = C_P \det (z_k^{i-1})_{1 \leq i, k \leq n}.$$

Ceci implique que ϕ_n s'écrit

$$\frac{1}{n!} C_n C_P^{-2} \det (P_{i-1}(z_k) e^{|z_k|^2/2})_{1 \leq i, k \leq n} \det (\overline{P_{i-1}(z_k)} e^{|z_k|^2/2})_{1 \leq i, k \leq n}.$$

On fait le choix $P_i(z) = \frac{1}{\sqrt{\pi i!}} z^i$, de sorte que l'on ait $C_P = \sqrt{C_n}$ et

$$\phi_n(z_1, \dots, z_n) = \frac{1}{n!} \det (K_n(z_i, z_j))_{1 \leq i, j \leq n},$$

où

$$K_n(x, y) := \sum_{i=1}^{n-1} P_i(x) \overline{P_j(y)} e^{-(|x|^2 + |y|^2)/2}.$$

La famille des polynômes P_i est orthonormale pour la mesure $e^{-|z|^2} dz$. Ceci implique les deux relations suivantes :

$$\begin{cases} \int_{\mathbb{C}} K_n(x, x) dx = n \\ K_n(x, y) = \int_{\mathbb{C}} K_n(x, u) K_n(u, y) du. \end{cases}$$

Nous admettons que ces deux identités suffisent à montrer le lemme suivant, dont on trouvera une preuve dans [15, Lemme 2.6.5].

Lemme 2.2. *Pour tout $n \geq 1$ et tout $0 \leq k \leq n - 1$:*

$$\int_{\mathbb{C}} \det (K_n(z_i, z_j))_{1 \leq i, j \leq k+1} dz_{k+1} = (n - k) \det (K_n(z_i, z_j))_{1 \leq i, j \leq k}.$$

Définissons la mesure moyennes $\mathbf{E}\mu_X$ comme l'unique mesure de probabilité telle que pour toute fonction $f : \mathbb{C} \rightarrow \mathbb{R}$ mesurable, on ait

$$\mathbf{E} \int_{\mathbb{C}} f d\mu_X = \int_{\mathbb{C}} f d\mathbf{E}\mu_X.$$

Le lemme 2.2 implique que $\mathbf{E}\mu_X$ a pour densité

$$\phi_{n,1}(z) = \frac{1}{n\pi} e^{-|z|^2} \sum_{i=1}^{n-1} \frac{|z|^{2i}}{i!}.$$

Un changement de variable montre que la densité de $\mathbf{E}\mu_{n-1/2X}$ est $n\phi_{n,1}(\sqrt{n}z)$. Il est possible de montrer que cette application converge uniformément sur tout compact vers $\pi^{-1} \mathbf{1}_{|z| \leq 1}$. On en déduit le théorème suivant :

Théorème 2.2. *La suite $\mathbf{E}[n^{-1/2}X]$ vérifie le théorème de la loi circulaire, autrement dit les mesures $\mathbf{E}\mu_{n-1/2X}$ convergent étroitement vers la loi circulaire.*

2.3 Convergence de $\mu_{n-1/2X}$

Il nous reste à montrer que $\mu_{n-1/2X}$ possède la même limite que sa moyenne $\mathbf{E}\mu_{n-1/2X}$. Il suffit de montrer que presque sûrement, pour toute fonction $f : \mathbb{C} \rightarrow \mathbb{R}$ continue à support compact,

$$\left| \int_{\mathbb{C}} f d\mu_{n-1/2X} - \int_{\mathbb{C}} f d\mathbf{E}\mu_{n-1/2X} \right| \xrightarrow{n \rightarrow +\infty} 0.$$

Comme l'ensemble des fonctions continues à support compact de \mathbb{C} dans \mathbb{R} muni de la norme $\|\cdot\|_\infty$ est séparable, il suffit de raisonner sur l'un de ses éléments f que l'on fixe dans ce qui suit. Notons $S_n = \int_{\mathbb{C}} f d\mu_{n^{-1/2}X}$ et supposons que

$$\mathbf{E}\left[(S_n - \mathbf{E}S_n)\right] = O(n^{-2}). \quad (9)$$

Dans ce cas on a le résultat cherché puisque le théorème de Fubini-Tonelli fournit

$$\mathbf{E} \sum_{n \geq 0} (S_n - \mathbf{E}S_n)^4 < \infty,$$

ce qui implique que presque sûrement, $|S_n - \mathbf{E}S_n|$ converge vers 0 lorsque n tend vers l'infini. Pour montrer l'estimation (9), on pose $Z_i = f(\lambda_i(n^{-1/2}X)) - \mathbf{E}f(\lambda_i(n^{-1/2}X))$ et on développe

$$\mathbf{E}[(S_n - \mathbf{E}S_n)^4] = \frac{1}{n^4} \mathbf{E} \left[\left(\sum_{i=1}^n Z_i \right)^4 \right]$$

en distinguant le nombre d'indices différents qui apparaissent dans les termes développés. Lorsqu'on regarde les termes où au plus deux indices interviennent, on obtient la contribution

$$\frac{1}{n^4} \sum_{1 \leq i_1 \leq n} \mathbf{E}[Z_{i_1}^4] + \frac{1}{n^4} \sum_{1 \leq i_1, i_2 \leq n} \left(4\mathbf{E}[Z_{i_1} Z_{i_2}^3] + 3\mathbf{E}[Z_{i_1}^2 Z_{i_2}^2] \right)$$

qui est un $O(n^{-2})$ puisque pour tout $1 \leq i \leq n$ on a $|Z_i| \leq 2\|f\|_\infty$.

La contribution où trois indices interviennent est

$$\frac{6}{n^4} \sum_{1 \leq i_1, i_2, i_3 \leq n} \mathbf{E}[Z_{i_1} Z_{i_2} Z_{i_3}^2].$$

Par ce qu'on a fait sur les fonctions de corrélations dans la partie 2.2, on a

$$\mathbf{E}[Z_{i_1} Z_{i_2} Z_{i_3}^2] = \int_{\mathbb{C}^3} \prod_{j=1}^3 f(z_{i_j}) \widetilde{\phi}_{n,3}(z_{i_1}, z_{i_2}, z_{i_3}) dz_{i_1} dz_{i_2} dz_{i_3},$$

où

$$\widetilde{\phi}_{n,3}(z_{i_1}, z_{i_2}, z_{i_3}) = n^3 \phi_{n,3}(\sqrt{n}z_{i_1}, \sqrt{n}z_{i_2}, \sqrt{n}z_{i_3}).$$

Une conséquence de l'inégalité d'Hadamard est que

$$\left| \det \left(\frac{(\sqrt{n}z_i)^{k-1}}{\sqrt{\pi}(k-1)!} e^{-|z_i|^2/2} \right)_{1 \leq i, k \leq 3} \right|^2 \leq n^3 \prod_{i=1}^3 \phi_{n,1}(\sqrt{n}z_i).$$

On en déduit la majoration

$$\begin{aligned} \phi_{n,3}(\sqrt{n}z_{i_1}, \sqrt{n}z_{i_2}, \sqrt{n}z_{i_3}) &= \frac{(n-3)!}{n!} \det (K_n(\sqrt{n}z_{i_k}, \sqrt{n}z_{i_l}))_{1 \leq k, l \leq 3} \\ &\leq \frac{(n-3)!}{n!} n^3 \prod_{j=1}^3 \phi_{n,1}(\sqrt{n}z_{i_j}). \end{aligned}$$

En réécrivant

$$\begin{aligned} \widetilde{\phi}_{n,3}(z_{i_1}, z_{i_2}, z_{i_3}) &= \frac{(n-3)!}{n!} n^3 \prod_{j=1}^3 \widetilde{\phi}_{n,1}(\sqrt{n}z_{i_j}) \\ &\quad - \left(\frac{(n-3)!}{n!} n^3 \prod_{j=1}^3 \widetilde{\phi}_{n,1}(\sqrt{n}z_{i_j}) - \widetilde{\phi}_{n,3}(z_{i_1}, z_{i_2}, z_{i_3}) \right), \end{aligned}$$

on obtient le résultat car :

$$|\mathbf{E}[Z_{i_1} Z_{i_2} Z_{i_3}^2]| \leq \frac{1}{n^4} n^2 (2\|f\|_\infty)^4 + \frac{1}{n^4} n^3 (2\|f\|_\infty)^4 \left(\frac{(n-3)!}{n!} n^3 - 1 \right)$$

est un $O(n^{-2})$.

Pour terminer nous admettons que la dernière contribution, faisant intervenir quatre indices :

$$\frac{1}{n^4} \sum_{1 \leq i_1, i_2, i_3, i_4 \leq n} \mathbf{E}[Z_{i_1} Z_{i_2} Z_{i_3} Z_{i_4}],$$

est bien un $O(n^{-2})$. Une preuve pourra être trouvée dans [12].

3 Uniforme intégrabilité du logarithme

On donne ici une preuve du deuxième point du lemme 1.1. On admettra pour cela le lemme 1.3. La version “en probabilité” de ce résultat peut s’obtenir de manière similaire, en utilisant l’estimation du lemme 1.4.

3.1 Réduction du problème

Rappelons que le but est de montrer que presque sûrement, pour presque tout $z \in \mathbb{C}$, on a

$$\lim_{r \rightarrow \infty} \lim_{n \rightarrow \infty} \int_{|\log t| > r} |\log t| d\nu_{n^{-1/2}X-zI}(t) = 0. \quad (10)$$

Les valeurs propres des matrices $n^{-1/2}X$ forment un ensemble dénombrable, donc pour presque tout $z \in \mathbb{C}$, p.s. z n’est pas une valeur propre de $n^{-1/2}X$ i.e. 0 n’est pas chargé par $\nu_{n^{-1/2}X-zI}$. Ceci implique qu’il suffit de montrer que

$$\lim_{r \rightarrow \infty} \limsup_{n \rightarrow \infty} \int_{|\log t| > r} |\log t| d\nu_{n^{-1/2}X-zI}(t) = 0. \quad (11)$$

En effet supposons que (11) est vérifiée. Soit $\varepsilon > 0$. Il existe $r_0 > 0$ et $n_0 \geq 1$ tels que pour tout $n \geq n_0$,

$$\int_{|\log t| > r_0} |\log t| d\nu_{n^{-1/2}X-zI}(t) \leq \varepsilon.$$

Par ailleurs, le raisonnement précédent implique que

$$\max_{1 \leq n \leq n_0} \int_{|\log t| > r_0} |\log t| d\nu_{n^{-1/2}X-zI}(t) < \infty,$$

on a donc (10) quitte à augmenter r_0 . Comme

$$\begin{aligned} & \int_{|\log t| > r} |\log t| d\nu_{n^{-1/2}X - zI}(t) \\ &= \int_{e^r}^{\infty} |\log t| d\nu_{n^{-1/2}X - zI}(t) + \int_0^{e^{-r}} |\log t| d\nu_{n^{-1/2}X - zI}(t), \end{aligned}$$

et que si $t > 0$ est suffisamment grand (resp. suffisamment petit) $\log s < s^p$ (resp. $|\log s| < s^{-p}$) pour tout exposant $p > 0$, on en déduit qu'il suffit de montrer qu'il existe $p > 0$ tel que

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n s_i(n^{-1/2}X - zI)^p < \infty \quad (12)$$

et

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n s_i(n^{-1/2}X - zI)^{-p} < \infty. \quad (13)$$

Remarquons que (12) est aisément obtenue pour $p = 2$ puisque

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n s_i(n^{-1/2}X - zI)^2 &\leq \frac{1}{n} \sum_{i=1}^n (s_i(n^{-1/2}X) + |z|)^2 \\ &\leq \frac{4}{n} \sum_{i=1}^n s_i(n^{-1/2}X)^2 + |z|^2 \\ &= \frac{4}{n^2} \sum_{1 \leq i, j \leq n} |X_{ij}|^2 + 4|z|^2 < \infty \end{aligned} \quad (14)$$

par la loi des grands nombres (on a utilisé le résultat (A.4) dans la première inégalité). Par ailleurs, le lemme 1.3 permet de majorer la partie de la somme (13) correspondant aux petites valeurs singulières. En effet, comme $-\sqrt{n}zI$ vérifie les hypothèses du lemme 1.3, en prenant $A > 1$ et le $B > 0$ correspondant, le lemme de Borel-Cantelli implique qu'il existe $b = B + 1/2 > 0$ tel que presque sûrement, pour tout n suffisamment grand, $s_n(n^{-1/2}X - zI) > n^{-b}$. Ainsi, pour tout $\gamma > 0$, p.s. la somme

$$\frac{1}{n} \sum_{i=n-n^{1-\gamma}+1}^n s_i(n^{-1/2}X - zI)^{-p} \leq n^{pb-\gamma} \quad (15)$$

tend vers 0 lorsque n tend vers l'infini dès que $p < \gamma/b$. Pour obtenir l'uniforme intégrabilité, il suffit donc de montrer que

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{n-n^{1-\gamma}} s_i(n^{-1/2}X - zI)^{-p} < \infty, \quad (16)$$

pour un $\gamma > 0$ et un $p < \gamma/b$ bien choisis.

3.2 Distance d'un vecteur aléatoire à un sous-espace

L'ingrédient essentiel pour montrer l'inégalité (16) est le contrôle de la distance d'une ligne de la matrice X au sous-espace engendré par certaines autres lignes. Comme les coefficients de X sont indépendants, ce problème revient à contrôler la distance d'un vecteur aléatoire à un sous-espace déterministe. Notons L_i les lignes de la matrice X , et d la distance euclidienne. Le sous-espace vectoriel engendré par des vecteurs v_1, \dots, v_l sera noté $\langle v_1, \dots, v_l \rangle$.

Lemme 3.1. *Il existe $\gamma > 0$ et $\delta > 0$ tels que pour tout n suffisamment grand et tout sous-espace vectoriel H de \mathbb{C}^n tel que $1 \leq \dim H \leq n - n^{1-\gamma}$, on ait*

$$\forall i \in \{1, \dots, n\}, \quad \mathbf{P} \left(d(L_i, H) \leq \frac{1}{2} \sqrt{n - \dim H} \right) \leq \exp(-n^\delta).$$

Preuve. La démonstration de ce lemme est basée sur l'inégalité de concentration de Talagrand (B.3). Pour pouvoir l'appliquer, il faut se ramener au cas de variables bornées, en effectuant un conditionnement judicieux. D'abord, remarquons que comme $d(L_i, H) \geq d(L_i, \langle H, \mathbf{E}L_i \rangle) = d(L_i - \mathbf{E}L_i, \langle H, \mathbf{E}L_i \rangle)$ et $\dim H \leq \dim \langle H, \mathbf{E}L_i \rangle + 1$, on peut supposer que $\mathbf{E}L_i = 0$ sans perdre de généralité. Dans ce cas $\mathbf{E}|X_{ij}|^2 = 1$ et l'inégalité de Markov entraîne que pour tout $\varepsilon > 0$:

$$\mathbf{E} \left[\sum_{j=1}^n \mathbf{1}_{|X_{ij}| \leq n^\varepsilon} \right] \geq n - n^{1-2\varepsilon}.$$

En notant $S_n = \sum_{j=1}^n \mathbf{1}_{|X_{ij}| \leq n^\varepsilon}$, on en déduit que

$$\begin{aligned} \mathbf{P}(S_n \leq n - n^{1-\varepsilon}) &\leq \mathbf{P}(S_n - \mathbf{E}S_n \leq n(n^{-2\varepsilon} - n^{-\varepsilon})) \\ &\leq \exp(-2n(n^{-2\varepsilon} - n^{-\varepsilon})^2), \end{aligned}$$

la deuxième inégalité étant une application de l'inégalité de Hoeffding (B.1). On en déduit que l'évènement "il existe $m = n - n^{1-\varepsilon}$ coefficients de la i -ème ligne dont la valeur absolue est majorée par n^ε " a une probabilité tendant vers 1 lorsque n tend vers l'infini. On peut donc raisonner conditionnellement à ce dernier. Par symétrie il suffit de raisonner conditionnellement à :

$$E_n = \{|X_{ij}| \leq n^\varepsilon, 1 \leq j \leq m\}.$$

Notons \mathcal{F}_m la tribu engendrée par les variables $X_{i,m+1}, \dots, X_{in}$ et \mathbf{E}_n l'espérance conditionnelle $\mathbf{E}[\cdot | E_n, \mathcal{F}_m]$. Le même raisonnement que celui fait en début de preuve montre qu'il suffit d'étudier la distance $d(Y, W)$, où

$$Y = (X_{i1} - \mathbf{E}_n X_{i1}, \dots, X_{im} - \mathbf{E}_n X_{im}, 0, \dots, 0)$$

et où W est le sous-espace vectoriel engendré par H et les vecteurs

$$(0, \dots, 0, X_{i,m+1}, \dots, X_{in}) \text{ et } (\mathbf{E}_n X_{i1}, \dots, \mathbf{E}_n X_{im}, 0, \dots, 0).$$

Comme Y est un vecteur dont les composantes sont bornées sous \mathbf{P}_n et que l'application $d(\cdot, W)$ est 1-lipschitzienne, l'inégalité de Talagrand nous donne

$$\forall t > 0, \quad \mathbf{P}_n \left(|d(Y, W) - \mathbf{M}(d(Y, W))| \geq t \right) \leq 4 \exp \left(-\frac{t^2}{16n^{2\varepsilon}} \right), \quad (17)$$

où $M(d(Y, W))$ désigne la médiane de $d(Y, W)$. En intégrant (17) on obtient en particulier que $\mathbf{E}_n[|d(Y, W) - M(d(Y, W))|^2] \leq Cn^{2\varepsilon}$ pour une constante strictement positive $C > 0$. Par l'inégalité de Cauchy-Schwarz :

$$\left(M(d(Y, W)) - \sqrt{\mathbf{E}_n[d(Y, W)^2]}\right)^2 \leq \mathbf{E}_n\left[|M(d(Y, W)) - d(Y, W)|^2\right].$$

Donc

$$M(d(Y, W)) \geq \sqrt{\mathbf{E}_n[d(Y, W)^2]} - C^{1/2}n^\varepsilon.$$

Par ailleurs, en notant P la matrice de la projection orthogonale sur l'orthogonal de W , on a :

$$\mathbf{E}_n[d(Y, W)^2] = \mathbf{E}\left[\sum_{1 \leq i, j \leq n} \bar{Y}_i P_{ij} Y_j\right] = \sum_{k=1}^m \mathbf{E}_n[Y_k^2] P_{kk},$$

car les Y_i sont indépendants, de moyennes nulles, et nuls dès que $i \geq m + 1$. Posons $\sigma^2 := \mathbf{E}|Y_i|^2$. Alors

$$\begin{aligned} \mathbf{E}_n[d(Y, W)^2] &= \sigma^2 \left(\text{Tr}P - \sum_{k=m+1}^n P_{kk} \right) \\ &\geq \sigma^2 \left(\text{Tr}P - \sum_{k=m+1}^n |P_{kk}| \right) \\ &\geq \sigma^2 (n - \dim H - 2 - n^{1-\varepsilon}), \end{aligned}$$

où l'on a utilisé que la trace d'un projecteur est égale à son rang, et que ses coefficients sont de valeurs propres bornées par 1. En choisissant $\gamma < \varepsilon$, un développement limité (et le fait que $\sigma^2 = 1 - o(1)$) nous permet d'obtenir, pour tout n suffisamment grand :

$$M(d(Y, W)) \geq c\sqrt{n - \dim H},$$

pour une constante $c > 1/2$. Le choix $1/2 < c < 1$ et $t = (c - 1/2)\sqrt{n - \dim H}$ dans (17) fournit finalement le résultat. \square

3.3 Conclusion

Pour obtenir la majoration (16), il faut contrôler les plus grandes valeurs singulières, à savoir les $s_{n-k}(n^{-1/2}X - zI)$ pour $k \in \{n^{1-\gamma}, \dots, n-1\}$. On raisonnera plutôt sur $k \in \{2n^{1-\gamma}, \dots, n-1\}$ et cela ne nous fait pas perdre en généralité quitte à augmenter γ . Fixons un tel indice k . Pour alléger les notations on notera $s_i = s_i(n^{-1/2}X - zI)$. Notons par ailleurs s'_i les valeurs singulières de la matrice $n^{-1/2}X - zI$ privée de ses $k/2$ dernières lignes. Les inégalités d'entrelacement (A.1) assurent que $s'_{n-k} \leq s_{n-k}$. On en déduit que

$$\frac{k}{2}s_{n-k}^{-2} \leq \frac{k}{2}s'_{n-k}{}^{-2} \leq \sum_{j=n-k}^{n-k/2} s'_{n-j}{}^{-2}.$$

Par le lemme A.1, il vient

$$\frac{k}{2}s_{n-k}^{-2} \leq n \sum_{j=1}^{n-k/2} d_j^{-2}, \quad (18)$$

où d_j désigne la distance du vecteur L_j au sous-espace H_j engendré par les vecteurs L_α , α étant différent de j et variant entre 1 et $n - k/2$. Comme la dimension de H_j vérifie $\dim H_j \leq n - n^{1-\gamma}$, l'estimation exponentiellement petite du lemme 3.1 assure que la série

$$\sum_{n \geq 1} \mathbf{P} \left(\bigcup_{k=2n^{1-\gamma}}^{n-1} \bigcup_{j=n-k}^{n-k/2} \left\{ d_j \leq \frac{1}{2} \sqrt{\frac{j}{2}} \right\} \right)$$

converge. Par le lemme de Borel-Cantelli, on en déduit que p.s, pour tout n suffisamment grand, tout $2n^{1-\gamma} \leq k \leq n - 1$ et tout $n - k \leq j \leq n - k/2$, on a la minoration $d_j \geq \sqrt{j}/(2\sqrt{2})$. La relation (18) mène finalement à l'existence d'une constante $C > 0$ telle que presque sûrement, pour tout n suffisamment grand,

$$\forall k \in \{2n^{1-\gamma}, \dots, n - 1\}, \quad s_{n-k} \geq C \frac{k}{n}.$$

Ainsi, on a la majoration (16) pour $0 < p < 1$ puisque dans ce cas

$$\frac{1}{n} \sum_{i=1}^{n-n^{1-\gamma}} s_i^{-p} \leq \frac{C}{n} \sum_{i=1}^n \left(\frac{i}{n}\right)^{-p} \xrightarrow{n \rightarrow \infty} C \int_0^1 s^{-p} ds < \infty,$$

le deuxième terme étant une somme de Riemann associée à l'intégrale de droite.

4 Contrôle presque sûr de $s_n(X + M)$

4.1 Problème de Littlewood-Offord

Historiquement, le problème de Littlewood-Offord s'intéresse à la concentration de la somme $\varepsilon_1 v_1 + \dots + \varepsilon_n v_n$ où les v_i sont des entiers fixés, et les ε_i des variables de Bernoulli indépendantes de paramètre $1/2$. Plus précisément, le but est de majorer la quantité

$$\sup_{k \in \mathbb{Z}} \mathbf{P} \left(\sum_{i=1}^n \varepsilon_i v_i = k \right).$$

Quel est le lien avec notre problème? Etudier la plus petite valeur singulière d'une matrice revient à étudier, comme on le verra, la distance d'une colonne C au sous-espace engendré par les autres colonnes. En notant ζ un vecteur orthogonal à ce sous-espace, cela revient à étudier la norme du produit scalaire entre C et ζ . Dans notre cas, on tombe sur la quantité

$$\zeta_1 x_1 + \dots + \zeta_n x_n,$$

où les x_i sont des v.a. indépendantes de même loi que X_{11} . Comme ζ est indépendant des x_i on le remplace par un vecteur $v \in \mathbb{C}^n$ quelconque. L'idée est alors de distinguer selon que la somme se concentre ou pas. Pour quantifier cela on introduit les "small ball probabilities" :

$$\forall v \in \mathbb{C}, \forall r > 0, \quad p_{r,v} := \sup_{z \in \mathbb{C}} \mathbf{P} \left(\left| \sum_{i=1}^n x_i v_i - z \right| \leq r \right). \quad (19)$$

On dira qu'un vecteur v est pauvre lorsque $p_{n^{-B+1/2}, v} \leq n^{-A-1}$. La constante $A > 0$ est celle du lemme 1.3 et on donnera une condition sur $B > 0$ dans les paragraphes qui suivent. Les autres vecteurs sont appelés riches, la terminologie n'étant pas anodine. En effet, ces derniers correspondent au cas où la somme $x_1 v_1 + \dots + x_n v_n$ se concentre, ce qui implique que l'ensemble des coordonnées de v possède une structure additive riche. L'article [16] parle de problème inverse de Littlewood-Offord².

Supposons qu'il existe une constante $C_1 > 0$ telle que $s_1(M) \leq n^{C_1}$. Pour obtenir le lemme, il nous suffit de montrer que :

$$\mathbf{P}(\exists v \in \mathbb{C}^n \text{ riche} : \|(X + M)v\|_2 \leq n^{-B}) \leq n^{-A}$$

et

$$\mathbf{P}(\exists v \in \mathbb{C}^n \text{ pauvre} : \|(X + M)v\|_2 \leq n^{-B}) \leq n^{-A}.$$

En effet pour toute matrice M , les égalités de Courant-Fisher (26) entraînent $s_1(M) = \|M\|$ où $\|\cdot\|$ désigne la norme d'opérateur associée à la norme euclidienne :

$$\|M\| = \max_{\|x\|_2=1} \|Mx\|_2.$$

4.2 Vecteurs pauvres

Le cas de l'ensemble des vecteurs pauvres est le plus simple. Notons E l'évènement

$$\text{il existe un vecteur pauvre } v \text{ tel que } \|(X + M)v\|_2 \leq n^{-B}.$$

Si E est réalisé, alors par dualité il existe un vecteur unitaire w tel que

$$\|w^*(X + M)\|_2 \leq n^{-B}.$$

Notons F l'évènement E auquel on ajoute la condition que w_n est la coordonnée de module maximal : $|w_n| \geq |w_i|$ pour tout $1 \leq i \leq n$. En particulier $|w_n| \geq n^{-1/2}$. Par symétrie

$$\mathbf{P}(E) \leq n\mathbf{P}(E \cap F).$$

Notons L_1, \dots, L_n les lignes de $X + M$ et plaçons nous sur $F \cap E$. Pour tout vecteur unitaire y on a $|w^*(X + M)y| \leq n^{-B}$, donc par inégalité triangulaire et hypothèse sur w :

$$\begin{aligned} |L_n| &\leq \frac{1}{|w_n|} \left(\sum_{i=1}^{n-1} |w_i| |L_i \cdot y| + n^{-B} \right) \\ &\leq \sum_{i=1}^{n-1} |L_i \cdot y| + n^{-B+1/2}, \end{aligned}$$

Finalement l'inégalité de Cauchy-Schwarz permet d'obtenir

$$|L_n \cdot y| \leq n^{-1/2} \left(\sum_{i=1}^{n-1} |L_i \cdot y|^2 \right)^{1/2} + n^{-B+1/2}.$$

². sachant que pour un v donné, la somme se concentre, on essaye d'extraire de l'information sur la structure de v .

La dernière étape consiste à conditionner par rapport aux $n-1$ premières lignes. Sans perte de généralité on se restreint au cas où $\mathbf{P}(E | L_1, \dots, L_{n-1}) > 0$. Il existe alors un vecteur aléatoire u pauvre, indépendant de L_n et tel que

$$\left(\sum_{i=1}^{n-1} |L_i \cdot u|^2 \right)^{1/2} \leq n^{-B}.$$

On a donc le résultat puisque ce qui précède entraîne :

$$\mathbf{P}(E \cap F) \leq \mathbf{E}[\mathbf{P}(|L_n \cdot u| \leq 2n^{-B+1/2} | L_1, \dots, L_n)] \leq n^{-A-1}.$$

4.3 Vecteurs riches

Traisons maintenant le cas des vecteurs riches. On ne rentrera pas dans les détails, le but étant seulement de donner une idée de la preuve. D'abord, comme les coefficients de la matrice X possèdent un moment d'ordre 2, on en déduit que pour un $\gamma > 0$ suffisamment grand, l'évènement

$$s_1(X + M) \leq n^\gamma \quad \text{et} \quad \sum_{i=1}^n |x_i| \leq n^\gamma$$

a lieu avec une probabilité tendant vers 1 lorsque n tend vers l'infini. On peut donc raisonner conditionnellement à celui-ci. Pour des raisons techniques, il faut supposer que $B > 2\gamma A + 3\gamma + 1/2$. Notons J l'unique entier tel que $2A + 2 < J \leq 2A + 3$. On pose $\delta = (A + 1)/J$ et $C = \gamma + 2\varepsilon$ pour un $\varepsilon > 0$ petit, de sorte que $B > JC + 1/2$. Soit v un vecteur riche. La suite des small ball probabilités $\{p_{n^{-B+Cj+1/2}, v}\}_{0 \leq j \leq J}$ (c.f. (19)) est croissante, comprise entre n^{-A-1} et 1, donc il existe au moins un indice j tel que

$$p_{n^{-B+C(j+1)+1/2}, v} \leq n^\delta p_{n^{-B+Cj+1/2}, v}.$$

L'ensemble des vecteurs riches peut donc s'écrire comme la réunion des

$$\Omega_{j,k} := \{v \text{ riche} : p_{n^{-B+C(j+1)+1/2}, v} \leq n^\delta p_{n^{-B+Cj+1/2}, v} \\ \text{et } n^{-k\varepsilon} \leq p_{n^{-B+Cj+1/2}, v} < n^{-(k-1)\varepsilon}\},$$

pour $0 \leq j \leq J - 1$ et $1 \leq k \leq \lceil (A + 1)/\varepsilon \rceil$. Comme le nombre de telles paires d'indices ne dépend pas de n , il suffit de montrer que

$$\mathbf{P}(\exists v \in \Omega_{j,k}, \|(X + M)v\|_2 \leq n^{-B}) = O(n^{-A})$$

pour une paire (j, k) fixée.

C'est à ce stade que le problème de Littlewood-Offord inverse apparaît. Remarquons que $\Omega_{j,k}$ est inclus dans l'ensemble des vecteurs unitaires v tels que $p_{n^{-B+Cj+1/2}, v} \geq n^{-k\varepsilon}$, que l'on note $S_{n^{-k\varepsilon}}$.

Lemme 4.1. *Il existe un ensemble V de taille finie³*

$$|V| \ll n^{-n/2+\varepsilon n} n^{nk\varepsilon} + \exp(o(n))$$

de sorte que pour tout $v \in S_{n^{-k\varepsilon}}$ il existe $v' \in V$ tel que $\|v - v'\|_\infty \leq n^{-B+Cj+1/2}$.

3. La notation $f \ll g$ signifie qu'il existe une constante $K > 0$ telle que $|f| \leq K|g|$.

La preuve de ce résultat pourra être trouvée dans [16]. Fixons un vecteur $v \in \Omega_{j,k}$ et notons $v' \in V$ le vecteur associé. Comme on s'est ramené au cas où $\|X + M\| \leq n^\gamma$, et par définition de v' ,

$$\|(X + M)(v - v')\|_2 \leq n^{-B+Cj+1+\gamma}.$$

Par inégalité triangulaire on obtient $\|(X + M)v'\| \leq 2n^{-B+Cj+1+\gamma}$. En notant L_i les lignes de la matrice $X + M$, une conséquence est qu'il existe au moins $m = n - n^{1-\varepsilon}$ coordonnées i telles que $|L_i \cdot v'| \leq n^{-B+Cj+1/2+\gamma+\varepsilon}$. Ainsi :

$$\begin{aligned} \mathbf{P}(\exists v \in \Omega_{j,k}, \|(X + M)v\|_2 \leq n^{-B}) \\ \ll \left(n^{-n/2+\varepsilon n} n^{nk\varepsilon} + \exp(o(n)) \right) \binom{n}{m} \\ \times \max_{v' \in V} \mathbf{P} \left(\left| \sum_{i=1}^n x_i v'_i \right| \leq n^{-B+Cj+1/2+\gamma+\varepsilon} \right), \end{aligned}$$

où l'on rappelle que les x_i sont des copies i.i.d. de X_{11} . La fin de la preuve consiste à majorer la probabilité du membre de droite, et ne pose pas de difficultés particulières, on renvoie à [16] pour les détails.

5 Contrôle en probabilité de $s_n(X + M)$

5.1 Une découpe astucieuse de la sphère unité

La preuve du lemme 1.4 repose sur un découpage de la sphère unité en deux parties : les vecteurs compressibles dont un nombre non négligeable de coordonnées sont presque nulles et les autres vecteurs appelés incompressibles. Plus précisément, on quantifie le fait d'être compressible ou pas à l'aide de deux paramètres $\delta, \rho > 0$. Le paramètre δ contrôle le nombre de coordonnées presque nulles :

$$\text{Sparse} = \text{Sparse}(\delta) := \{x \in \mathbb{C}^n : |\text{Supp}(x)| \leq \delta n\},$$

où $\text{Supp}(x)$ est l'ensemble des coordonnées non nulles de x . On définit alors l'ensemble des vecteurs compressibles comme l'ensemble des vecteurs unitaires ρ -éloignés de Sparse :

$$\text{Comp}(\rho, \delta) = \text{Comp} := \{x \in \mathbb{S}^{n-1} : \text{dist}(x, \text{Sparse}) \leq \rho\}.$$

On pose enfin $\text{Incomp} := \mathbb{S}^{n-1} \setminus \text{Comp}$, l'ensemble des vecteurs incompressibles.

Dans tout ce qui suit, on notera $A = X + M$, $b := \mathbf{P}(|X_{11}| \leq a)$ et $\sigma^2 := \text{Var}(X_{11} \mathbf{1}_{|X_{11}| \leq a})$ pour une constante $a > 0$. Comme $|X_{11}|$ a un second moment, on peut supposer que a est tel que $b, \sigma^2 > 0$. On va estimer $\|Ax\|_2$ en distinguant selon que x est compressible ou pas.

5.2 Vecteurs compressibles

Soit $x \in \text{Comp}$. On note y le vecteur le plus proche de x dans l'ensemble Sparse. Par inégalité triangulaire, on a :

$$\|Ax\|_2 \geq \|Ay\|_2 - \rho s_1(A).$$

Si $\pi \subset \{1, \dots, n\}$ est l'ensemble des indices des coordonnées non nulles de y , alors en notant $A_{|\pi}$ la matrice obtenue en enlevant les colonnes d'indices $i \in \pi$ de A :

$$\|Ax\|_2 \geq (1 - \rho)s_n(A_{|\pi}) - \rho s_1(A).$$

Par ailleurs, on dispose de la minoration

$$s_n(A_{|\pi}) \geq \frac{1}{\sqrt{|\pi|}} \min_{i \in \pi} \text{dist}(C_i, H_i),$$

où C_i désigne la i -ème colonne de A et H_i le sous-espace vectoriel engendré par les colonnes de A différentes de C_i : $H_i = \langle C_j, j \neq i \rangle$. En effet soit $z = (z_1, \dots, z_n) \in \mathbb{C}^n$ et $i_0 \in \{1, \dots, n\}$ une coordonnée telle que $|z_{i_0}|$ est maximal, si $P_{i_0}^\perp$ désigne la projection orthogonale sur H_{i_0} , alors

$$\begin{aligned} \|A_{|\pi} z\|_2^2 &\geq \|z_{i_0}(C_{i_0} - P_{i_0}^\perp C_{i_0})\|_2^2 \\ &= |z_{i_0}|^2 \text{dist}(C_{i_0}, H_{i_0})^2 \\ &\geq \frac{1}{|\pi|} \sum_{i \in \pi} |z_i|^2 \text{dist}(C_{i_0}, H_{i_0})^2. \end{aligned}$$

On en déduit finalement que

$$\min_{x \text{ Comp}} \|Ax\|_2 \geq (1 - \rho) \frac{1}{\sqrt{|\pi|}} \min \left\{ \text{dist}(C_i, H_i); \right. \\ \left. \pi \subset \{1, \dots, n\}, |\pi| \leq \delta n, i \in \pi \right\} - \rho s_1(A).$$

Le résultat [4, lemme A.1] concernant la distance d'un vecteur aléatoire à un sous-espace vectoriel est analogue au lemme 3.1. Il assure que quitte à réduire $\delta > 0$, il existe une constante $K > 0$ telle que pour tout $n \gg 1$:

$$\forall i \in \pi, \forall \varepsilon > 0, \quad \mathbf{P}(\text{dist}(C_i, H_i) \leq \varepsilon \sigma \sqrt{n}) \leq K \delta n e^{-K \sigma^2 n}.$$

Un jeu avec les constantes δ et ρ mène finalement à la majoration :

$$\mathbf{P}\left(\min_{x \in \text{Comp}} \|Ax\|_2 \leq \frac{\varepsilon \sigma}{2\delta^{1/2}}; s_1(A) \leq s \right) \leq K \binom{n}{\delta n} \delta n e^{-K \sigma^2 n},$$

Le coefficient binomial provient du fait que l'on doit choisir un sous-ensemble π de taille plus petite que δn , et le facteur δn apparaît en utilisant la borne de l'union sur l'ensemble des indices $i \in \pi$. En utilisant la formule de Stirling, il est possible d'obtenir une borne exponentielle sur la probabilité précédente : il existe une constante $c > 0$ telle que

$$\mathbf{P}\left(\min_{x \in \text{Comp}} \|Ax\|_2 \leq \frac{\varepsilon \sigma}{2\delta^{1/2}}; s_1(A) \leq s \right) \leq e^{-cn}.$$

Ceci permet de conclure le cas des vecteurs compressibles.

Remarque 5.1. *Pour être précis, on peut par exemple prendre $\delta = c_1 \sigma^2 |\log \sigma|^{-1}$ pour une constante $c_1 > 0$ suffisamment petite et $\rho = (1 \wedge (\varepsilon \sigma)(s\delta)^{-1})/4$.*

5.3 Vecteurs incompressibles

Le cas des vecteurs incompressibles est plus compliqué et est à l'origine du terme non sommable $n^{-1/2}$ dans le lemme 1.4. Ce terme provient d'une application de l'estimation de Berry-Esseen, énoncée sous forme de lemme.

Lemme 5.1. *Il existe une constante $K > 0$ telle que si $\{Z_i\}_{1 \leq i \leq n}$ est une famille de v.a. indépendantes, centrées et possédant un troisième moment, alors pour tout $t \geq 0$:*

$$\sup_{z \in \mathbb{C}} \mathbf{P} \left(\left| \sum_{i=1}^n Z_i - z \right| \leq t \right) \leq \frac{Kt}{\sqrt{\sum_{i=1}^n \mathbf{E}[|Z_i|^2]}} + \frac{K \sum_{i=1}^n \mathbf{E}[|Z_i|^3]}{(\sum_{i=1}^n \mathbf{E}[|Z_i|^2])^{3/2}}.$$

Preuve. Notons $\tau^2 = \sum_{i=1}^n \mathbf{E}[|Z_i|^2]$ et supposons que $\sum_{i=1}^n \mathbf{E}[|\Re Z_i|^2] \geq \tau^2/2$ où $\Re Z_i$ désigne la partie réelle de Z_i (si ce n'est pas le cas le même raisonnement s'appliquera en remplaçant partie réelle par partie imaginaire). Comme

$$\sup_{z \in \mathbb{C}} \mathbf{P} \left(\left| \sum_{i=1}^n Z_i - z \right| \leq t \right) \leq \sup_{z \in \mathbb{C}} \mathbf{P} \left(\left| \sum_{i=1}^n \Re Z_i - z \right| \leq t \right),$$

on peut supposer que Z_i est à valeurs réelles, quitte à diviser c par 2. Notons \mathcal{N} une variable aléatoire gaussienne centrée de variance τ^2 . Alors le théorème de Berry-Esseen, dont un énoncé pourra être trouvé dans [8], fournit l'existence d'une constante $K > 0$ telle que

$$\sup_{t \in \mathbb{R}} \left| \mathbf{P} \left(\sum_{i=1}^n Z_i \leq t \right) - \mathbf{P}(\mathcal{N} \leq t) \right| \leq K\tau^{-3/2} \sum_{i=1}^n \mathbf{E}[|Z_i|^3].$$

On en déduit donc le résultat par inégalité triangulaire et parce que la densité de \mathcal{N} est bornée. En effet pour tout $x \in \mathbb{R}$ et tout $t > 0$:

$$\begin{aligned} \mathbf{P} \left(\left| \sum_{i=1}^n Z_i - x \right| \leq t \right) &\leq \mathbf{P}(|\mathcal{N} - x| \leq t) + 2K\tau^{-3/2} \sum_{i=1}^n \mathbf{E}[|Z_i|^3] \\ &\leq \frac{2t}{\sqrt{2\pi\tau^2}} + 2K\tau^{-3/2} \sum_{i=1}^n \mathbf{E}[|Z_i|^3]. \end{aligned}$$

D'où le résultat, quitte à augmenter la constante $K > 0$. \square

On va se ramener à ce lemme dans les paragraphes qui suivent. Un premier constat important est qu'un vecteur incompressible est "étalé" : un grand nombre de ses coordonnées sont de l'ordre $n^{-1/2}$. En effet soit $x \in \text{Incomp}$. Comme x est unitaire, il possède au plus $\delta n/2$ coordonnées de module supérieur à $\sqrt{2/(\delta n)}$. De plus comme x est au moins ρ -éloigné de Sparse, il possède au moins δn coordonnées de module supérieur à ρ/\sqrt{n} . En conséquence, il existe un sous-ensemble π de $\{1, \dots, n\}$ tel que $|\pi| \geq \delta n/2$ et

$$\forall i \in \pi, \quad \frac{\rho}{\sqrt{n}} \leq |x_i| \leq \sqrt{\frac{2}{\delta n}}.$$

Comme $\|Ax\|_2 \geq \max_{1 \leq i \leq n} |x_i| \text{dist}(C_i, H_i)$ on en déduit la minoration

$$\|Ax\|_2 \geq \frac{\rho}{\sqrt{n}} \max_{i \in \pi} \text{dist}(C_i, H_i).$$

On se ramène alors à un problème de distance d'un vecteur aléatoire à un sous-espace puisque pour tout $t \geq 0$:

$$\begin{aligned} \mathbf{P}\left(\min_{x \in \text{Incomp}} \|Ax\|_2 \leq \frac{t\rho}{\sqrt{n}}\right) &\leq \max_{|\pi| \geq \delta n/2} \mathbf{P}\left(\max_{i \in \pi} \text{dist}(C_i, H_i) \leq t\right) \\ &\leq \max_{|\pi| \geq \delta n/2} \mathbf{E}\left[\frac{1}{|\pi|} \sum_{i \in \pi} \mathbf{1}_{\{\text{dist}(C_i, H_i) \leq t\}}\right] \\ &\leq \frac{2}{\delta n} \sum_{1 \leq i \leq n} \mathbf{P}(\text{dist}(C_i, H_i) \leq t). \end{aligned}$$

Fixons un indice $i \in \{1, \dots, n\}$. Soit $\zeta \in \mathbb{S}^{n-1} \cap H_i^\perp$ un vecteur unitaire orthogonal à H_i . Comme $|\langle \zeta, C_i \rangle| \leq \text{dist}(C_i, H_i)$, le problème se réduit à contrôler la norme de $\langle \zeta, C_i \rangle$: il suffit de montrer que pour tout $t \geq 0$,

$$\mathbf{P}(|\langle \zeta, C_i \rangle| \leq t) \leq \frac{c}{\sigma} \sqrt{\frac{|\log \rho|}{\delta}} \left(\frac{t}{\rho} + \frac{1}{\sqrt{n}}\right). \quad (20)$$

En effet si (20) est vrai, alors les inégalités ci-dessus entraîne que

$$\mathbf{P}\left(\min_{x \in \text{Incomp}} \|Ax\|_2 \leq \frac{t\rho}{\sqrt{n}} \text{ et } s_1(A) \leq s\right) \leq c\sqrt{|\log cs|} \left(ts^2 + \frac{1}{\sqrt{n}}\right).$$

Le lemme 1.4 en découle en prenant $t = s^{-2}n^{-1/2}$ et $s = n^r$ pour un $r > 0$.

La fin de la preuve consiste à conditionner le problème à un sous-ensemble convenable, de sorte que l'on pourra appliquer le lemme 5.1. D'abord, on remarque que le vecteur orthogonal ζ est incompressible avec une probabilité tendant vers 1 lorsque n tend vers l'infini. En effet, si \tilde{A} désigne la matrice A^* privée de sa i -ème ligne, alors $\tilde{A}\zeta = 0$. Comme

$$\min_{x \in \text{Comp}} \|\tilde{A}x\|_2 \geq \min_{x \in \text{Comp}} \|A^*x\|_2,$$

on en déduit que

$$\mathbf{P}(\zeta \in \text{Comp} \text{ et } s_1(A^*) \leq s) \leq \mathbf{P}\left(\min_{x \in \text{Comp}} \|A^*x\|_2 = 0 \text{ et } s_1(A^*) \leq s\right).$$

Cette dernière probabilité est exponentiellement petite par ce qui a été fait dans le cas des vecteurs compressibles, dans la partie 5.2. On peut donc raisonner conditionnellement à $\{\zeta \in \text{Comp}\}$. Notons π l'ensemble des coordonnées de ζ de l'ordre de $n^{-1/2}$. Par symétrie on peut supposer que cet ensemble constitue les premières coordonnées de ζ . De plus, comme $|\pi| \geq \delta n/2$, une application de l'inégalité de Hoeffding fournit

$$\mathbf{P}\left(\sum_{i \in \pi} \mathbf{1}_{|X_{ik}| \leq a} \leq \frac{\delta nb}{4} \mid \zeta \in \text{Incomp}\right) \leq e^{-|\pi|^2 b^2 / 2},$$

Autrement dit, on peut supposer que $m := \lfloor \delta nb/4 \rfloor$ coordonnées $i \in \pi$ sont telles que $|X_{ik}| \leq a$. Par symétrie on peut supposer que ce sont les m premières. Introduisons l'évènement

$$E := \{\zeta \in \text{Incomp}\} \cap \{|X_{ik}| \leq a, 1 \leq i \leq m\},$$

de sorte qu'il suffit de raisonner sur la probabilité conditionnée $\mathbf{P}_E(\cdot) = \mathbf{P}(\cdot|E)$. Remarquons que l'on peut écrire le produit scalaire entre ζ et C comme somme de deux variables aléatoires indépendantes :

$$\langle \zeta, C \rangle = \sum_{i=1}^m \bar{\zeta}_i (X_{ik} - \mathbf{E}[X_{ik}|E]) + U,$$

En notant \mathcal{L}_U la loi de U :

$$\begin{aligned} \mathbf{P}_E(|\langle \zeta, C \rangle| \leq \rho t) &= \int \mathbf{P}_E \left(\left| \sum_{i=1}^m \bar{\zeta}_i (X_{ik} - \mathbf{E}[X_{ik}|E]) + z \right| \leq \rho t \right) \mathcal{L}_U(dz) \\ &\leq \sup_{z \in \mathbb{C}} \mathbf{P}_E \left(\left| \sum_{i=1}^m \bar{\zeta}_i (X_{ik} - \mathbf{E}[X_{ik}|E]) - z \right| \leq \rho t \right). \end{aligned}$$

Sous \mathbf{P}_E , les variables aléatoires $\{Z_i\}_{1 \leq i \leq m} := \{\bar{\zeta}_i (X_{ik} - \mathbf{E}[X_{ik}|E])\}_{1 \leq i \leq m}$ sont i.i.d. et bornées. Par le lemme 5.1

$$\mathbf{P}_E(|\langle \zeta, C \rangle| \leq \rho t) \leq \frac{K \rho t}{\sqrt{\sum_{i=1}^m \mathbf{E}[|Z_i|^2]}} + \frac{K \sum_{i=1}^m \mathbf{E}[|Z_i|^3]}{(\sum_{i=1}^m \mathbf{E}[|Z_i|^2])^{3/2}}.$$

Notons que le même raisonnement s'applique en remplaçant l'ensemble des indices $\{1, \dots, m\}$ par n'importe quel sous-ensemble $\alpha \subset \{1, \dots, m\}$. Ce constat va nous permettre de conclure. Posons $L = 2^{-1} \log_2(2/(\delta \rho))$. Par le principe des tiroirs, il existe un $1 \leq j \leq L$ tel que le sous-ensemble d'indices :

$$\alpha_j := \left\{ 1 \leq i \leq m : \frac{2^{j-1} \rho}{\sqrt{n}} \leq |\zeta_i| \leq \frac{2^j \rho}{\sqrt{n}} \right\}$$

est de cardinal au moins m/L . Finalement, le choix des constantes δ et ρ fait dans la remarque 5.1 fournit l'existence d'un $c' > 0$ tel que $L \leq c' |\log \rho|$, menant à l'inégalité (20) recherchée puisque :

$$\begin{aligned} \mathbf{P}_E(|\langle \zeta, C \rangle| \leq \rho t) &\leq \frac{K t}{\sqrt{\sum_{i=1}^m \mathbf{E}[|Z_i|^2]}} + \frac{K \sum_{i=1}^m \mathbf{E}[|Z_i|^3]}{(\sum_{i=1}^m \mathbf{E}[|Z_i|^2])^{3/2}} \\ &\leq \frac{K t \sqrt{n}}{\sqrt{2^{2j-2} \rho^2 \sigma^2 |\alpha_j|}} + \frac{2^{j+1} K \rho a}{\sqrt{n} \sqrt{\sum_{i=1}^m \mathbf{E}[|Z_i|^2]}} \\ &\leq \frac{2 K t}{\rho \sqrt{2^{2j-2} \rho^2 \sigma^2}} \sqrt{\frac{c' |\log \rho|}{\delta}} + \frac{2^{j+1} K \rho a}{\sqrt{2^{2j-2} \rho^2 \sigma^2 |\alpha_j|}} \\ &\leq \frac{c}{\sigma} \sqrt{\frac{|\log \rho|}{\delta}} \left(\frac{t}{\rho} + \frac{1}{\sqrt{n}} \right), \end{aligned}$$

pour une certaine constante $c > 0$.

6 Convergence des mesures singulières

On montre ici la convergence des mesures singulières $\nu_{n^{-1/2}X-z}$ pour tout complexe z , où X est défini au début de la sous-partie 1.1. Le point important est que les lois limites ν_z ne dépendent que de z : on parle d'universalité. Dans tout ce qui suit, $z \in \mathbb{C}$ est un complexe fixé. Remarquons que la suite des $\nu_{n^{-1/2}X-z}$ est tendue puisque la suite de leurs seconds moments est bornée par l'inégalité (14).

6.1 Réduction du problème

Dans cette sous-partie, on effectue trois réductions successives, régulièrement faites lors de l'étude asymptotique des spectres de matrices aléatoires.

D'abord, on peut supposer que les coefficients de X sont bornés. Pour le voir, introduisons la matrice tronquée $Y = (Y_{ij}) = (X_{ij} \mathbf{1}_{\{|X_{ij}| \leq C\}})$, où C est une constante positive. Par l'inégalité de Hoffman-Wielandt (théorème A.5) :

$$\frac{1}{n} \sum_{k=1}^n |s_k(n^{-1/2}Y - z) - s_k(n^{-1/2}X - z)|^2 \leq \frac{1}{n^2} \sum_{1 \leq i, j \leq n} |X_{ij}|^2 \mathbf{1}_{\{|X_{ij}| > C\}}.$$

Comme X_{11} est de variance finie, $\mathbf{E}[|X_{ij}|^2 \mathbf{1}_{\{|X_{ij}| > C\}}]$ converge vers 0 lorsque C tend vers l'infini. Par la loi des grands nombres, le terme de droite converge donc vers 0 lorsque C tend vers l'infini. Mais le terme de gauche n'est autre que le carré la distance de Wasserstein W_2 entre $\nu_{n^{-1/2}Y-z}$ et $\nu_{n^{-1/2}X-z}$ (définie comme l'infimum des quantité $\mathbf{E}[(U - V)^2]^{-1/2}$, où (U, V) est un couplage des deux mesures singulières), et la convergence en norme de Wasserstein implique la convergence faible de mesures.

Ensuite, on peut supposer que les coefficients de X sont centrés. En effet, si $Z = X - \mathbf{E}X$, la conséquence (28) des inégalités de Lidskii assure que

$$\sup_{t \geq 0} |\nu_{n^{-1/2}X-z}([0, t]) - \nu_{n^{-1/2}Z-z}([0, t])| \leq \frac{\text{rg}(X - Z)}{n} \leq \frac{1}{n}.$$

Enfin, on montre qu'il suffit de raisonner sur la mesure singulière moyennisée $\mathbf{E}\nu_{n^{-1/2}X-z}$. Ceci est une conséquence du résultat de concentration suivant, qui repose sur l'indépendance des coefficients de X . Rappelons que la variation totale d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est la norme définie par

$$\|f\|_{TV} := \sup \sum_{k \in \mathbb{Z}} |f(x_{k+1}) - f(x_k)|,$$

où le supremum est pris sur l'ensemble des suites croissantes $\{x_k\}_{k \in \mathbb{Z}}$.

Lemme 6.1. *Soit M une matrice aléatoire complexe de taille n dont les colonnes sont indépendantes. Pour toute fonction $g : \mathbb{R} \rightarrow \mathbb{R}$ telle que $\|g\|_{TV} \leq 1$ et $g(x) \rightarrow 0$ lorsque x tend vers $\pm\infty$, on a la concentration exponentielle :*

$$\forall t > 0, \quad \mathbf{P} \left(\left| \int g d\nu_M - \mathbf{E} \int g d\nu_M \right| \geq t \right) \leq 2e^{-2nt^2}.$$

Preuve. Le lemme est une conséquence de l'inégalité de McDiarmid (B.2) appliquée à la fonction

$$G : \mathbb{C}^n \times \dots \times \mathbb{C}^n \longrightarrow \mathbb{R} \\ (C_1, \dots, C_n) \longmapsto \int g d\nu_{[C_1, \dots, C_n]},$$

où $[C_1, \dots, C_n]$ est la matrice formée des colonnes C_i . Il suffit de vérifier l'hypothèse des différences bornées. Soient $C_1, \dots, C_i, C'_i, \dots, C_n \in \mathbb{C}^n$. Si g est lisse, une intégration par partie fournit

$$\Delta_i := |G(C_1, \dots, C_i, \dots, C_n) - G(C_1, \dots, C'_i, \dots, C_n)| \\ = \left| \int g'(t) \left(F_{\nu_{[C_1, \dots, C_i, \dots, C_n]}}(t) - F_{\nu_{[C_1, \dots, C'_i, \dots, C_n]}}(t) \right) dt \right|,$$

où F_ν désigne la fonction de répartition de la mesure ν . Comme la variation totale de g n'excède pas 1, on en déduit que

$$\Delta_i \leq \|F_{\nu_{[C_1, \dots, C_i, \dots, C_n]}} - F_{\nu_{[C_1, \dots, C'_i, \dots, C_n]}}\|_\infty.$$

Cette inégalité est encore vérifiée pour une fonction g générale vérifiant les hypothèses du lemme par un argument d'approximation. La conséquence (28) des inégalités de Lidskii entraîne $\Delta_i \leq 1/n$, ce qui permet de conclure. \square

6.2 Matrice hermitisée

Plutôt que d'étudier la convergence des mesures singulières $\nu_{n^{-1/2}X-z}$, on s'intéresse à leurs symétrisées

$$\tilde{\nu}_{n^{-1/2}X-z}(\cdot) := \frac{\nu_{n^{-1/2}X-z}(\cdot) + \nu_{n^{-1/2}X-z}(\cdot)}{2}.$$

La raison de ce choix est que l'étude des mesures symétrisées nous ramène au cadre des matrices hermitiennes.

Lemme 6.2. *Soit A une matrice carrée complexe de taille n . Les valeurs propres de la matrice hermitisée*

$$H(A) := \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$$

sont les $\{\pm s_i(A)\}_{1 \leq i \leq n}$. En particulier $\mu_{H(A)} = \tilde{\nu}_A$.

Preuve. Pour tout $\lambda \in \mathbb{C}$, on a

$$\det \begin{pmatrix} -\lambda & A \\ A^* & -\lambda \end{pmatrix} = \det(\lambda^2 - AA^*),$$

d'où le résultat par définition des valeurs singulières de A . \square

Notons que $\nu_{n^{-1/2}X-z}$ converge si et seulement si $\tilde{\nu}_{n^{-1/2}X-z}$ puisque la symétrisation $\nu \mapsto \tilde{\nu}$ établie une bijection entre les mesures à support dans les réels positifs et les mesures symétriques réelles. Par la réduction faite avant, il suffit de raisonner sur les mesures moyennes. Pour montrer la convergence des $\mathbf{E}\tilde{\nu}_{n^{-1/2}X-z}$, on va considérer leurs transformées de Stieltjes.

Définition 6.1. Soit μ une mesure réelle. La transformée de Stieltjes m_μ de μ est la fonction définie par

$$m_\mu : \eta \in \mathbb{C} \setminus \mathbb{R} \mapsto \int \frac{1}{\lambda - \eta} d\mu(\lambda).$$

La tension de la suite $\tilde{\nu}_{n-1/2X-z}$ et le [1, théorème 2.4.4] assurent que si $m_{\mathbf{E}\tilde{\nu}_{n-1/2X-z}}(\eta)$ converge vers $m(\eta)$ pour tout $\eta \in \mathbb{C} \setminus \mathbb{R}$, alors m est la transformée de Stieltjes d'une mesure de probabilité ν_∞ et $\mathbf{E}\tilde{\nu}_{n-1/2X-z}$ converge étroitement vers ν_∞ . Pour terminer, notons que la transformée de Stieltjes d'une mesure symétrisée $\tilde{\nu}_A$ est reliée à ν par la formule suivante

$$m_{\tilde{\nu}}(\eta) = \frac{1}{2} \int_{\mathbb{R}_+} \left(\frac{1}{\lambda - \eta} - \frac{1}{\lambda + \eta} \right) d\nu(\lambda) = \int_{\mathbb{R}} \frac{\eta}{\lambda^2 - \eta^2} d\nu(\lambda). \quad (21)$$

6.3 Résolvante quaternionique

Pour montrer la convergence de $m_{\mathbf{E}\tilde{\nu}_{n-1/2X-z}}(\eta)$, les auteurs de [4] utilisent une méthode de résolvante. Cette méthode est classique dans l'analyse des mesures spectrales limites de matrices aléatoires hermitiennes, ce qui explique l'importance de la matrice hermitisée introduite dans la sous-partie précédente. Le point nouveau est que la méthode de résolvante est ici appliquée à une matrice dont les coefficients sont des matrices 2×2 . La présence de deux paramètres complexes a incité les auteurs à baptiser cet objet résolvante quaternionique. Le terme peut être trompeur : il n'y aura pas de quaternions dans ce qui suit.

Soit A comme dans le lemme 6.2. La résolvante quaternionique de A est définie par

$$\forall q \in \mathbb{H}_+, \quad R_A(q) = \begin{pmatrix} -\eta & A - z \\ A^* - \bar{z} & -\eta \end{pmatrix},$$

où \mathbb{H}_+ est l'ensemble des matrices 2×2

$$\left\{ \begin{pmatrix} \eta & z \\ \bar{z} & \eta \end{pmatrix}; z, \eta \in \mathbb{C}, \operatorname{Im}(\eta) > 0 \right\}.$$

On considèrera $R_A(q)$ comme un élément de $\mathcal{M}_n(\mathcal{M}_2(\mathbb{C}))$. Posons alors

$$\Gamma_A(q) := \frac{1}{n} \operatorname{Tr} R_A(q) = \frac{1}{n} \sum_{k=1}^n R_A(q)_{kk} = \begin{pmatrix} a(q) & b(q) \\ c(q) & d(q) \end{pmatrix}.$$

Nous allons voir que Γ_A permet de retrouver les mesures singulières ν_{A-z} (au travers des $\tilde{\nu}_{A-z}$). C'est une conséquence de la formule $m_{\tilde{\nu}_{A-z}} = (a(q) + d(q))/2$. Mais on peut être plus précis en exploitant les symétries de la résolvante quaternionique. Notons momentanément $A_z = A - z$. Comme

$$\begin{pmatrix} -\eta & A_z \\ A_z^* & -\eta \end{pmatrix}^{-1} = - \begin{pmatrix} \eta(\eta^2 - A_z A_z^*)^{-1} & A_z(\eta^2 - A_z^* A_z)^{-1} \\ A_z^*(\eta^2 - A_z A_z^*)^{-1} & \eta(\eta^2 - A_z^* A_z)^{-1} \end{pmatrix},$$

on en déduit que si τ désigne l'opérateur $\frac{1}{n} \operatorname{Tr}$:

$$\Gamma_A(q) = - \begin{pmatrix} \eta\tau(\eta^2 - A_z A_z^*)^{-1} & \tau A_z(\eta^2 - A_z^* A_z)^{-1} \\ \tau A_z^*(\eta^2 - A_z A_z^*)^{-1} & \eta\tau(\eta^2 - A_z^* A_z)^{-1} \end{pmatrix}.$$

La formule (21) entraîne

$$m_{\bar{\nu}_{A_z}}(\eta) = \int_{\mathbb{R}} \frac{\eta}{\lambda - \eta^2} d\mu_{A_z^* A_z}(\lambda) = \eta \tau(A_z^* A_z - \eta^2)^{-1}.$$

Comme $A_z^* A_z$ et $A_z A_z^*$ ont le même spectre, on en déduit que $a(q) = d(q)$. En fait, il existe un résultat plus général :

Lemme 6.3. *On a que $\Gamma_A(q) \in \mathbb{H}_+$. De plus $m_{\bar{\nu}_{A-z}} = a(q)$, et au sens des distributions $\pi \mu_A(\cdot) = -\lim_{t \searrow 0} \partial b(q(\cdot, it))$.*

Preuve. Les deux premiers points sont une conséquence direct de ce qui précède, l'égalité $\overline{b(q)} = c(q)$ étant due à la définition de τ . Pour la dernière affirmation, on renvoie à [4, lemme 4.19] pour les détails, le point important étant que si $t > 0$:

$$\bar{\partial} \int \log |s + it| d\nu_{A-z}(s) = -\frac{1}{2} b(q(z, it)),$$

où $\bar{\partial}$ est l'opérateur $(\partial_x + i\partial_y)/2$. En notant ∂ l'opérateur $(\partial_x - i\partial_y)/2$, et en utilisant la relation $4\partial\bar{\partial} = \Delta$, le résultat en découle aisément car l'intégrale de gauche converge vers le potentiel logarithmique de μ_A évalué en z lorsque t tend vers 0, par le théorème de convergence monotone. \square

Dans le cas où $A = n^{-1/2}X$, une méthode classique de résolvante, basée sur l'inversion par blocs de Schur et utilisant les réductions effectuées dans la partie 6.1, permet d'obtenir

$$\mathbf{E} \begin{pmatrix} a(q) & b(q) \\ \overline{b(q)} & a(q) \end{pmatrix} = - \left(q + \mathbf{E} \begin{pmatrix} a(q) & 0 \\ 0 & a(q) \end{pmatrix} \right)^{-1} + \varepsilon,$$

où ε est une matrice qui tend vers la matrice nulle lorsque n tend vers l'infini. Les détails du calcul se trouvent dans [4]. Ainsi, tout point d'accumulation $m(\eta)$ de $m_{\bar{\nu}_{n^{-1/2}X-z}}(\eta) = a(q)$ vérifie l'équation

$$m(\eta) = \frac{m(\eta) + \eta}{|z|^2 - (m(\eta) + \eta)^2}. \quad (22)$$

En particulier tout point d'accumulation ne dépend que de z . Pour conclure, il faut montrer qu'il existe une unique transformée de Stieltjes solution de (22). On utilise pour cela un fait général : la transformée de Stieltjes d'une mesure, vue comme application du demi-plan complexe supérieur $\mathbb{C}_+ = \{z \in \mathbb{C} : \text{Im}(z) > 0\}$, est à valeurs dans \mathbb{C}_+ . Par le théorème de Vitali, on en déduit qu'un point d'accumulation m est une fonction holomorphe de \mathbb{C}_+ dans \mathbb{C}_+ . Il s'agit donc de montrer qu'une unique solution de (22) est à valeurs dans \mathbb{C}_+ . Il suffit pour cela de prouver que deux solutions coïncident sur $i\mathbb{R}_+$, puisqu'elles seront alors égales par analyticité. La définition de la transformée de Stieltjes entraîne que pour tout $t > 0$, $m(it) \in i\mathbb{R}_+$. Notons alors $h(t)$ le réel strictement positif tel que $m(it) = ih(t)$. Ce dernier vérifie

$$1 = \frac{1 + th^{-1}}{|z|^2 - (h + t)^2},$$

donc est unique puisque le terme de droite est une fonction décroissante de h , dont la limite est $+\infty$ (resp. 0) lorsque h tend vers 0 (resp. $+\infty$).

7 Preuve alternative : méthode des moments

Nous donnons ici une preuve alternative de l'existence et universalité des mesures singulières limites ν_z du lemme 1.2. Tout comme avec la résolvante quaternionique, l'idée est de raisonner sur la matrice hermitisée, vue avec des coefficients à valeurs dans $\mathcal{M}_2(\mathbb{C})$. Dans le cas $z = 0$, notre méthode fournit une manière de retrouver la loi de Marcenko-Pastur pour les matrices carrées en utilisant uniquement la connaissance de la loi semi-circulaire. On commence par introduire quelques notations et résultats qui nous serviront dans la preuve.

7.1 Quelques notations

Nous utiliserons ici les réductions effectuées dans la sous-partie 6.1 : on peut supposer que les coefficients de X sont centrés et bornés. Quitte à renormaliser la matrice X , on suppose que $\mathbf{E}|X_{11}|^2 = 1$. La matrice hermitisée⁴ de X est

$$H(X) := \frac{1}{\sqrt{n}} \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix}.$$

Notons P la matrice de la permutation σ de $\{1, \dots, 2n\}$ définie par

$$\forall k \in \{1, \dots, n\}, \quad \sigma(2k-1) = k \text{ et } \sigma(2k) = n+k.$$

La matrice $PH(X)P = (H_{ij})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathcal{M}_2(\mathbb{C}))$ a pour coefficients

$$H_{ij} = \frac{1}{\sqrt{n}} \begin{pmatrix} 0 & X_{ij} \\ X_{ji} & 0 \end{pmatrix},$$

et sa mesure spectrale (cette fois vue comme matrice $2n \times 2n$) est encore $\tilde{\nu}_{n^{-1/2}X}$. Plus généralement, pour tout complexe $z \in \mathbb{C}$, la matrice $L := PH(X - \sqrt{n}z)P$, dont les coefficients sont

$$L_{ij} = H_{ij} - \mathbf{1}_{i=j} \begin{pmatrix} 0 & z \\ \bar{z} & 0 \end{pmatrix},$$

a pour mesure spectrale $\tilde{\nu}_{n^{-1/2}X-z}$. Fixons un entier $k \geq 1$ et notons $M_n(k)$ le moment d'ordre k de $\mathbf{E}\tilde{\nu}_{n^{-1/2}X-z}$.

Remarque 7.1. *Comme $\mathbf{E}\tilde{\nu}_{n^{-1/2}X-z}$ est symétrique, ses moments d'ordre impair sont nuls. Pour cette raison, on supposera désormais que k est pair.*

Un peu d'algèbre linéaire mène à la formule suivante

$$M_n(k) = \frac{1}{2n} \sum_{1 \leq i_1, \dots, i_k \leq n} \mathbf{E} \left[\text{Tr} \left(L_{i_1 i_2} L_{i_2 i_3} \cdots L_{i_k i_1} \right) \right].$$

S'inspirant de la preuve de théorème de Wigner dans [1], on associe au mot fermé $i_1 i_2 \dots i_k i_1$ le graphe $G(i_1 \dots i_k i_1) = (V, E)$ dont l'ensemble des sommets est l'ensemble des indices i_j et dont les arêtes sont les paires consécutives d'indices. Autrement dit :

$$V = \{i_j, 1 \leq j \leq k\} \text{ et } E = \{\{i_j, i_{j+1}\}, 1 \leq j \leq k\}.$$

4. la définition diffère d'un facteur $n^{-1/2}$ de celle donnée dans le lemme 6.2.

Une arête de la forme $\{i_j, i_{j+1} = i_j\}$ sera appelée boucle. Deux mots sont dits équivalents si il existe une bijection entre l'ensemble des lettres des deux mots. Pour tout $B \in \{1, \dots, k\}$ on notera \mathcal{W}^B un ensemble de représentants des mots fermés $i_1 \dots i_k i_1$, dont le graphe associé possède B boucles. Par soucis de clarté on notera \mathbf{i} le mot fermé générique $i_1 \dots i_k i_1$. Le moment d'ordre k se réécrit

$$M_n(k) = \sum_{B=1}^k \frac{1}{2n} \sum_{\mathbf{i} \in \mathcal{W}^B} C(\mathbf{i}) \mathbf{E} \left[\text{Tr} \left(L_{i_1 i_2} L_{i_2 i_3} \cdots L_{i_k i_1} \right) \right],$$

où $C(\mathbf{i}) = n(n-1) \cdots (n - |\mathbf{V}| + 1)$ est le nombre de mots équivalents à \mathbf{i} . On notera

$$M_n(k, B) := \frac{1}{2n} \sum_{\mathbf{i} \in \mathcal{W}^B} C(\mathbf{i}) \mathbf{E} \left[\text{Tr} \left(L_{i_1 i_2} L_{i_2 i_3} \cdots L_{i_k i_1} \right) \right]. \quad (23)$$

Le point important est que de nombreux termes s'annulent dans la somme ci-dessus, car les coefficients de X sont supposés centrés. D'autres encore auront une contribution asymptotique nulle. Commençons par traiter le cas $B = 0$, qui revient, quand $n \rightarrow +\infty$, à choisir $z = 0$.

7.2 Le cas $z = 0$: le quart de cercle

Cela revient à étudier la quantité

$$M_n(k, 0) = \frac{1}{2n} \sum_{\mathbf{i} \in \mathcal{W}^0} C(\mathbf{i}) \mathbf{E} \left[\text{Tr} \left(H_{i_1 i_2} H_{i_2 i_3} \cdots H_{i_k i_1} \right) \right].$$

En effet, on justifiera (remarque 7.2) que les contributions associées aux mots \mathbf{i} possédant des boucles sont asymptotiquement nulles dans ce cas. Commençons par remarquer que comme $\mathbf{E}X_{11} = 0$, si une arête $\{i, j\}$ apparaît une unique fois dans le mot \mathbf{i} , alors la contribution est nulle, par indépendance des coefficients de X . En particulier toute arête est parcourue au moins deux fois. Ainsi $|\mathbf{E}| \leq k/2$. Comme \mathbf{i} est fermé, le graphe G est connexe ce qui implique que $|\mathbf{V}| \leq |\mathbf{E}| + 1$. On en déduit l'inégalité $|\mathbf{V}| \leq k/2 + 1$. Ceci permet de réécrire

$$M_n(k, 0) = \frac{1}{2n} \sum_{s=1}^{\lfloor k/2 \rfloor + 1} C(s) \sum_{a=s-1}^{\lfloor k/2 \rfloor} \sum_{\mathbf{i} \in \mathcal{W}_{a,s}^0} \mathbf{E} \left[\text{Tr} \left(H_{i_1 i_2} H_{i_2 i_3} \cdots H_{i_k i_1} \right) \right],$$

où $\mathcal{W}_{a,s}^0$ est un ensemble de représentants des mots \mathbf{i} dont le graphe possède a arêtes, s sommets et 0 boucles. Un élément de $\mathcal{W}_{a,s}^0$ possède $C(s) \sim n^s$ mots équivalents. Un terme associé à un élément de $\mathcal{W}_{a,s}^0$ contribuera donc à la limite seulement si sa norme d'opérateur est d'ordre plus grand que n^{1-s} . Le lemme suivant est une conséquence immédiate du fait que les coefficients de X sont supposés bornés, disons par une constante $C > 0$.

Lemme 7.1. *Notons $\|\cdot\|$ la norme d'opérateur associée à la norme euclidienne. Pour tout $i, j \in \{1, \dots, n\}$, on a*

$$\forall k \geq 1, \quad \|\mathbf{E}|H_{ij}|^k\| \leq \frac{C^k}{n^{k/2}}.$$

Fixons $\mathbf{i} \in \mathcal{W}_{a,s}^0$. Alors

$$\left\| \frac{C(s)}{2n} H_{i_1 i_2} H_{i_2 i_3} \cdots H_{i_k i_1} \right\| = O(n^{s-1-k/2}),$$

et on en déduit que les seules contributions asymptotiques non nulles correspondent au cas où $s = k/2 + 1$ et donc $a = k/2$. Dans ce cas, le graphe associé à \mathbf{i} est un arbre (dont le contour est \mathbf{i}). En particulier chaque arête est parcourue exactement deux fois.

Remarque 7.2. *En fait, en considérant un mot quelconque \mathbf{i} possédant a arêtes et s sommets, on obtient la même estimation, ce qui implique que pour donner lieu à une contribution asymptotique non nulle, \mathbf{i} ne doit pas posséder de boucles.*

Reste à calculer la contribution limite.

Lemme 7.2. *Soit $\mathbf{i} = i_1 \dots i_k i_1$ un représentant donnant lieu à une contribution asymptotique non nulle. Alors*

$$\mathbf{E} \left[H_{i_1 i_2} H_{i_2 i_3} \cdots H_{i_k i_1} \right] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Preuve. En les regroupant deux par deux, on voit que le produit des matrices $H_{i_j i_{j+1}}$ vaut

$$\begin{pmatrix} X_{i_1 i_2} \overline{X_{i_3 i_2}} & 0 \\ 0 & \overline{X_{i_2 i_1}} X_{i_2 i_3} \end{pmatrix} \begin{pmatrix} X_{i_3 i_4} \overline{X_{i_5 i_4}} & 0 \\ 0 & \overline{X_{i_4 i_3}} X_{i_4 i_5} \end{pmatrix} \cdots \\ \cdots \begin{pmatrix} X_{i_{k-1} i_k} \overline{X_{i_1 i_k}} & 0 \\ 0 & \overline{X_{i_k i_{k-1}}} X_{i_k i_1} \end{pmatrix},$$

car k est pair. Donc, avec la convention $k+1 = 1$, celui-ci s'écrit sous la forme

$$\begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} := \begin{pmatrix} \prod_{j=0}^{k/2-1} X_{i_{2j+1} i_{2(j+1)}} \overline{X_{i_{2j+3} i_{2(j+1)}}} & 0 \\ 0 & \prod_{j=0}^{k/2-1} \overline{X_{i_{2(j+1)} i_{2j+1}}} X_{i_{2(j+1)} i_{2j+3}} \end{pmatrix}. \quad (24)$$

Fixons une arête $(i, j) = (i_\alpha, i_{\alpha+1})$ associée au mot \mathbf{i} . Alors H_{ij} et H_{ji} apparaissent exactement une fois dans le produit $H_{i_1 i_2} \cdots H_{i_k i_1}$, donc il existe un unique indice β tel que $(j, i) = (i_\beta, i_{\beta+1})$. Les coefficients X_{ij} et $\overline{X_{ij}}$ associés se retrouvent dans le même produit (P_1 ou P_2) si et seulement s'il y a un nombre pair de matrices séparant les blocs H_{ij} et H_{ji} dans le mot de matrices $H_{i_1 i_2} \cdots H_{i_k i_1}$. Plus précisément, en définissant la distance entre H_{ij} et H_{ji} par $d(H_{ij}, H_{ji}) = \beta - \alpha - 1$, on a que X_{ij} et son conjugué $\overline{X_{ij}}$ interviennent dans le même produit si et seulement si $d(H_{ij}, H_{ji})$ est pair. De même pour les coefficients X_{ji} et $\overline{X_{ji}}$.

Il reste alors à se rappeler que $i_1 \dots i_k i_1$ est le contour d'un arbre. Ceci implique que la condition de parité sur $d(H_{ij}, H_{ji})$ est vérifiée pour toute arête $\{i, j\}$ de ce mot. Ainsi, P_1 (resp. P_2) est un produit de $k/2$ variables aléatoires indépendantes de même loi que $|X_{11}|^2$. En prenant l'espérance, on obtient l'énoncé du lemme. \square

On en déduit que

$$M(k, 0) := \lim_{n \rightarrow +\infty} M_n(k, 0) = \frac{1}{2} |\mathcal{W}_{k/2, k/2+1}^0| \operatorname{Tr} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

L'ensemble $\mathcal{W}_{k/2+1, k/2}^0$ étant en bijection avec l'ensemble des arbres enracinés possédant $k/2 + 1$ arêtes, $M(k, 0)$ est égal à $\operatorname{Cat}(k/2)$, le $k/2$ -ème nombre de Catalan. Rappelons le résultat suivant :

Proposition 7.1. *La loi semi-circulaire ρ est la mesure de probabilité définie par la densité (par rapport à la mesure de Lebesgue)*

$$x \mapsto \frac{1}{2\pi} \sqrt{4 - x^2} \mathbf{1}_{\{|x| \leq 2\}}.$$

Ses moments d'ordre impairs sont nuls car la fonction de x ci-dessus est impaire. Les moments d'ordre pair de ρ sont donnés par

$$\forall k \geq 1, \quad \int x^{2k} \rho(dx) = \operatorname{Cat}(k),$$

où $\operatorname{Cat}(k) = \frac{1}{k+1} \binom{2k}{k}$ est le k -ème nombre de Catalan.

On en déduit que les moments de $\mathbf{E}\tilde{\nu}_{n-1/2X}$ convergent vers les moments de la loi semi-circulaire. Cette dernière étant caractérisée par ses moments, on en déduit que $\mathbf{E}\tilde{\nu}_{n-1/2X}$ converge étroitement vers la loi semi-circulaire par le théorème 30.2 de [3]. D'où la convergence et l'universalité dans le cas $z = 0$.

Remarque 7.3. *Les mesures $\mathbf{E}\nu_{n-1/2X}$ convergent étroitement vers la loi du quart de cercle*

$$\frac{1}{\pi} \sqrt{4 - x^2} \mathbf{1}_{\{0 \leq x \leq 2\}} dx.$$

Le changement de variable $y = x^2$ nous permet de retrouver la loi de Marcenko-Pastur qui est la limite des mesures spectrales de la suite $n^{-1}XX^*$:

$$\frac{1}{2y} \sqrt{(4 - y)y} \mathbf{1}_{\{0 \leq y \leq 4\}} dy.$$

7.3 Le cas général

Nous supposons ici que $B \geq 1$. Le produit de matrices générique intervenant dans le calcul du moment (23) est

$$L_{i_1 i_2} \cdots L_{i_{\alpha_1} i_{\alpha_1+1}} \cdots L_{i_{\alpha_j} i_{\alpha_j+1}} \cdots L_{i_{\alpha_B} i_{\alpha_B+1}} \cdots L_{i_k i_1},$$

pour des indices $\{\alpha_j\}_{1 \leq j \leq B}$ tels que $i_{\alpha_j} = i_{\alpha_j+1}$. Le produit se réécrit

$$\begin{aligned} \Pi := & \begin{pmatrix} 0 & X_{i_1 i_2} \\ \overline{X}_{i_2 i_1} & 0 \end{pmatrix} \cdots \begin{pmatrix} 0 & X_{i_{\alpha_1} i_{\alpha_1+1}} \\ \overline{X}_{i_{\alpha_1+1} i_{\alpha_1}} & 0 \end{pmatrix} - \begin{pmatrix} 0 & z \\ \overline{z} & 0 \end{pmatrix} \cdots \\ & \cdots \begin{pmatrix} 0 & X_{i_{\alpha_B} i_{\alpha_B+1}} \\ \overline{X}_{i_{\alpha_B+1} i_{\alpha_B}} & 0 \end{pmatrix} - \begin{pmatrix} 0 & z \\ \overline{z} & 0 \end{pmatrix} \cdots \begin{pmatrix} 0 & X_{i_k i_1} \\ \overline{X}_{i_1 i_k} & 0 \end{pmatrix}. \end{aligned}$$

Pour tout sous-ensemble \mathcal{J} de $\{1, \dots, B\}$, introduisons

$$\Pi(\mathcal{J}) = H_{i_1 i_2} \cdots M_{i_{\alpha_1} i_{\alpha_1+1}}^{\mathcal{J}} \cdots M_{i_{\alpha_j} i_{\alpha_j+1}}^{\mathcal{J}} \cdots M_{i_{\alpha_B} i_{\alpha_B+1}}^{\mathcal{J}} \cdots H_{i_k i_1},$$

où

$$M_{i_{\alpha_j} i_{\alpha_j+1}}^{\mathcal{J}} = H_{i_{\alpha_j} i_{\alpha_j+1}} \mathbf{1}_{j \notin \mathcal{J}} - Z \mathbf{1}_{j \in \mathcal{J}} \quad \text{et} \quad Z = \begin{pmatrix} 0 & z \\ \bar{z} & 0 \end{pmatrix}.$$

En développant le produit définissant Π ci-dessus, on obtient

$$\Pi = \sum_{\mathcal{J} \subset \{1, \dots, B\}} \Pi(\mathcal{J}).$$

On appellera contribution asymptotique associée à $\Pi(\mathcal{J})$ la limite de

$$\text{Tr} \frac{1}{2n} C(\mathbf{i}) \mathbf{E}[\Pi(\mathcal{J})].$$

Lemme 7.3. *Soit $\mathcal{J} \subset \{1, \dots, B\}$. La contribution asymptotique de $\Pi(\mathcal{J})$ est nulle si B est impair ou si B est pair et $\mathcal{J} \not\subset \{1, \dots, B\}$.*

Preuve. Soit $\mathcal{J} \subset \{1, \dots, B\}$. Par définition et par le lemme 7.1 on a

$$\|\Pi(\mathcal{J})\| \leq C n^{-(k-|\mathcal{J}|)/2}$$

pour une constante $C > 0$. Par ailleurs, le graphe associé au mot $i_1 \dots i_k i_1$ a autant de sommet que le graphe associé à ce même mot privé des lettres i_{α_j} pour tout $j \in \mathcal{J}$ puisqu'on a supposé que $i_{\alpha_j} = i_{\alpha_j+1}$. Notons \mathbf{i}' ce nouveau mot et G' le graphe associé. Les arêtes e de ce dernier correspondent aux termes H_e présents dans $\Pi(\mathcal{J})$. Comme $\mathbf{E}H_e = 0$ chacune de ces arêtes doit être parcourue au moins deux fois dans \mathbf{i}' . Ceci entraîne la majoration $E' \leq \lfloor (k - |\mathcal{J}|)/2 \rfloor$ puis $|V'| \leq \lfloor (k - |\mathcal{J}|)/2 \rfloor + 1$ par connexité de G' . On en déduit l'équivalent $C(\mathbf{i}) \sim n^{\lfloor (k-B)/2 \rfloor}$ et l'estimation

$$\left\| \frac{1}{2n} C(\mathbf{i}) \Pi(\mathcal{J}) \right\| = O\left(n^{\lfloor (k-B)/2 \rfloor - (k-|\mathcal{J}|)/2} \right).$$

En particulier si B est impair le terme ci-dessus converge vers 0, de même lorsque B est pair et $\mathcal{J} \not\subset \{1, \dots, B\}$. \square

Par conséquent, seul le terme $\Pi(B) := \Pi(\{1, \dots, B\})$ pour B pair est susceptible de donner lieu à une contribution asymptotique non nulle. Remarquons que dans ce cas, le graphe G' que l'on a introduit dans la preuve précédente est un arbre puisque $V' = E' + 1$. Le mot \mathbf{i} est donc le contour d'un arbre planaire auquel on a ajouté B boucles (c.f. figure 1). Chaque arête de \mathbf{i}' est parcourue exactement deux fois, et de la même manière que dans la preuve du lemme 7.2, on a effectivement une contribution si et seulement si les blocs H_{ij} et H_{ji} sont séparés par un nombre pair de matrices dans le produit $\Pi(B)$, pour toute arête $\{i, j\}$ de \mathbf{i}' . À la différence du cas $z = 0$, cette condition n'est pas toujours remplie à cause de la présence des matrices Z faisant intervenir z .

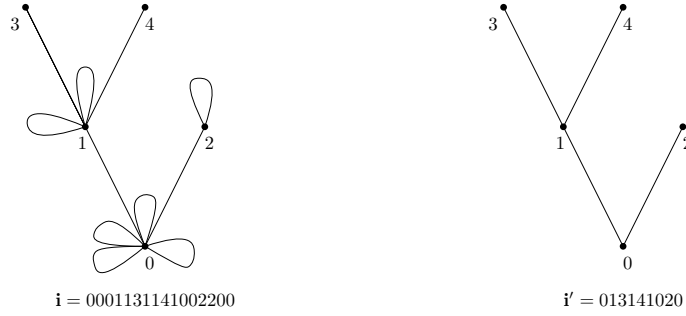


FIGURE 1 – Exemple d’arbre à boucles

Lemme 7.4. *On a que $d(H_{ij}, H_{ji})$ est paire pour toute arête $\{i, j\} \in E'$ si et seulement si chaque sommet de G possède un nombre pair de boucles.*

Remarque 7.4. *La distance d , introduite dans la preuve du lemme 7.2, est prise dans le mot \mathbf{i} donc prend en compte les blocs Z .*

Preuve. Soit $\{i, j\}$ une arête de E' . La distance entre H_{ij} et H_{ji} dans \mathbf{i}' est paire car G' est un arbre. Donc la distance entre H_{ij} et H_{ji} dans \mathbf{i} est paire si et seulement s’il existe un nombre pair de bloc faisant intervenir z entre ces deux matrices. Plaçons nous dans ce cas et notons $0, \dots, g$ les générations de l’arbre G' , et P_i la propriété “les sommets des générations $i, i+1, \dots, g$ possèdent chacun un nombre pair de boucles”. Alors P_g est vérifiée en choisissant les arêtes extrémales de l’arbre. De plus si P_{i+1} est vérifiée, alors P_i aussi. En effet soit x un sommet de la génération i . Notons y le parent de x . Il y a un nombre pair de boucles (i.e. de blocs contenant z) entre H_{yx} et H_{xy} , et tous les sommets (distincts de x) de l’arbre induit par x en possèdent un nombre pair. Donc x lui-même possède un nombre pair de boucles. On en déduit le lemme en effectuant une récurrence descendante sur les générations de G' . \square

Si B est pair et que chaque sommet de G possède un nombre pair de boucles, calculons le produit $\Pi(B)$. Comme \mathbf{i}' est le contour d’un arbre, ce dernier possède une structure récursive. Notons r_\emptyset le nombre de boucles que possède la racine de G' . On numérote les enfants de la racine $1, \dots, l$. Alors

$$\Pi(B) = Z^{r_\emptyset} H_{\emptyset 1} \Pi^{(1)} H_{1 \emptyset} Z^{r_1} H_{\emptyset 2} \Pi^{(2)} H_{2 \emptyset} \cdots Z^{r_{l-1}} H_{\emptyset l} \Pi^{(l)} H_{l \emptyset} Z^{r_l},$$

où $\Pi^{(j)}$ est le produit induit par l’enfant j de la racine, et $r_\emptyset + \cdots + r_l = r_\emptyset$. La structure d’arbre et l’indépendance des coefficients de X entraînent que les variables aléatoires $H_{\emptyset j} \Pi^{(j)} H_{j \emptyset}$ sont indépendantes. De plus, ce sont des matrices diagonales car $\Pi^{(j)}$ contient un nombre pair de termes antidiagonaux. Écrivons

$$H_{\emptyset j} \Pi^{(j)} H_{j \emptyset} = \begin{pmatrix} a_j & 0 \\ 0 & b_j \end{pmatrix}.$$

La formule (24) montre que pour tout $1 \leq j \leq l$, a_j et b_j sont identiquement distribués. Remarquons que $H_{\emptyset j} \Pi^{(j)} H_{j \emptyset} Z^{r_j}$ vaut

$$\begin{pmatrix} a_j |z|^{r_j} & 0 \\ 0 & b_j |z|^{r_j} \end{pmatrix} \mathbf{1}_{\{r_j \text{ pair}\}} + \begin{pmatrix} 0 & a_j |z|^{r_j-1} z \\ b_j |z|^{r_j-1} \bar{z} & 0 \end{pmatrix} \mathbf{1}_{\{r_j \text{ impair}\}}.$$

Supposons que $j_1, \dots, j_p \in \{1, \dots, l\}$ sont les indices associés aux r_j impairs, et que p est pair (le cas impair se traitant de manière similaire). Alors

$$\begin{aligned} \Pi(B) = Z^{r_0} & \left(\prod_{1 \leq j < j_1} \begin{pmatrix} a_j |z|^{r_j} & 0 \\ 0 & b_j |z|^{r_j} \end{pmatrix} \right) \begin{pmatrix} 0 & a_{j_1} |z|^{r_{j_1}-1} z \\ b_{j_1} |z|^{r_{j_1}-1} \bar{z} & 0 \end{pmatrix} \dots \\ & \dots \begin{pmatrix} 0 & a_{j_p} |z|^{r_{j_p}-1} z \\ b_{j_p} |z|^{r_{j_p}-1} \bar{z} & 0 \end{pmatrix} \left(\prod_{j_p < j \leq l} \begin{pmatrix} a_j |z|^{r_j} & 0 \\ 0 & b_j |z|^{r_j} \end{pmatrix} \right). \end{aligned}$$

Pour toute paire d'indices (p, q) , notons $a^{[p,q]} = a_p a_{p+1} \dots a_q$, et définissons de même les produits $b^{[p,q]}$, $a^{[p,q]}$, etc. En posant $j_0 = 0$, on obtient

$$\Pi(B) = Z^{r_0} \begin{pmatrix} \prod_{0 \leq i \leq p-2} a^{[j_i, j_{i+1}]} b^{[j_{i+1}, j_{i+2}]} & 0 \\ 0 & \prod_{0 \leq i \leq p-2} b^{[j_i, j_{i+1}]} a^{[j_{i+1}, j_{i+2}]} \end{pmatrix}.$$

L'indépendance des $H_{\emptyset_j} \Pi^{(j)} H_{j_\emptyset}$ et les égalités en loi $a_j \stackrel{(d)}{=} b_j$ nous fournissent donc la formule suivante

$$\mathbf{E} \Pi(B) = Z^{r_0} \prod_{1 \leq j \leq l} \mathbf{E} \left[H_{\emptyset_j} \Pi^{(j)} H_{j_\emptyset} \right].$$

En traitant chaque $\Pi^{(j)}$ de la même manière, on trouve

$$\mathbf{E} \Pi(B) = Z^B \mathbf{E} \left[H_{i'_1 i'_2} \dots H_{i'_{k-N} i'_1} \right] = \begin{pmatrix} |z|^B & 0 \\ 0 & |z|^B \end{pmatrix},$$

où l'on a écrit $\mathbf{i}' = i'_1 i'_2 \dots i'_{k-N} i'_1$. L'espérance de $H_{i'_1 i'_2} \dots H_{i'_{k-N} i'_1}$ vaut l'identité par le lemme 7.2.

Définition 7.1. *Pour tout $k \geq 1$ et $l \geq 0$, on introduit $T_k(2l, P)$ le cardinal de l'ensemble des arbres possédant k arêtes et $2l$ boucles telles que chaque sommet possède un nombre pair de boucles, et $T_k(2l+1, P)$ le cardinal de l'ensemble des arbres possédant k arêtes et $2l+1$ boucles telles que chaque sommet possède un nombre pair de boucles, sauf la racine qui en possède un nombre impair.*

Notons $M(k, B)$ la limite de $M_n(k, B)$ lorsque n tend vers l'infini. Le théorème suivant résume ce que l'on vient de démontrer.

Théorème 7.1. *Pour tout $k \geq 1$ et tout $1 \leq B \leq k$,*

$$M(k, B) = \begin{cases} 0 & \text{si } k \text{ est impair} \\ 0 & \text{si } k \text{ est pair et } B \text{ impair} \\ T_{\frac{k-B}{2}}(B, P) |z|^B & \text{si } k \text{ est pair et } B \text{ pair.} \end{cases}$$

7.4 Série génératrice

Nous avons vu que la suite $\mathbf{E} \nu_{n-1/2} X_{-z}$ est tendue. Notons ν un point d'accumulation. Par ce qui précède les moments de ν , notés $M(k)$, vérifient

$$M(k) = \sum_{0 \leq B \leq k} M(k, B).$$

Notons $G(z, \eta) = \sum_{k \geq 0} M(k) \eta^k$ la série génératrice des moments de ν . Alors en utilisant le théorème 7.1 on en déduit que

$$\begin{aligned} G(z, \eta) &= \sum_{k \geq 0} M(2k) \eta^{2k} \\ &= \sum_{k \geq 0} \sum_{0 \leq l \leq k} \mathbb{T}_{k-l}(2l, P) |z|^{2l} \eta^{2k} \\ &= \sum_{l \geq 0} \sum_{k-l \geq 0} \mathbb{T}_{k-l}(2l, P) |z|^{2l} \eta^{2k} \\ &= \sum_{k, l \geq 0} \mathbb{T}_k(2l, P) (|z| \eta)^{2l} \eta^{2k}. \end{aligned}$$

Notons que l'on a effectué un changement de variable dans la dernière égalité. Nous allons établir une équation sur G . Pour cela, il sera utile de raisonner sur les séries génératrices suivantes :

$$P(x, y) := \sum_{k, l \geq 0} \mathbb{T}_k(2l, P) x^l y^k \quad \text{et} \quad I(x, y) := \sum_{k, l \geq 0} \mathbb{T}_k(2l+1, I) x^l y^k.$$

Remarquons que l'on a l'égalité $G = P(|z|\eta^2, \eta^2)$.

Lemme 7.5. *Pour tout $k \geq 0$ et $l \geq 0$, on a*

$$\begin{aligned} \mathbb{T}_{k+1}(2l, P) &= \sum_{p=0}^l \sum_{k_1+k_2=k} \sum_{(l_1, l_2) \in \mathcal{C}_{p,l}} \mathbb{T}_{k_1}(l_1, P) \mathbb{T}_{k_2}(l_2, P) \\ &\quad + \sum_{p=0}^{l-1} \sum_{k_1+k_2=k} \sum_{(l_1, l_2) \in \mathcal{C}'_{p,l}} \mathbb{T}_{k_1}(l_1, P) \mathbb{T}_{k_2}(l_2, I) \end{aligned}$$

et

$$\begin{aligned} \mathbb{T}_{k+1}(2l, I) &= \sum_{p=0}^l \sum_{k_1+k_2=k} \sum_{(l_1, l_2) \in \mathcal{C}'_{p,l}} \mathbb{T}_{k_1}(l_1, P) \mathbb{T}_{k_2}(l_2, I) \\ &\quad + \sum_{p=0}^{l-1} \sum_{k_1+k_2=k} \sum_{(l_1, l_2) \in \mathcal{C}_{p,l}} \mathbb{T}_{k_1}(l_1, P) \mathbb{T}_{k_2}(l_2, P), \end{aligned}$$

où $\mathcal{C}_{p,l}$ est l'ensemble des couples d'entiers pairs (l_1, l_2) satisfaisant $l_1 + l_2 = 2l - 2p$, tandis que $\mathcal{C}'_{p,l}$ est l'ensemble des couples d'entiers (l_1, l_2) tels que l_1 est pair, l_2 est impair et $l_1 + l_2 = 2l - 2p - 1$.

Corollaire 7.1. *Les fonctions génératrices P et I vérifient*

$$\begin{cases} P &= \frac{1}{1-x} + \frac{y}{1-x} P^2 + \frac{yx}{1-x} PI \\ I &= \frac{x}{1-x} + \frac{yx}{1-x} P^2 + \frac{y}{1-x} PI. \end{cases}$$

Preuve du lemme 7.5. On montre la première formule, la preuve de la seconde étant similaire. Soient k et l deux entiers positifs. Soit T un arbre possédant $k+1$ arêtes et $2l$ boucles. Notons $0 \leq q \leq 2l$ le nombre de boucles au début

du contour de T , 0 la racine et 1 le premier enfant de la racine. Alors T se décompose en deux sous-arbres à boucles. Plus précisément, son contour s'écrit

$$\underbrace{00\dots 0}_{q \text{ fois}}01\text{Cont}(T_1)0\text{Cont}(T'),$$

où Cont désigne le contour d'un arbre à boucles, T_1 l'arbre à boucles issu de l'enfant 1 et T' l'arbre à boucles obtenu en retirant les q premières boucles de T et la première lignée de 0 (i.e. l'arête 01 et T_1). La figure 2 résume ce que l'on vient de décrire.

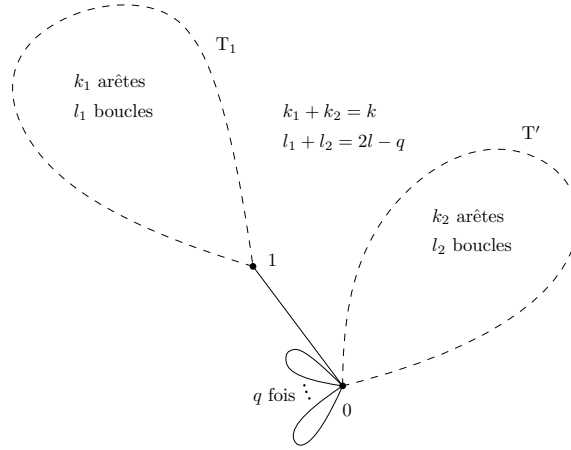


FIGURE 2 – Décomposition de l'arbre à boucle T

Notons k_1 (resp. k_2) le nombre d'arêtes et l_1 (resp. l_2) le nombre de boucles de T_1 (resp. T'). Un premier constat est que $k_1 + k_2 = k$ puisque l'arête 01 n'est pas comptée, et $l_1 + l_2 = 2l - q$. Ensuite, remarquons que T_1 possède un nombre pair de boucles sur chacun de ses sommets. Par contre, le nombre de boucles q' de la racine de T' dépend de la parité de q . En effet, puisque les sommets de T possèdent chacun un nombre pair de boucles, $q + q'$ est pair. Donc q et q' ont la même parité. Le résultat découle de toutes ces observations. \square

Preuve du corollaire 7.1. On montre seulement la première équation car la seconde résulte du même raisonnement. On commence par faire apparaître le premier terme :

$$\begin{aligned} P(x, y) &= \sum_{l \geq 0} x^l + \sum_{k \geq 1} \sum_{l \geq 0} T_k(2l, P) x^l y^k \\ &= \frac{1}{1-x} + y \sum_{k \geq 0} \sum_{l \geq 0} T_{k+1}(2l, P) x^l y^k. \end{aligned}$$

En utilisant la relation de récurrence du lemme 7.5, la somme ci-dessus se décompose en deux termes :

$$S_1 := \sum_{k \geq 0} \sum_{l \geq 0} \sum_{p=0}^l \sum_{k_1+k_2=k} \sum_{(l_1, l_2) \in \mathcal{C}_{p, l}} T_{k_1}(l_1, P) T_{k_2}(l_2, P) x^l y^k$$

et

$$S_2 := \sum_{k \geq 0} \sum_{l \geq 0} \sum_{p=0}^{l-1} \sum_{k_1+k_2=k} \sum_{(l_1, l_2) \in \mathcal{C}'_{p,l}} \mathbb{T}_{k_1}(l_1, P) \mathbb{T}_{k_2}(l_2, I) x^l y^k.$$

En intervertissant les sommes et en posant $L = l - p$, on voit que

$$\begin{aligned} S_1 &= \sum_{k \geq 0} \sum_{p \geq 0} x^p \sum_{L \geq 0} \sum_{k_1+k_2=k} \sum_{(l_1, l_2) \in \mathcal{C}'_{p,l}} \mathbb{T}_{k_1}(l_1, P) \mathbb{T}_{k_2}(l_2, P) x^l y^k \\ &= \left(\sum_{p \geq 0} x^p \right) \sum_{k \geq 0} \sum_{L \geq 0} \sum_{k_1+k_2=k} \sum_{(l_1, l_2) \in \mathcal{C}'_{p,l}} \mathbb{T}_{k_1}(l_1, P) \mathbb{T}_{k_2}(l_2, P) x^l y^k \\ &= \frac{1}{1-x} P(x, y)^2. \end{aligned}$$

Pour S_2 , on introduit l'_2 l'unique entier pair tel que $l_2 = l'_2 + 1$. On remarque alors que $(l_1, l'_2) \in \mathcal{C}_{p-1, l}$, donc en posant $L = l - p - 1$, on obtient :

$$\begin{aligned} S_2 &= \sum_{k \geq 0} \sum_{p \geq 0} x^{p+1} \sum_{L \geq 0} \sum_{k_1+k_2=k} \sum_{(l_1, l'_2) \in \mathcal{C}'_{p,l}} \mathbb{T}_{k_1}(l_1, P) \mathbb{T}_{k_2}(l'_2 + 1, I) x^l y^k \\ &= \left(\sum_{p \geq 0} x^{p+1} \right) \sum_{k \geq 0} \sum_{L \geq 0} \sum_{k_1+k_2=k} \sum_{(l_1, l'_2) \in \mathcal{C}'_{p,l}} \mathbb{T}_{k_1}(l_1, P) \mathbb{T}_{k_2}(l'_2 + 1, I) x^l y^k \\ &= \frac{x}{1-x} P(x, y) I(x, y). \end{aligned}$$

□

Le corollaire 7.1 permet de déduire une équation cubique sur P . En effet, en utilisant que $I - xP = yIP$ dans la première équation, on obtient après simplification

$$y^2 P^3 - 2yP^2 + (y - x + 1)P - 1 = 0.$$

Donc la série génératrice G vérifie

$$\eta^4 G^3 - 2\eta^2 G^2 + (\eta^2 - (\eta|z|)^2 + 1)G - 1 = 0.$$

La transformée de Stieltjes de ν est reliée à la série génératrice G par la formule $G(\eta) = -\eta^{-1} m_\nu(\eta^{-1})$. Après un calcul élémentaire, on obtient finalement

$$m_\nu(\eta) = \frac{m_\nu(\eta) + \eta}{|z|^2 - (m_\nu(\eta) + \eta)^2}.$$

On reconnaît l'équation 22. Le lemme 1.2 en découle.

Remarque 7.5. *On pouvait conclure à l'existence et universalité sans avoir recours à cette sous-partie sur les séries génératrices. En effet, les moments limites $M(k)$ vérifient la condition de Carleman :*

$$\sum_{k \geq 0} M(2k)^{-\frac{1}{2k}} = \infty. \quad (25)$$

Par le [2, lemme B.3], on en déduit que la mesure de probabilité limite est unique, puisqu'elle est entièrement caractérisée par ses moments. Pour obtenir 25, on peut par exemple remarquer que $M(2k) \geq \text{Cat}(k)$.

A Analyse matricielle

Soit H une matrice hermitienne de taille n . Quitte à effectuer un changement de base on peut supposer que $H = \text{diag}(\lambda_1(A), \dots, \lambda_n(A))$. Remarquons que pour $k \in \{1, \dots, n\}$:

$$\begin{aligned} \max_{\substack{x \neq 0 \\ x \perp w_1, \dots, w_{k-1}}} \frac{x^* H x}{x^* x} &= \max_{\substack{\|x\|_2=1 \\ x \perp w_1, \dots, w_{k-1}}} \sum_{i=1}^n \lambda_i(H) |x_i|^2 \\ &\geq \max_{\substack{x \neq 0 \\ x \perp w_1, \dots, w_{k-1} \\ x_j=0 \text{ et } \forall j < k}} \sum_{i=1}^n \lambda_i(H) |x_i|^2 \\ &\geq \lambda_k(H), \end{aligned}$$

où les $w_i \in \mathbb{C}^n$ sont quelconques. En prenant le minimum, la borne inférieure est atteinte dans les inégalités ci-dessus. Par exemple pour les choix de vecteurs $w_i = (0, \dots, 0, 1, 0, \dots, 0)$ où le 1 est en i -ème position, qui autorisent $x = (0, \dots, 0, 1, 0, \dots, 0)$ avec le 1 en k -ème position. On obtient les relations de Courant-Fisher :

$$\forall k \in \{1, \dots, n\}, \quad \lambda_k(A) = \min_{w_1, \dots, w_{n-k} \in \mathbb{C}^n} \max_{\substack{x \neq 0 \\ x \perp w_1, \dots, w_{n-k}}} \frac{x^* H x}{x^* x}. \quad (26)$$

Pour une matrice complexe A de taille $n \times n$, on en déduit une formule variationnelle pour les valeurs singulières : pour tout $k \in \{1, \dots, \min\{m, n\}\}$:

$$s_k(A) = \min_{w_1, \dots, w_{n-k} \in \mathbb{C}^n} \max_{\substack{\|x\|_2=1 \\ x \perp w_1, \dots, w_{n-k}}} \|Ax\|_2. \quad (27)$$

Notons en particulier que $s_1(A) = \|A\|$. Ces relations permettent d'obtenir un résultat d'entrelacement entre les valeurs singulières de A et de A_r qui désigne la matrice A privée de r de ses colonnes ou de ses lignes.

Théorème A.1. *En convenant que $s_k(X) = 0$ pour toute matrice X de taille $i \times j$ et tout $k > \min\{i, j\}$, on a l'entrelacement suivant :*

$$\forall 1 \leq k \leq \min\{m, n\}, \quad s_k(A) \leq s_k(A_r) \leq s_{k+r}(A).$$

Le théorème A.1 permet de comparer les valeurs propres et les valeurs singulières de A . Plus précisément, on a les inégalités de Weyl :

Théorème A.2. *Supposons que A est de taille $n \times n$. Pour tout $1 \leq k \leq n$:*

$$\prod_{i=1}^k |\lambda_i(A)| \leq \prod_{i=1}^k s_i(A).$$

Preuve. Fixons un entier positif $k \leq n$. On utilise la décomposition de Schur : il existe une matrice unitaire U telle que $UAU^* = D = \text{diag}(\lambda_1(A), \dots, \lambda_n(A))$. Notons $U_k \in \mathcal{M}_{n \times k}(\mathbb{C})$ la matrice constituée des k premières colonnes de U . En effectuant un produit par blocs, on voit que

$$D = \begin{pmatrix} U_k A U_k^* & 0 \\ 0 & \text{diag}(\lambda_{k+1}(A), \dots, \lambda_n(A)) \end{pmatrix}.$$

En particulier

$$\prod_{i=1}^k |\lambda_i(A)| = |\det(U_k A U_k^*)| = \prod_{i=1}^k s_i(U_k A U_k^*).$$

On remarque finalement que $U_k A U_k^*$ est obtenue en retirant $n-k$ lignes puis $n-k$ colonnes à $U A U^*$. Le théorème A.1 implique donc que $s_i(U_k A U_k^*) \leq s_i(U A U^*)$ pour tout $1 \leq i \leq k$. La formule 27 implique que les valeurs singulières d'une matrice sont invariantes par conjugaison par une matrice unitaire. Donc pour tout $1 \leq i \leq n$ on a $s_i(U A U^*) = s_i(A)$. Le théorème en découle. \square

On admettra la généralisation suivante, dont une preuve pour être trouvée dans [11, théorème 3.3.13].

Théorème A.3. *Soit f une fonction telle que $t \mapsto f(e^t)$ est croissante et convexe sur l'intervalle $[s_n(A), s_1(A)]$. Alors pour tout $k \in \{1, \dots, n\}$:*

$$\sum_{i=1}^k f(|\lambda_i(A)|) \leq \sum_{i=1}^k f(s_i(A)).$$

En particulier, le choix $f(t) = t^2$ entraîne l'inégalité 1.

Donnons deux résultats concernant la comparaison des valeurs singulières de deux matrices. Le premier est une conséquence de la formule variationnelle 27, connu sous le nom des inégalités de Lidskii :

Théorème A.4. *Soient A et B deux matrices complexes de taille $m \times n$. Notons $q = \min\{m, n\}$. Alors pour tous indices $1 \leq i, j \leq q$ tels que $i + j - 1 \leq q$, on a*

$$s_{i+j-1}(A+B) \leq s_i(A) + s_j(B).$$

Prenons A et B deux matrices carrées de taille $n \times n$ et notons F_A et F_B les fonctions de répartition des mesures spectrales μ_A et μ_B . Une conséquence des inégalités de Lidskii est l'estimation suivante :

$$\|F_A - F_B\|_\infty \leq \frac{\text{rg}(A-B)}{n}. \quad (28)$$

Le second résultat est appelé inégalité de Hoffman-Wielandt, dont on pourra trouver une preuve dans [1, lemme 2.1.19].

Théorème A.5. *Soient A et B deux matrices complexes de taille $n \times n$. En notant $\|X\|_2 = \sqrt{\text{Tr}(X X^*)}$, on a*

$$\sum_{i=1}^n |s_i(A) - s_i(B)|^2 \leq \|A - B\|_2^2.$$

Pour terminer, le lemme suivant nous a permis d'obtenir l'uniforme intégrabilité du logarithme. Une preuve pourra être trouvée dans [17, Lemme A.4].

Lemme A.1. *Soient $m \leq n$ et A une matrice de taille $m \times n$ et de rang m . Notons L_i les lignes de A et d_i la distance entre L_i et le sous-espace engendré par les autres lignes de A . Alors*

$$\sum_{i=1}^m s_i(A)^{-2} = \sum_{i=1}^m d_i^{-2}.$$

B Concentration de la mesure

Nous avons utilisé trois résultats généraux de concentration de la mesure dans ce document. Leurs preuves pourront être trouvées dans [15, partie 2.1] ou dans [5] qui fournit les constantes exactes. Le premier résultat est l'inégalité de Hoeffding :

Théorème B.1. *Soient X_1, \dots, X_n des v.a. bornées et indépendantes. Pour tout $1 \leq i \leq n$, notons $d_i := \max X_i - \min X_i$. Posons $d^2 = d_1^2 + \dots + d_n^2$ et $S_n = X_1 + \dots + X_n$. Alors pour tout $t \geq 0$:*

$$\mathbf{P}(|S_n - \mathbf{E}S_n| > t) \leq \exp\left(\frac{-2t^2}{d^2}\right).$$

Le deuxième résultat est l'inégalité de McDiarmid :

Théorème B.2. *Soient X_1, \dots, X_n des variables aléatoires indépendantes à valeurs dans des espaces $\mathcal{X}_1, \dots, \mathcal{X}_n$. Soit $G : \mathcal{X}_1 \times \dots \times \mathcal{X}_n \rightarrow \mathbb{R}$ une fonction mesurable. Supposons que pour tout $1 \leq k \leq n$ il existe une constante $c_k > 0$ telle que pour tout $x_1, \dots, x_k, x'_k, \dots, x_n \in \mathcal{X}_1 \times \dots \times \mathcal{X}_k^2 \times \dots \times \mathcal{X}_n$:*

$$|G(x_1, \dots, x_k, \dots, x_n) - G(x_1, \dots, x'_k, \dots, x_n)| \leq c_k.$$

Alors pour tout $t \geq 0$, en notant $c^2 = c_1^2 + \dots + c_n^2$:

$$\mathbf{P}\left(|G(X_1, \dots, X_n) - \mathbf{E}G(X_1, \dots, X_n)| \geq t\right) \leq 2 \exp\left(\frac{-2t^2}{c^2}\right).$$

Enfin, on a eu recours à l'inégalité de Talagrand :

Théorème B.3. *Soit $r > 0$, $\mathcal{D} = \{z \in \mathbb{C} : |z| \leq r\}$, et X_1, \dots, X_n des variables aléatoires indépendantes à valeurs dans \mathcal{D} . Soit $F : \mathcal{D}^n \rightarrow \mathbb{R}$ une fonction 1-lipschitzienne. En notant M la médiane de la variable aléatoire $F(X_1, \dots, X_n)$, on a, pour tout $t \geq 0$:*

$$\mathbf{P}\left(|F(X_1, \dots, X_n) - M| \geq t\right) \leq 4 \exp\left(\frac{-t^2}{16r^2}\right).$$

Références

- [1] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118. Cambridge university press, 2010.
- [2] Zhidong Bai and Jack W. Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20. Springer, 2010.
- [3] P. Billingsley. *Probability and Measure*. New York [etc.] : Wiley, 1995.
- [4] Charles Bordenave, Djalil Chafaï, et al. Around the circular law. *Probability surveys*, 9, 2012.
- [5] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities : A nonasymptotic theory of independence*. Oxford university press, 2013.
- [6] Percy Deift. *Orthogonal polynomials and random matrices : a Riemann-Hilbert approach*, volume 3. American Mathematical Soc., 2000.
- [7] R. Brent Dozier and Jack W. Silverstein. On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices. *Journal of Multivariate Analysis*, 98(4) :678–694, 2007.
- [8] William Feller. An introduction to probability and its applications, vol. ii. *Wiley, New York*, 1971.
- [9] Jean Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *Journal of Mathematical Physics*, 6(3) :440–449, 1965.
- [10] Vyacheslav L. Girko. Circular law. *Teoriya Veroyatnostei i ee Primeneniya*, 29(4) :669–679, 1984.
- [11] Roger A. Horn and Charles R. Johnson. *Topics in matrix analysis*. Cambridge University Press, 1994.
- [12] Chii-Ruey Hwang et al. A brief survey on the spectral radius and the spectral distribution of large random matrices with iid entries. *Contemporary Mathematics AMS. 50*, pages 145–152, 1986.
- [13] John Lee. *Introduction to Smooth Manifolds*, volume 218. Springer Science & Business Media, 2012.
- [14] Madan Lal Mehta. *Random matrices*, volume 142. Academic press, 2004.
- [15] Terence Tao. *Topics in random matrix theory*, volume 132. American Mathematical Society Providence, RI, 2012.
- [16] Terence Tao and Van Vu. Random matrices : the circular law. *Communications in Contemporary Mathematics*, 10(02) :261–307, 2008.
- [17] Terence Tao, Van Vu, and Manjunath Krishnapur. Random matrices : universality of esds and the circular law. *The Annals of Probability*, 38(5) :2023–2065, 2010.
- [18] Frank W. Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94. Springer Science & Business Media, 2013.