Astérisque

ALICE GUIONNET

Grandes matrices aléatoires et théorèmes d'universalité [d'après Erdos, Schlein, Tao, Vu et Yau]

Astérisque, tome 339 (2011), Séminaire Bourbaki,
exp. n^{o} 1019, p. 203-237

<http://www.numdam.org/item?id=AST_2011__339__203_0>

Séminaire BOURBAKI Mars 2010
62e année, 2009-2010, n° 1019, p. 203 à 237

GRANDES MATRICES ALÉATOIRES ET THÉORÈMES D'UNIVERSALITÉ [d'après Erdős, Schlein, Tao, Vu et Yau]

par **Alice GUIONNET**

INTRODUCTION

Les grandes matrices aléatoires sont tout d'abord apparues en statistique dans les travaux de Wishart [**37**] pour décrire des tableaux de données aléatoires, puis en physique dans les travaux de Wigner [**36**] et Dyson pour approximer un opérateur modélisant le Hamiltonien d'un noyau excité. Montgomery, soutenu par les simulations d'Odlyzko, a conjecturé que leurs valeurs propres étaient également reliées aux fameux zéros de la fonction de Riemann sur la droite critique, les espacements de ceux-ci loin de l'axe réel étant distribués de façon identique. Les matrices aléatoires sont depuis intervenues dans de nombreux contextes, et leur spectre a attiré un intérêt grandissant. On a notamment étudié sa convergence globale, la distribution des espacements des valeurs propres à l'intérieur du spectre et les fluctuations des valeurs propres extrêmes, quand la taille des matrices tend vers l'infini. Le cas de matrices aléatoires à coefficients gaussiens s'est révélé plus facile à appréhender, compte tenu de formules explicites pour la loi jointe des valeurs propres. Néanmoins, comme pour le théorème central limite, il est attendu que ces théorèmes limites sont universels et ne dépendent que très peu de la nature des coefficients. Le but de cet exposé est de montrer les récents efforts fournis pour étudier cette universalité. Nous considérerons les matrices dites de Wigner qui sont hermitiennes et avec des coefficients indépendants et équidistribués (modulo l'hypothèse de symétrie). Plus précisément, on notera $\mathbf{X}_N$ une matrice hermitienne $N \times N$ dont les coefficients $(X_{ij})_{1\le i<j\le N}$ (resp. $(X_{ii})_{1\le i\le N}$) sont indépendants et de loi ν (resp. μ) sur $\mathbb{C}$ (resp. $\mathbb{R}$) et on étudiera les valeurs propres $(\lambda_1, \ldots, \lambda_N) \in \mathbb{R}^N$ de cette matrice. Nous supposerons que ν et μ sont centrées et telles que $\int |x|^2 d\nu(x) = \int x^2 d\mu(x) = 1$. Dans ce cas il est connu depuis les travaux de Wigner dans les années cinquante que le nombre moyen de valeurs propres de

$\mathbf{Y}_N = N^{-\frac{1}{2}}\mathbf{X}_N$ tombant dans un intervalle $[a,b]$ de la droite réelle est de l'ordre de $N\sigma([a,b])$ où σ est la loi semi-circulaire ;

(1)
$$\lim_{N\to\infty}\frac{1}{N}\sharp\{i : N^{-\frac{1}{2}}\lambda_i \in [a,b]\} = \sigma([a,b]) = \int_a^b \rho_{sc}(x)dx := \int_{[a\vee -2, b\wedge 2]} \sqrt{4-x^2}\frac{dx}{2\pi} \quad \text{p.s.}$$

L'étude des propriétés locales du spectre s'est avérée beaucoup plus complexe et n'a été entreprise que très récemment dans une certaine généralité sur les distributions μ et ν des coefficients. Nous discuterons tout particulièrement le résultat suivant. Supposons que μ et ν ont des queues de distributions sous-exponentielles, dans le sens où il existe des constantes $C, C' > 0$ telles que, pour tout $t > C$

$$\nu\left(|z| > t^{C'}\right) \le e^{-t} \qquad \mu\left(|z| > t^{C'}\right) \le e^{-t}. \tag{2}$$

Si ν est une mesure sur $\mathbb{C}$, nous supposerons que les lois de la partie imaginaire et de la partie réelle sont indépendantes, et, si ν est une mesure sur la droite réelle, que $\int x^3 d\nu(x)$ s'annule ou que ν a au moins trois points dans son support. Nous distinguerons ces deux cas par un paramètre β qui sera égal à deux (resp. un) si le support est dans $\mathbb{C}$ (resp. $\mathbb{R}$).

THÉORÈME 0.1. — *Soit B un intervalle compact de $\mathbb{R}$ et $x \in (-2,2)$.*

La probabilité qu'aucune valeur propre de $\mathbf{X}_N$ ne tombe dans un intervalle $N^{\frac{1}{2}}x + N^{-\frac{1}{2}}\rho_{sc}(x)^{-1}B$ converge quand la dimension N tend vers l'infini vers une limite non triviale qui ne dépend que de β et de B (voir la description de cette limite en (11) *dans le cas où $\beta = 2$).*

La probabilité qu'aucune valeur propre de $\mathbf{X}_N$ ne tombe dans l'ensemble $2N^{\frac{1}{2}} + N^{-\frac{1}{6}}B$ converge quand la dimension N tend vers l'infini vers une limite non triviale qui ne dépend que de β et de B (voir (12) *dans le cas où $\beta = 2$).*

Le second résultat décrit les fluctuations des valeurs propres de $\mathbf{Y}_N$ au bord du spectre (proche de 2), qui sont d'ordre $N^{-\frac{2}{3}}$, alors que le premier concerne celles des valeurs propres à l'intérieur du spectre, qui sont beaucoup plus petites, d'ordre N^{-1}. Ce dernier point permet d'étudier la convergence de la loi des espacements des valeurs propres à l'intérieur du spectre ; on a par exemple le résultat suivant (voir [**24**, p.84] et [**1**, Theorem 4.2.49] avec [**7**] pour le cas gaussien et [**10**, Theorem 3] pour le cas général) dans le cas où ν est une mesure sur $\mathbb{C}$.

THÉORÈME 0.2. — *Pour tout $x \in (-2,2)$, toute fonction l_N tendant vers l'infini plus lentement que N, le nombre moyen de valeurs propres de $\mathbf{Y}_N$ à distance inférieure à $l_N/N\rho_{sc}(x)$ de x et dont les espacements sont plus petits que $s/\rho_{sc}(x)N$ est approximativement égal à $P_{\mathrm{Gaudin}}([0,s])l_N$, où P_{Gaudin} est la distribution de Gaudin.*

Le second point du théorème 0.1 donne les fluctuations de la plus grande valeur propre.

THÉORÈME 0.3. — *Pour tout $t \in \mathbb{R}$, la plus grande valeur propre λ_1 de $\mathbf{X}_N$ est telle que*

$$\lim_{N\to\infty} P\left(N^{\frac{2}{3}}(\frac{\lambda_1}{\sqrt{N}} - 2) \leq t\right) = F_\beta(t)$$

avec F_β la fonction de répartition de la loi de Tracy-Widom.

Ces théorèmes ont été démontrés pour des matrices gaussiennes depuis une quinzaine d'années, notamment après les travaux de M. Mehta [**24**], P. Forrester [**17**] et C. Tracy et H. Widom [**33, 34**], grâce à la formule explicite de la loi jointe des valeurs propres et sa propriété de loi déterminantale. Les premiers pas vers l'universalité ont été franchis par K. Johansson [**21**] qui a considéré des matrices obtenues comme somme d'une matrice gaussienne et d'une matrice de Wigner plus générale, de nouveau grâce à une structure déterminantale de la loi jointe des valeurs propres. Dans ce dernier cas, il est primordial de supposer que les coefficients sont complexes hors de la diagonale (cas hermitien). L'universalité des fluctuations au bord du spectre a été démontrée par des techniques de moments et de combinatoire par A. Soshnikov. Cette approche ne pouvant permettre d'analyser les espacements des valeurs propres à l'intérieur du spectre, il a fallu attendre l'été dernier pour que cette question soit enfin résolue par deux équipes indépendantes, d'une part L. Erdős, B. Schlein et H.T. Yau et leurs collaborateurs, et d'autre part T. Tao et V. Vu. Les approches de ces deux équipes, quoique différentes, reposent sur des raisonnements beaucoup plus probabilistes et montrent une sorte de régularité des espacements des valeurs propres en fonction des coefficients de la matrice. Les travaux presque simultanés de T. Tao et V. Vu [**32**] et de L. Erdős, B. Schlein, S. Péché, J. Ramirez et H.T. Yau [**9**] ont permis d'étendre ces résultats d'une part sous la condition que les quatre premiers moments de ν et μ sont les mêmes que ceux de la gaussienne (ou plus généralement que ceux d'une matrice étudiée par K. Johansson [**21**]), et d'autre part, sous des conditions de régularité de la densité de ν et μ. Alliant ces résultats, L. Erdős, J. Ramirez, B. Schlein, T. Tao, V. Vu et H.T. Yau [**10**] ont pu démontrer l'universalité dans le cas hermitien sous des conditions de queues exponentielles seulement. Afin de pouvoir considérer des ensembles plus généraux, et en particulier les matrices à coefficients réels (cas symétrique), L. Erdős, B. Schlein et H.T. Yau [**14**] ont introduit une approche dynamique qui permet de largement généraliser les résultats de K. Johansson [**21**] sans utiliser de formule explicite ou de structure déterminantale. Ce point de vue a été étendu dans les travaux récents de L. Erdős, H.T. Yau et Y. Yin [**15, 16**] pour étudier les matrices de Wishart et à bande, et réduire considérablement les hypothèses sur les troisième et quatrième moments. L'universalité est attendue à l'intérieur (resp. au bord) du spectre seulement sous une condition de second (resp. quatrième) moment fini, voir **[22]** pour ce résultat dans le cas de coefficients complexes ayant une composante gaussienne. Le but de ces notes est de présenter les idées des preuves de ces différents résultats, en précisant particulièrement celles des travaux de L. Erdős, B. Schlein et H.T. Yau et T. Tao et V. Vu. Nous évoquerons également leurs liens avec les propriétés de délocalisation des vecteurs propres de X_N. Nous ne parlerons pas des diverses généralisations à d'autres ensembles de matrices tels que les matrices de Wishart **[25]**, **[31]** et **[15]** ou les matrices à bande généralisées **[16]**. Par ailleurs, les résultats que nous développerons concernent la situation de coefficients indépendants, totalement différente par exemple de celle rencontrée pour les modèles de matrices avec un potentiel non quadratique pour laquelle nous renvoyons le lecteur à **[6]**. Un des problèmes qui semble encore ouvert pour des matrices à coefficients indépendants concerne le cas de matrices à bande dont les coefficients sont nuls en dehors d'une bande de largeur W bien plus petite que N. Il est attendu que pour W bien plus grand que $\sqrt{N}$ les vecteurs propres sont délocalisés comme pour les matrices de Wigner, voir le corollaire 4.4, alors que pour $W \ll \sqrt{N}$, ils sont localisés. Cette transition serait un modèle simple de la transition d'Anderson.

1. MATRICES GAUSSIENNES ET LOIS DÉTERMINANTALES

Nous considérons dans cette section le cas de matrices de Wigner $\mathbf{G}_N$ dites du GUE (pour Gaussian Unitary Ensemble), c'est-à-dire les matrices $\mathbf{X}_N$ telles que ν (resp. μ) est une mesure gaussienne standard sur $\mathbb{C}$ (resp. $\mathbb{R}$) ; pour toute fonction test f

(3)
$$\int f(x)d\mu(x) = \int_{\mathbb{R}} f(x)\frac{dxe^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \quad \text{et} \quad \int f(z)d\nu(z) = \int f(\frac{x+iy}{\sqrt{2}})d\mu(x)d\mu(y).$$

Comme nous allons le voir, la particularité de ces matrices est que la loi jointe de leurs valeurs propres est explicite et est une loi déterminantale. Cette structure très particulière, que nous allons bientôt décrire, permet une analyse relativement aisée des propriétés locales du spectre.

Il n'est pas difficile de constater que la loi de la matrice $\mathbf{G}_N$ est laissée invariante par la conjugaison $U\mathbf{G}_N U^*$ par une matrice unitaire, et par conséquent que les valeurs propres $D^N = (\lambda_1, \ldots, \lambda_N)$ de $\mathbf{G}_N$ sont indépendantes de ses vecteurs propres, dont la matrice V^N suit la mesure uniforme sur le groupe unitaire. En calculant le Jacobien du changement de coordonnés qui à $\mathbf{G}_N$ associe D^N et une paramétrisation de V^N, on trouve que la loi des valeurs propres est décrite par la mesure de probabilité

$$d\mathcal{P}_N(x_1,\dots,x_N) = \frac{1}{Z_N}\prod_{1\le i<j\le N}|x_i-x_j|^2\prod_{i=1}^N e^{-x_i^2/2}dx_i\,. \tag{4}$$

Cette formule exacte permet d'étudier les lois jointes de p valeurs propres et de montrer qu'elles ont une forme très particulière ; elles sont décrites par le déterminant d'un noyau. On notera $\mathcal{P}_{p,N}$ la loi de p valeurs propres non ordonnées donnée pour tout $f\in C_b(\mathbb{R}^p)$, par

$$\int f(\theta_1,\dots,\theta_p)d\mathcal{P}_{p,N}(\theta_1,\dots,\theta_p) = \int f(\theta_1,\dots,\theta_p)d\mathcal{P}_N(\theta_1,\dots,\theta_N)\,.$$

Le résultat principal est le suivant.

LEMME 1.1. — *Pour tout $p\le N$, la loi $\mathcal{P}_{p,N}$ est absolument continue par rapport à la mesure de Lebesgue, et de densité*

$$\rho_p(\theta_1,\dots,\theta_p) = \frac{(N-p)!}{N!}\det\left(K^{(N)}(\theta_k,\theta_l)\right)_{k,l=1}^p\,,$$

où

$$K^{(N)}(x,y) = \sum_{k=0}^{N-1}\psi_k(x)\psi_k(y) \tag{5}$$

avec ψ_k la fonction d'Hermite

$$\psi_k(x) := \frac{e^{-\frac{x^2}{4}}P_k(x)}{(2\pi)^{\frac14}\sqrt{k!}} \qquad si \qquad P_k(x) := (-1)^k e^{x^2/2}\frac{d^k}{dx^k}e^{-x^2/2}\,.$$

ρ_p est appelée la fonction de corrélations à p points, et $\mathcal{P}_N$ est dite déterminantale comme ses fonctions de corrélations sont données par le déterminant d'un noyau.

Preuve. — Le point principal est de remarquer que la densité ρ_N dépend du déterminant de Vandermonde qui, comme les polynômes P_k sont moniques, s'écrit

$$\Delta_N(x) := \prod_{1\le i<j\le N}(x_j-x_i) = \det(x_i^{j-1})_{i,j=1}^N = \det\left(P_{j-1}(x_i)\right)_{i,j=1}^N\,. \tag{6}$$

Dans le cas où $p=N$, nous en déduisons immédiatement que

$$\begin{aligned}\rho_N(\theta_1,\dots,\theta_N) &= C_N\left(\det\left(P_{j-1}(\theta_i)\right)_{i,j=1}^N\right)^2\prod_{i=1}^N e^{-\theta_i^2/2} \\ &= \tilde{C}_N\left(\det\left((\psi_{j-1}(\theta_i))_{i,j=1}^N\right)\right)^2 = \tilde{C}_N\det\left(\left(K^{(N)}(\theta_i,\theta_j)\right)_{i,j=1}^N\right)\,,\end{aligned} \tag{7}$$

où nous avons finalement utilisé la formule $\det(AB) = \det(A)\det(B)$. Pour $p < N$, en utilisant (6), on a, avec $x_i = \theta_i$ si $i \leq p$ et ζ_i sinon, et en notant $S_{\nu_1,\ldots,\nu_p}$ l'ensemble des bijections de $\{1,\ldots,p\}$ dans $\{\nu_1,\ldots,\nu_p\}$,

$$\begin{aligned}
\rho_p(\theta_1,\ldots,\theta_p) &= \tilde{C}_p \int (\det((\psi_{j-1}(x_i))_{i,j=1}^N))^2 \prod_{i=p+1}^N d\zeta_i \\
&= \tilde{C}_p \sum_{\sigma,\tau\in S_{1,\ldots,N}} \varepsilon(\sigma)\varepsilon(\tau) \int \prod_{j=1}^N \psi_{\sigma(j)-1}(x_j)\psi_{\tau(j)-1}(x_j) \prod_{i=p+1}^N d\zeta_i \\
&= \tilde{C}_p \sum_{1\leq\nu_1<\cdots<\nu_p\leq N} \sum_{\sigma,\tau\in S_{\nu_1,\ldots,\nu_p}} \varepsilon(\sigma)\varepsilon(\tau) \prod_{i=1}^p \psi_{\sigma(i)-1}(\theta_i)\psi_{\tau(i)-1}(\theta_i) \\
&= \tilde{C}_p \sum_{1\leq\nu_1<\cdots<\nu_p\leq N} \left(\det\left((\psi_{\nu_j-1}(\theta_i))_{i,j=1}^p\right)\right)^2,
\end{aligned}$$

où nous avons simplement utilisé que $\int \psi_k(x)\psi_\ell(x)dx = 1_{k=\ell}$. La formule de Cauchy-Binet permet de conclure. Le calcul des constantes $\tilde{C}_p$ est aisé et laissé au lecteur. □

La forme particulière des fonctions de corrélations permet d'écrire la probabilité que toutes les valeurs propres sont dans un ensemble donné comme un déterminant de Fredholm.

LEMME 1.2. — *Pour tout ensemble mesurable A de $\mathbb{R}$,*

$$\mathscr{P}_N(\cap_{i=1}^N\{\lambda_i \in A\}) = 1 + \sum_{k=1}^\infty \frac{(-1)^k}{k!} \int_{A^c} \cdots \int_{A^c} \det\left((K^{(N)}(x_i,x_j))_{i,j=1}^k\right) \prod_{i=1}^k dx_i. \tag{8}$$

Preuve. — En utilisant le lemme précédent, l'orthogonalité des fonctions d'Hermite et de nouveau le théorème de Cauchy-Binet, on a

$$\begin{aligned}
\mathscr{P}_N(\cap_{i=1}^N\{\lambda_i \in A\}) &= \frac{1}{N!}\int_A \cdots \int_A \left(\det\left((\psi_i(x_j))_{i,j=0}^{N-1}\right)\right)^2 \prod dx_i \\
&= \det\left(\left(\int_A \psi_i(x)\psi_j(x)dx\right)_{i,j=0}^{N-1}\right) \\
&= \det\left(\left(\delta_{ij} - \int_{A^c} \psi_i(x)\psi_j(x)dx\right)_{i,j=0}^{N-1}\right) \\
&= 1 + \sum_{k=1}^N (-1)^k \sum_{0\leq\nu_1<\cdots<\nu_k\leq N-1} \det\left(\left(\int_{A^c} \psi_{\nu_i}(x)\psi_{\nu_j}(x)dx\right)_{i,j=1}^k\right) \\
&= 1 + \sum_{k=1}^\infty \frac{(-1)^k}{k!}\int_{A^c}\cdots\int_{A^c} \det(K^{(N)}(x_i,x_j))_{i,j=1}^k \prod_{i=1}^k dx_i\,. \qquad (9)
\end{aligned}$$

□

Cette formule peut être utilisée pour obtenir des asymptotiques. En effet, le déterminant de Fredholm

$$\Delta(A,K) := 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_A \cdots \int_A \det\left((K(x_i,x_j))_{i,j=1}^k\right) \prod_{i=1}^k dx_i$$

est continu sur l'ensemble des noyaux K muni de la topologie uniforme. Prenant $A^c = N^{\frac{1}{2}}x + N^{-\frac{1}{2}}B$ pour un ensemble compact B et $x \in (-2,+2)$, on voit que la première partie du théorème 0.1 est une conséquence de la convergence (uniforme sur les compacts)

(10)

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} K^{(N)}\left(N^{\frac{1}{2}}x + \frac{y_1}{\rho_{sc}(x)\sqrt{N}}, N^{\frac{1}{2}}x + \frac{y_2}{\rho_{sc}(x)\sqrt{N}}\right) = \frac{\sin(y_1-y_2)}{\pi(y_1-y_2)} =: S(y_1,y_2)$$

qui implique que

$$\lim_{N\to\infty} \mathscr{P}_N\left(\lambda_i \notin N^{\frac{1}{2}}x + N^{-\frac{1}{2}}\rho_{sc}(x)B, i=1,\ldots,N\right) = \Delta(B,S)\,. \tag{11}$$

De même, en prenant $A = 2N^{\frac{1}{2}} + N^{-\frac{1}{6}}[t,t']$, on démontre que

$$\lim_{N\to\infty} \mathscr{P}_N\left(\lambda_i \notin 2N^{\frac{1}{2}} + N^{-\frac{1}{6}}[t,t'], i=1,\ldots,N\right) = \Delta([t,t'],\mathscr{A}) \tag{12}$$

car

$$\lim_{N\to\infty} \frac{1}{N^{\frac{1}{6}}} K^{(N)}(2\sqrt{N} + \frac{x}{N^{\frac{1}{6}}}, 2\sqrt{N} + \frac{y}{N^{\frac{1}{6}}}) = \frac{Ai(x)Ai'(y) - Ai'(x)Ai(y)}{x-y} =: \mathscr{A}(x,y)$$

où Ai est la fonction de Airy. Ces asymptotiques sont démontrées grâce à la formule

$$K^{(N)}(x,y) = \sqrt{N}\frac{\psi_N(x)\psi_{N-1}(y) - \psi_{N-1}(x)\psi_N(y)}{x-y} \tag{13}$$

et aux asymptotiques de $\psi_N(x_N)$ et $\psi_{N-1}(x_N)$ en $x_N = N^{\frac{1}{2}}x + N^{-\frac{1}{2}}y$ et $x_N = 2N^{\frac{1}{2}} + N^{-\frac{1}{6}}x$. De telles asymptotiques peuvent se démontrer en écrivant la fonction d'Hermite ψ_k comme une intégrale oscillante

$$(-1)^k\sqrt{\sqrt{2\pi}k!}e^{-\frac{x^2}{4}}\psi_k(x) = \frac{d^k}{dx^k}e^{-\frac{x^2}{2}} = \frac{d^k}{dx^k}\int e^{ixy}d\mu(y) = \int (iy)^k e^{ixy}d\mu(y)$$

et en utilisant une méthode du col. On notera à ce propos que le noyau sinus apparaît quand le point de selle est du second ordre, alors qu'il est du troisième ordre pour le noyau d'Airy, voir [**1**, chapitre 3] pour plus de détails.

L'étude des statistiques locales des matrices gaussiennes a été étendue au cas de matrices symétriques et symplectiques pour lesquelles la loi jointe des valeurs propres est également explicite (41) ; elle est néanmoins plus compliquée, le déterminant de Vandermonde n'apparaissant pas au carré et les fonctions de corrélations n'étant plus déterminantales, et fait appel à la théorie des pfaffiens [**1**, section 3.9].

2. UNIVERSALITÉ POUR UN CERTAIN ENSEMBLE DE MATRICES DE WIGNER

Au cours de la section précédente, l'étude précise du spectre des matrices aléatoires gaussiennes hermitiennes a pu être effectuée grâce à la structure déterminantale de la loi jointe des valeurs propres de ces matrices. Suite aux travaux de Brézin-Hikami [**4, 5**], K. Johansson [**21**] fit une remarque fondamentale; la loi d'une matrice de Wigner dont les coefficients ont une petite composante gaussienne a également une structure déterminantale. Plus précisément, soient $\mathbf{G}_N$ une matrice du GUE et $\mathbf{V}_N$ une matrice de Wigner de loi P_N, indépendante de $\mathbf{G}_N$. Pour tout $\varepsilon \in (0,1)$, nous considérons les matrices de Wigner données par

$$\mathbf{X}_N = \sqrt{1-\varepsilon}\mathbf{V}_N + \sqrt{\varepsilon}\mathbf{G}_N\,, \tag{14}$$

c'est-à-dire dont les lois des coefficients ν et μ sont obtenues par convolution par les distributions gaussiennes (3) ; on dit que ces lois sont divisibles par une gaussienne. Nous allons voir que la loi jointe $\mathscr{P}_N^\varepsilon$ des valeurs propres de $\mathbf{X}_N$ est une moyenne de lois déterminantales. En effet, comme on peut décomposer $\mathbf{G}_N$ comme $U\Lambda U^*$ avec U qui suit la mesure de Haar sur le groupe unitaire et Λ une matrice diagonale, indépendante de U, dont les coefficients sur la diagonale ont pour loi (4), on a

$$\begin{aligned}\int_{\mathbb{R}^N} f(\lambda)\mathscr{P}_N^\varepsilon(d\lambda) &= \frac{1}{Z_N}\int f(\text{valeurs propres } \mathbf{X}_N)e^{-\frac{1}{2\varepsilon}\mathrm{Tr}(\mathbf{X}_N-\sqrt{1-\varepsilon}V)^2}dP_N(V)d\mathbf{X}_N \\ &= \frac{1}{Z'_N}\int f(\Lambda)\int_{U(N)} e^{-\frac{1}{2\varepsilon}\mathrm{Tr}(U\Lambda U^*-\sqrt{1-\varepsilon}V)^2}dU dP_N(V)d\mathscr{P}_N(\Lambda),\end{aligned}$$

où dU est la mesure de Haar sur le groupe unitaire. On reconnaît l'intégrale de Harish-Chandra–Itzykson-Zuber qui est elle aussi donnée par un déterminant [**19, 20**]

$$\int e^{\frac{1}{\varepsilon}\mathrm{Tr}(U\Lambda U^* V)}dU = \Delta_N(\rho)\times\frac{\det(e^{\frac{1}{\varepsilon}\Lambda_{ii}\lambda_j(V)})_{1\le i,j\le N}}{\Delta_N(\Lambda)\Delta_N(\lambda(V))}$$

avec $\rho_i = i-1$, $1\le i\le N$, et $(\lambda_k(V))_{1\le k\le N}$ les valeurs propres de V. Ceci permet d'écrire la formule

$$\int_{\mathbb{R}^N} f(\lambda)\mathscr{P}_N^\varepsilon(d\lambda) = \frac{1}{Z'_N}\int f(\eta)\frac{\Delta_N(\eta)}{\Delta_N(\lambda(V))}\det((e^{-\frac{1}{2\varepsilon}(\eta_j-\sqrt{1-\varepsilon}\lambda_k(V))^2})_{j,k=1}^N)d\eta dP_N(V).$$

Cette structure particulière permet de montrer que les fonctions de corrélations sont des moyennes, sous la loi P_N, de déterminants. En effet, un peu d'algèbre montre que pour tout $\lambda = (\lambda_1,\ldots,\lambda_N)\in\mathbb{R}^N$, la mesure de probabilité absolument continue par rapport à la mesure de Lebesgue et de densité donnée par

$$q_S(x:\lambda) = \frac{1}{Z'_N}\frac{\Delta_N(x)}{\Delta_N(\lambda)}\det((e^{-(x_j-\lambda_k)^2})_{j,k=1}^N)$$

est déterminantale, i.e. que ses fonctions de corrélations à p points s'écrivent comme le déterminant d'un noyau K_N^λ qui dépend de λ. Ce résultat généralise le lemme 1.1. K. Johansson a montré [**21**, Lemma 3.2] que l'ensemble des λ dont la mesure empirique $N^{-1}\sum_{i=1}^N \delta_{\lambda_i}$ approxime suffisamment bien la loi semi-circulaire (notamment tel que la convergence dans (1) se fasse polynomialement en N) est tel que K_N^λ, correctement rééchelonné, converge vers le noyau sinus, i.e satisfait (10). On peut alors étendre (11) en montrant que les valeurs propres de $\mathbf{V}_N$ satisfont cette condition presque sûrement, ce qui est vérifié dès lors que les coefficients de $\mathbf{V}_N$ ont plus de six moments finis.

Ainsi, la preuve que les espacements des valeurs propres à l'intérieur du spectre sont asymptotiquement décrits par le noyau sinus repose à nouveau sur une structure déterminantale, mais qui demande que les coefficients des matrices aient une petite composante gaussienne complexe et des moments d'ordre six finis. Ce résultat vient d'être étendu par K. Johansson [**22**] qui a montré que de telles matrices ont des fluctuations universelles à l'intérieur du spectre dès lors que le second moment des coefficients de $\mathbf{V}_N$ est fini alors que les fluctuations au bord du spectre sont universelles sous la condition optimale [**2**] des quatre premiers moments finis. A. Soshnikov [**29**] a abordé la question de l'universalité au bord du spectre en utilisant des estimations de moments, fondées sur des techniques de combinatoire. Il a ainsi pu contourner l'hypothèse de divisibilité par une gaussienne.

3. UNIVERSALITÉ AU BORD DU SPECTRE

Une méthode assez naturelle pour étudier les matrices aléatoires, et par exemple démontrer (1), est d'estimer les moments de la mesure empirique des valeurs propres

$$m_N(k) := \mathbb{E}[\frac{1}{N}\sum_{i=1}^N (\frac{\lambda_i}{\sqrt{N}})^k] = \mathbb{E}[\frac{1}{N}\mathrm{Tr}(\mathbf{Y}_N^k)] \quad \text{où} \quad \mathrm{Tr}(A) = \sum_{i=1}^N A_{ii}\,.$$

En effet, le terme de droite s'écrit en fonction des moments des coefficients de la matrice, et est donc à priori calculable. Démontrons (1) afin d'avoir une idée de ces techniques.

THÉORÈME 3.1. — *Supposons que ν et μ ont tous leurs moments finis. Alors, $m_N(k)$ converge quand N tend vers l'infini vers $\int x^k d\sigma(x)$ pour tout $k \in \mathbb{N}$.*

Ce théorème peut être généralisé pour démontrer la convergence presque sûre des moments, simplement en montrant que leurs covariances sont suffisamment petites pour que le lemme de Borel-Cantelli permette de montrer l'existence d'un espace de probabilités où leurs convergences ont lieu presque sûrement.

Preuve. — La preuve de ce théorème repose sur le fait que $\int x^k d\sigma(x)$ est nul si k est impair et égal au nombre de Catalan sinon, c'est-à-dire au nombre d'arbres enracinés orientés avec $k/2$ branches. Ce nombre va naturellement apparaître dans la combinatoire des moments

$$m_N(k) = \mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\left(\mathbf{Y}_N^k\right)\right] = \sum_{i_1,\ldots,i_k=1}^{N} N^{-\frac{k}{2}-1}\mathbb{E}[X_{i_1i_2}X_{i_2i_3}\cdots X_{i_ki_1}].$$

Nous allons associer aux indices $\mathbf{i} = \{i_1,\ldots,i_k\}$ le graphe $G(\mathbf{i})$ de sommets $V(\mathbf{i}) = \{i_1,\ldots,i_k\}$ et d'arêtes $E(\mathbf{i}) = \{(i_1,i_2),\ldots,(i_{k-1},i_k),(i_k,i_1)\}$. Par définition, $G(\mathbf{i})$ est connexe. Par ailleurs, par indépendance et centrage, $\mathbb{E}[X_{i_1i_2}X_{i_2i_3}\cdots X_{i_ki_1}] = 0$ si $E(\mathbf{i})$ possède une arête non orientée de multiplicité un. Par conséquent, le nombre $|E(\mathbf{i})|$ d'arêtes non orientées différentes du graphe $G(\mathbf{i})$ contribuant à la somme (15) est au plus $2^{-1}k$, et de fait au plus la partie entière $[2^{-1}k]$ de ce nombre. Or, comme $G(\mathbf{i})$ est connexe, l'inégalité

$$|V(\mathbf{i})| \leq |E(\mathbf{i})| + 1 \tag{15}$$

implique que $|V(\mathbf{i})| \leq [\frac{k}{2}] + 1$, et donc il y a au plus $N^{[\frac{k}{2}]+1}$ indices contribuant à la somme (15), qui est alors au plus d'ordre $N^{-\frac{1}{2}}$ si k est impair. Si k est pair, le même raisonnement montre que les indices contribuant à la somme (15) sont tels que $|V(\mathbf{i})| = |E(\mathbf{i})| + 1 = 2^{-1}k + 1$. Mais l'égalité dans (15) n'est possible que si $G(\mathbf{i})$ est un arbre. Par ailleurs, comme $|E(\mathbf{i})| = 2^{-1}k$, chaque arête non orientée apparaît exactement deux fois, et en fait une fois dans chaque orientation. Dans ce cas, $\mathbb{E}[X_{i_1i_2}X_{i_2i_3}\cdots X_{i_ki_1}]$ vaut un comme c'est un produit de covariances. Enfin, comme sont associés à cet arbre $G(\mathbf{i})$ une racine (par exemple (i_1,i_2)) et un ordre d'exploration (décrit par le chemin $(i_1 \to i_2 \cdots \to i_k \to i_1)$), on conclut que le premier ordre du terme de droite de (15) est donné exactement par le nombre d'arbres enracinés orientés avec $k/2$ branches, c'est-à-dire le nombre de Catalan. □

Pour étudier le bord du spectre, l'idée est simplement d'arriver à estimer des moments d'ordre k tendant à l'infini suffisamment rapidement. En effet, clairement, la plus grande valeur propre est plus grande que $N^{\frac{1}{2}}(2-\varepsilon)$ pour tout $\varepsilon > 0$ car sinon aucune valeur propre ne pourrait être dans $[N^{\frac{1}{2}}(2-\varepsilon), N^{\frac{1}{2}}]$, ce qui démentirait la convergence de la mesure empirique vers la loi semi-circulaire. Par ailleurs, l'inégalité de Tchebychev implique que

$$P\left(\max_{1\leq i\leq N}\lambda_i \geq \sqrt{N}(2+\varepsilon)\right) \leq \mathbb{E}\left[\left(\frac{\max_{1\leq i\leq N}\lambda_i}{\sqrt{N}(2+\varepsilon)}\right)^{2k}\right] \leq \mathbb{E}\left[\mathrm{Tr}\left(\frac{\mathbf{Y}_N}{(2+\varepsilon)}\right)^{2k}\right] \tag{16}$$

et il suffit donc de montrer que, pour $k \gg \log N$,

$$\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{Y}_N\right)^{2k}\right] = \sqrt{\frac{8}{\pi}}\frac{2^{2k}N}{(2k)^{\frac{3}{2}}}(1+o(1)) \tag{17}$$

pour déduire que le terme de droite de (16) tend vers zéro suffisamment vite pour que les valeurs propres de $\mathbf{Y}_N$ soient plus petites que $(2+\varepsilon)$ pour N suffisamment grand presque sûrement. Bien sûr, ce dernier résultat ne nécessite pas une estimation aussi précise que (17), au contraire de l'étude des fluctuations exactes de la plus grande valeur propre qui fait appel à une estimation de ce type pour k de l'ordre de $N^{\frac{2}{3}}$ comme nous allons le voir. L'estimation (17) a été effectuée dans **[28]** pour $k = o(N^{\frac{1}{2}})$ et dans le cas où les mesures μ et ν sont symétriques et ont des queues sous-gaussiennes, si bien que leurs moments ne croissent pas trop vite. La preuve de ce résultat consiste à montrer que, comme dans la preuve du théorème 3.1, les indices contribuant au premier ordre du développement du moment d'ordre k sont représentés par des arbres enracinés orientés, si bien que ce moment est équivalent au moment d'ordre k de la loi semi-circulaire. En développant ce type d'idées, on voit que les fluctuations de la plus grande valeur propre sont décrites par les moments d'ordre $k = tN^{\frac{2}{3}}$ dans le sens où

$$\mathbb{E}\left[\operatorname{Tr}\left(2^{-1}\mathbf{Y}_N\right)^{2[tN^{\frac{2}{3}}]}\right] = \mathbb{E}\left[\sum_{i=1}^{N} e^{tN^{\frac{1}{6}}(\lambda_i - 2\sqrt{N})}\right](1+o(1)).$$

Par contre, l'estimation de ce moment ne donne pas directement la transformée de Laplace de $\theta := (\max_i \lambda_i - 2\sqrt{N})N^{\frac{1}{6}}$ mais plutôt la somme des transformées de Laplace des valeurs propres proches de deux. Afin d'obtenir la transformée de Laplace de θ seulement, l'idée est d'estimer le moment plus général

$$\mathbb{E}\left[\prod_{i=1}^{m}\operatorname{Tr}\left(2^{-1}\mathbf{Y}_N\right)^{2[t_i N^{\frac{2}{3}}]}\right]$$

pour tout m et des paramètres t_i arbitraires, ce qui décrit grosso modo la loi jointe des plus grandes valeurs propres, recentrées par leur limite et rééchelonnées. Le résultat central de [**29**, Corollary 5] est d'avoir montré que ces moments sont approximativement les mêmes que dans le cas gaussien quand les lois ν et μ sont symétriques et de moments sous-gaussiens, en affinant les idées de **[28]**. Cette approche nécessite une maîtrise très fine de la combinatoire des moments, qui n'a pu encore être généralisée **[26]** au cas où les lois μ et ν ne sont pas symétriques, cette hypothèse annulant un grand nombre de moments. L'hypothèse sur les moments a pu être affaiblie en une hypothèse de trente-sixième (resp. douzième) plus epsilon moment fini en procédant par approximation **[27]** (resp. **[23]**). L'universalité des fluctuations sous la condition que les quatre premiers moments de la distribution des coefficients sont égaux à ceux de la distribution gaussienne, mais sans condition de symétrie des lois, vient d'être obtenue **[30]** par des méthodes qui généralisent celles développées dans la section 4.2.

4. UNIVERSALITÉ À L'INTÉRIEUR DU SPECTRE ; LE THÉORÈME DES QUATRE MOMENTS

Afin de montrer l'universalité des espacements sous une condition générale sur les moments, T. Tao et V. Vu **[32]** ont eu l'idée simple mais lumineuse de comparer les statistiques locales de deux matrices de Wigner en remplaçant une à une les entrées de la matrice et en montrant que ce changement n'affecte guère les statistiques locales, suivant ainsi la stratégie de remplacement que Lindeberg introduisit pour démontrer l'universalité du théorème central limite. Pour mener à bien ce programme, une étape essentielle est de prouver que les vecteurs propres de la matrice sont « délocalisés », phénomène lié à une étude fine de la convergence vers la loi semi-circulaire (1). Ce point a été établi dans [**12**, Theorem 3.1]. Nous développons maintenant ces arguments. Cette section s'applique aussi bien dans le cas de matrices symétriques (μ mesure sur la droite réelle) que hermitiennes, au contraire de la section 2.

4.1. Loi semi-circulaire locale et phénomène de délocalisation

Cette étude remonte à [**12**, Theorem 3.1]. La loi semi-circulaire locale établit que le nombre de valeurs propres de $\mathbf{Y}_N$ tombant dans un intervalle centré en $x \in (-2,2)$ de largeur ℓ supérieure à $N^{-1}(\log N)^4$, mais tendant vers zéro avec N, est proche de $N\rho_{sc}(x)\ell$ avec grande probabilité. Plus précisément, nous avons

THÉORÈME 4.1. — *Supposons que μ et ν ont des queues sous-gaussiennes. Soit $\mathcal{N}_I$ le nombre de valeurs propres de $\mathbf{Y}_N$ dans $I \subset]-2,2[$. Alors, il existe une constante $c \in (0,\infty)$ telle que, pour tout $\kappa \in (0,2)$, tout $\delta \leq c\kappa$, tout $\eta \in [\frac{(\log N)^4}{N}, \kappa/2]$, on a*

$$P\left(\sup_{|E|\leq 2-\kappa}\left|\frac{\mathcal{N}_{[E-\eta,E+\eta]}}{N} - \sigma([E-\eta,E+\eta])\right| > 2\eta\delta\right) \leq Ce^{-c\delta^2\sqrt{N\eta}}. \tag{18}$$

La preuve de ce théorème repose sur l'étude de la transformée de Cauchy-Stieltjes qui est définie comme suit.

DÉFINITION 4.2. — *Pour une mesure de probabilité P sur $\mathbb{R}$, on note pour $z \in \mathbb{C}\backslash\mathbb{R}$*

$$G_P(z) = \int \frac{1}{z-x} dP(x)$$

la transformée de Cauchy-Stieltjes de P.
En particulier, nous noterons $G^N(z) = N^{-1}\mathrm{Tr}\left((z-\mathbf{Y}_N)^{-1}\right)$ et $G(z) = G_\sigma(z)$ les transformées de Cauchy-Stieltjes de la mesure spectrale de $\mathbf{Y}_N$ et de la loi semi-circulaire.

La transformée de Cauchy-Stieltjes joue un rôle central dans l'étude des matrices aléatoires car c'est une fonction génératrice des moments qui satisfait souvent des équations algébriques, liées aux propriétés combinatoires des moments que nous avons utilisées dans la section 3. La comparaison des équations satisfaites par G^N et G permet ainsi souvent de comparer ces deux fonctions comme nous allons le voir. Nous allons démontrer le résultat suivant, qui permet de démontrer le théorème (4.1), voir (29).

PROPOSITION 4.3. — *Sous les hypothèses du théorème précédent, il existe des constantes $c, C > 0$ finies telles que, pour tout $\kappa \in (0,2)$, tout $\eta \in [\frac{(\log N)^4}{N}, \kappa/2]$, tout $\delta \leq C\kappa$, tout $N \geq 2$,*

$$(19) \qquad P\left(\sup_{E\in[-2+\kappa,2-\kappa]} \left|G^N(E+i\eta) - G(E+i\eta)\right| > \delta\right) \leq Ce^{-c\delta\sqrt{N\eta}}.$$

Ces résultats ont été ultérieurement étendus à des queues sous-exponentielles. Ils entraînent le phénomène de délocalisation des vecteurs propres suivant.

COROLLAIRE 4.4. — *Sous les mêmes hypothèses, pour tous $\kappa > 0$ et $K > 0$, il existe des constantes C, c telles que pour tout $N \geq 2$,*

$$P\left(\exists v \text{ tel que } \mathbf{Y}_N v = \mu v,\ \|v\|_2 = 1,\ \mu \in [-2+\kappa, 2-\kappa] \text{ et } \|v\|_\infty \geq \frac{(\log N)^{\frac{9}{2}}}{N^{\frac{1}{2}}}\right) \leq Ce^{-c(\log N)^2}.$$

Preuve de la proposition 4.3. — Cette preuve utilise la formule d'algèbre linéaire

$$(20) \qquad (z - \mathbf{Y}_N)^{-1}_{kk} = \left(z - N^{-\frac{1}{2}}X_{kk} - N^{-1}\langle X_k, (z - \mathbf{Y}^{(k)})^{-1}X_k\rangle\right)^{-1}$$

où $\mathbf{Y}^{(k)}$ est la matrice carrée de taille $N-1$ obtenue en supprimant la ligne et la colonne k de la matrice $\mathbf{Y}_N$, et X_k est le vecteur $(X_{k1}, \ldots, X_{kk-1}, X_{kk+1}, \ldots, X_{kN})$. Une preuve classique de la convergence vers la loi semi-circulaire utilise l'indépendance de X_k et $\mathbf{Y}^{(k)}$ qui permet d'établir, par la loi des grands nombres, que $N^{-1}\langle X_k, (z - \mathbf{Y}^{(k)})^{-1}X_k\rangle$ est proche de sa moyenne par rapport à X_k, qui est elle-même égale à $G^{N,k}(z) = \frac{1}{N}\mathrm{Tr}((z-\mathbf{Y}^{(k)})^{-1})$. Par la propriété d'entrelacement des valeurs propres de Weyl, chaque valeur propre de $\mathbf{Y}^{(k)}$ est située entre deux valeurs propres de $\mathbf{Y}_N$, et par conséquent $G^{N,k}(z)$ est proche de $G^N(z)$, ce qui permet avec (20) et en sommant sur k, de déduire que

$$(21) \qquad G^N(z) = \frac{1}{N}\sum_{k=1}^{N} \frac{1}{z - G^N(z) - \varepsilon_k^N(z)},$$

avec $\varepsilon_k^N(z)$ tendant vers zéro avec N pour tout z hors de l'axe réel. Asymptotiquement, on déduit que les points limites de $G^N(z)$ satisfont l'équation

$$G(z) = (z - G(z))^{-1}$$

qui a une unique solution tendant vers zéro quand la partie imaginaire de z tend vers l'infini. Pour obtenir le théorème 4.1, il faut quantifier précisément les erreurs effectuées. Classiquement, ces erreurs étaient estimées en bornant l'opérateur $(z - \mathbf{Y}^{(k)})^{-1}$ par l'inverse de la partie imaginaire de z. La remarque fondamentale de **[12]** a été de remarquer qu'on pouvait minorer beaucoup plus finement la partie imaginaire de $z - \mathbf{Y}^{(k)}$, pourvu qu'on puisse estimer a priori le nombre de valeurs propres autour de la partie réelle de z. Pour obtenir cette estimation, on remarque tout d'abord que, pour toute mesure de probabilité P, tout $\eta \in (0, 2)$,

$$\begin{aligned} \Im \int (E + i\eta - x)^{-1} dP(x) &= \int \frac{-\eta}{\eta^2 + (E-x)^2} dP(x) \\ &\leq \int_{E-\eta}^{E+\eta} \frac{-\eta}{\eta^2 + (E-x)^2} dP(x) \leq \frac{-1}{2\eta} P([E-\eta, E+\eta]) \end{aligned} \tag{22}$$

si bien qu' en choisissant P la mesure spectrale de $\mathbf{Y}_N$, il suffit de minorer $\Im G^N(E+i\eta)$ par $-M$ pour montrer qu'il y a au plus $2N\eta M$ valeurs propres dans $[E-\eta, E+\eta]$. Mais, en utilisant (20), on trouve, comme la partie imaginaire de $\langle X_k, (z-\mathbf{Y}^{(k)})^{-1} X_k\rangle$ a le signe opposé de celui de η, que, avec $z = E + i\eta$,

$$\left|\Im G^N(z)\right| \leq (-\Im N^{-1} \langle X_k, (z-\mathbf{Y}^{(k)})^{-1} X_k\rangle)^{-1} = (-N^{-1} \sum_{i=1}^N \Im (z - \lambda_i^k)^{-1} |\langle X_k, u_i^k\rangle|^2)^{-1}$$

où u_i^k est le vecteur propre de $\mathbf{Y}^{(k)}$ pour la valeur propre λ_i^k. En utilisant la minoration $-\Im(z - \lambda_i^k)^{-1} \geq (2\eta)^{-1}$ si $|\lambda_i^k - E| \leq \eta$, on déduit que

$$\left|\Im G^N(E+i\eta)\right| \leq 2\eta (N^{-1} \sum_{i=1}^N 1_{|\lambda_i^k - E| \leq \eta} |\langle X_k, u_i^k\rangle|^2)^{-1}. \tag{23}$$

Par la propriété d'entrelacement de Weyl, le nombre de valeurs propres de $\mathbf{Y}^{(k)}$ dans $[E-\eta, E+\eta]$ est au moins $\mathcal{N}_{[E-\eta, E+\eta]} - 1$. Si ce nombre tend vers l'infini avec N, la loi des grands nombres implique que $(\mathcal{N}_{[E-\eta,E+\eta]})^{-1} \sum_{i=1}^N 1_{|\lambda_i^k - E| \leq \eta} |\langle X_k, u_i^k\rangle|^2$ converge presque sûrement vers un par l'indépendance de X_k avec (u_i, λ_i). Nous pouvons quantifier cette convergence par le lemme de Hanson-Wright **[18]** (voir aussi **[14**, Proposition 4.5] pour les hypothèses actuelles) qui implique que, si v est un vecteur aléatoire de coefficients indépendants et équidistribués selon la loi ν, et A une matrice déterministe,

$$P\left(|\langle v, Av\rangle - \mathbb{E}[\langle v, Av\rangle]| \geq \delta\right) \leq 4e^{-\min\{\delta C, \delta^2 C^2\}} \tag{24}$$

avec $C^{-2} = \mathrm{Tr}(AA^*)$. Conditionnellement à $A := \sum_{i=1}^N 1_{|\lambda_i^k - E| \leq \eta} u_i^k (u_i^k)^*$, et donc $C^{-2} := \sum_{i=1}^N 1_{|\lambda_i^k - E| \leq \eta} = \mathbb{E}[\sum_{i=1}^N 1_{|\lambda_i^k - E| \leq \eta} |\langle X_k, u_i^k\rangle|^2]$, nous déduisons que pour tout $\delta > 0$

$$\sum_{i=1}^{N} 1_{|\lambda_i^k - E| \leq \eta} |\langle X_k, u_i^k \rangle|^2 \geq C^{-2} - \delta C^{-1}$$

avec probabilité supérieure à $1 - e^{-\min\{\delta,\delta^2\}}$. Si $C^{-2} \leq 2\eta N$, alors

$$\mathcal{N}_{[E-\eta,E+\eta]} - 1 \leq 2\eta N \tag{25}$$

et nous avons l'estimation désirée. Sinon, en prenant $\delta = 2^{-1}C^{-1} \geq (\eta N/2)^{\frac{1}{2}}$, on obtient

$$\sum_{i=1}^{N} 1_{|\lambda_i^k - E| \leq \eta} |\langle X_k, u_i^k \rangle|^2 \geq 2^{-1}(\mathcal{N}_{[E-\eta,E+\eta]} - 1)$$

avec probabilité supérieure ou égale à $1 - e^{-2^{-1}\sqrt{N\eta}}$. Dans ce cas, (22) et (23) impliquent qu'avec cette grande probabilité

$$(\mathcal{N}_{[E-\eta,E+\eta]} - 1)^2 \leq 4(N\eta)^2 ,$$

ce qui permet d'obtenir de nouveau (25), avec probabilité supérieure ou égale à $1 - e^{-2^{-1}\sqrt{N\eta}}$. Nous avons donc dans tous les cas cette estimation. Pour obtenir le résultat (19), nous estimons les erreurs ε_k^N dans (21) qui valent

$$\frac{1}{N}\sum_{i=1}^{N}(z-\lambda_i^k)^{-1}[|\langle X_k, u_i^k\rangle|^2 - 1] + \frac{1}{N}\mathrm{Tr}((z-\mathbf{Y}^{(k)})^{-1} - (z-\mathbf{Y}_N)^{-1}) + \frac{X_{kk}}{\sqrt{N}} =: \varepsilon_k^{N,1} + \varepsilon_k^{N,2} + \varepsilon_k^{N,3} .$$

Par la propriété d'entrelacement de Weyl, le second terme $\varepsilon_k^{N,2}$ est au plus d'ordre $1/N\eta$ alors que $\varepsilon_k^{N,3}$ est inférieur à $\delta > 0$ avec probabilité supérieure à $1 - e^{-c\delta^2 N}$. Pour $\varepsilon_k^{N,1}$, nous appliquons le lemme de Hanson-Wright à $v(i) = X_k(i)$ et conditionnellement à $A = N^{-1}\sum_{j=1}^N (z-\lambda_j^k)^{-1} u_j^k (u_j^k)^*$ on obtient, avec $I_n = [E - 2^n\eta, E + 2^n\eta]$, et comme $|\lambda_i^k - E|^2 + \eta^2 \geq (2^n+1)\eta^2$ si $\lambda_i^k \in I_n$,

$$C^{-2} = \mathrm{Tr}(AA^*) = \frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{|\lambda_i^k - E|^2 + \eta^2} \leq \frac{1}{N^2}\sum_{n\geq 0}\frac{\mathcal{N}_{I_n}}{\eta^2(2^n+1)} \leq \frac{C}{N\eta}$$

par (25). Par conséquent, (24) implique que, pour $\delta > 1/\sqrt{N\eta}$

$$P\left(\max_{1\leq k\leq N} |\varepsilon_k^N| \geq \delta\right) \leq 4Ne^{-c\delta\sqrt{N\eta}} . \tag{26}$$

Pour en déduire une estimation sur $G^N - G$, remarquons déjà que (21) implique

$$G^N(z) - \frac{1}{z - G^N(z)} = \eta_N(z) := \frac{1}{N}\sum_{k=1}^{N}\frac{\varepsilon_k^N(z)}{(z - G^N(z))(z - G^N(z) + \varepsilon_k^N(z))} . \tag{27}$$

À $\eta_N(z)$ donné, il y a une unique solution à cette équation telle que $\Im G^N(z) < 0$ pour $\Im z > 0$, et elle est donnée par

$$G^N(z) = \frac{1}{2}(z - \sqrt{z^2 - 4 - 2z\eta_N(z) + \eta_N(z)^2}) - \frac{\eta_N(z)}{2}.$$

Pour z de partie réelle dans $]-2+\kappa, 2-\kappa[$, $|z^2 - 4| \geq C\kappa$, si bien que la racine est Lipschitz à ce point et que

$$|G^N(z) - G(z)| \leq C(\kappa)\eta_N(z). \tag{28}$$

Pour conclure, il nous faut donc montrer que $\eta_N(z)$ est bien petit avec grande probabilité, fait pour lequel il nous suffit de montrer que $z - G^N(z)$ n'est pas trop petit d'après (27) et (26). Comme les parties imaginaires de z et $G^N(z)$ sont de signes opposés, cela est vrai quand la partie imaginaire de z est plus grande que un, et $\varepsilon_N(z)$ est alors majoré par δ avec probabilité supérieure à $1 - 4Ne^{-c\delta\sqrt{N\eta}}$. Pour étendre ce résultat à z proche de l'axe réel, l'idée est d'utiliser que la partie imaginaire de $G_N(z)$ doit être grande, comme on s'attend à ce qu'elle converge vers la densité de la loi semi-circulaire qui est strictement positive dans la zone considérée. Nous pouvons effectuer cette estimation en partant de $z = E + i\eta$ avec $\eta \geq 1$ et en posant $z_n = E + i2^{-n}\eta$. On note $c(\kappa) = \inf_{|x|\leq 2-\kappa, \eta\in[0,1]} \Im G(x+i\eta) > 0$. Supposons que l'estimée $|G^N(z) - G(z)| \leq \epsilon < 4^{-1}c(\kappa)$ est vraie pour $z = z_n$; alors, par (35), on trouve

$$\Im G^N(z_{n+1}) \geq \frac{1}{2}\Im G^N(z_n) \geq \frac{1}{2}c(\kappa) - \varepsilon \geq \frac{1}{4}c(\kappa).$$

(28) donne donc l'estimée pour $z = z_{n+1}$ avec probabilité supérieure à $1 - 4Ne^{-c\delta\sqrt{N\eta 2^{-n-1}}}$ pourvu qu'on puisse poursuivre la récurrence, i.e.

$$C(\kappa)\eta_N(z_{n+1}) \leq C(\kappa)\frac{4\delta}{(c(\kappa))(c(\kappa) - 4\delta)} \leq \varepsilon < 4^{-1}c(\kappa)$$

ce qui est vrai pour δ assez petit et montre que,

$$\sup_{n\leq p} |G^N(z_n) - G(z_n)| \leq C(\kappa) \max_{1\leq k\leq N} |\varepsilon_k^N|$$

pour une constante finie $C(\kappa)$ et avec probabilité supérieure à $1 - 4N\sum_{n\leq p} e^{-c\delta\sqrt{N\eta 2^{-n-1}}}$. Ceci conclut l'argument en prenant $2^{-p-1} \geq (\log N)^4/N\eta$ pour que cette probabilité soit grande.

On peut rendre la convergence uniforme sur les variables E et η en s'appuyant sur la régularité des fonctions de Cauchy-Stieltjes et une méthode de maillage classique. □

Preuve du théorème 4.1. — Pour déduire ce théorème de la proposition 4.3, il faut comparer la probabilité d'un intervalle sous une probabilité P et la partie imaginaire de la transformée de Cauchy-Stieltjes G_P. Le point fondamental est l'égalité

$$\int_{-\eta^*}^{\eta^*} \Im G_P(E+x+i\eta)\frac{dx}{\pi} = -P * P_\eta([E-\eta^*, E+\eta^*])$$

avec P_γ la loi de Cauchy de paramètre γ. Si $P([x, x+\ell]) \leq C\ell$ pour tout $[x, x+\ell] \subset (-2+\kappa, 2-\kappa)$, nous déduisons que, pour tout $E \in (-2+2\kappa, 2-2\kappa)$, tout $a, \eta \in (0, \kappa)$

$$\begin{aligned} |P([E-\eta^*, E+\eta^*]) - P * P_\eta([E-\eta^*, E+\eta^*])| &\leq P_\eta([-\kappa,\kappa]^c) \\ +P_\eta([-\kappa,\kappa]\backslash[-a,a]) \sup_{y\in[-\kappa,\kappa]} P([E-\eta^*+y, E+\eta^*+y]) \\ + \sum_{E'=E\pm\eta^*} P([E'-a, E'+a]) \quad &\leq \frac{\eta}{\kappa} + \frac{\eta}{a}2C\eta^* + 2Ca. \end{aligned} \tag{29}$$

Si P est la mesure spectrale (ou la loi semi-circulaire) notre hypothèse est satisfaite par (25) (resp. par densité bornée) avec grande probabilité. En prenant $\eta = C(\kappa)\delta^2\eta^*$, $a = C(\kappa)\delta\kappa\eta^*$ pour une constante $C(\kappa)$ suffisamment petite, on déduit (18) de (19). Notons ici que la vitesse sous-exponentielle des déviations de la transformée de Stieljes devient une vitesse sous gaussienne pour les nombres d'occupation comme on doit prendre η de l'ordre de δ^2. □

Preuve du Corollaire 4.4. — Celle-ci est maintenant aisée. Par une simple formule d'algèbre linéaire, si $\mathbf{Y}_N v = \mu v$ pour un vecteur $v = (v_1, \tilde{V})$ de $\mathbb{C} \times \mathbb{C}^{N-1}$, on a

$$|v_1|^2 = \left(1 + N^{-1}\langle X_1, (\mu - \mathbf{Y}^{(1)})^{-2} X_1\rangle\right)^{-1} \leq \left(1 + N^{-1}\eta^{-2} \sum_{i:|\lambda_i^1 - \mu|\leq\eta} |\langle u_i^1, X_1\rangle|^2\right)^{-1}.$$

En utilisant de nouveau la propriété d'entrelacement des valeurs propres, les estimées de loi des grands nombres (24) et le théorème 4.1 on conclut qu'avec grande probabilité

$$|v_1|^2 \leq \left(1 + N^{-1}\eta^{-2}(\mathcal{N}_{[\mu-\eta,\mu+\eta]} - 1)\right)^{-1} \leq c\eta.$$

Nous choisissons finalement η de l'ordre de $N^{-1}(\log N)^4$ pour conclure. Cette estimation est uniforme sur $\mu \in [-2+\kappa, 2-\kappa]$ et sur les indices du vecteur v par symétrie.

4.2. Méthode des quatre moments d'après Tao et Vu

Dans **[32]**, T. Tao et V. Vu ont démontré que les statistiques locales des valeurs propres d'une matrice de Wigner dans l'intérieur du spectre sont les mêmes que celles d'une autre matrice de Wigner pourvu que les moments des coefficients coïncident jusqu'au degré quatre. T. Tao et V. Vu montrèrent qu'alors on peut comparer les statistiques locales de ces deux matrices en remplaçant deux par deux les coefficients de la matrice initiale par celles de la deuxième matrice. Afin de montrer ce résultat, l'idée est simplement d'estimer les dérivées des valeurs propres le long de ces interpolations et de montrer qu'elles sont petites devant le nombre N^2 des coefficients. Nous considérons deux matrices de Wigner $\mathbf{X}_N$ et $\mathbf{X}'_N$ dont les coefficients sont indépendants (mais non nécessairement équidistribués) pour $i \leq j$, centrés et de variance un. Dans le cas complexe, les parties imaginaire et réelle de la loi des coefficients sont supposées indépendantes. Par ailleurs, nous supposerons qu'elles ont des queues sous-exponentielles (2). Enfin, nous supposerons que les moments $C(\ell, p) = \mathbb{E}[\Re(X_{ij})^\ell \Im(X_{ij})^p]$ sont les mêmes que ceux de $\mathbf{X}'_n$ pour tout $\ell + p \leq 4$. On peut affaiblir cette hypothèse en supposant que les moments d'ordre trois (resp. quatre) diffèrent d'au plus $N^{-\frac{1}{2}-\delta}$ (resp. $N^{-\delta}$) pour un $\delta > 0$; cette généralisation sera cruciale dans la dernière section pour supprimer les hypothèses sur les moments d'ordre trois et quatre. Sous ces conditions, nous avons, si nous notons $\lambda_1(M_N) \leq \lambda_2(M_N) \cdots \leq \lambda_{N-1}(M_N) \leq \lambda_N(M_N)$ les valeurs propres ordonnées d'une matrice hermitienne M_N, le

THÉORÈME 4.5 (Theorem 15 de [**32**]). — *Pour tout $c_0 > 0$ suffisamment petit, pour tout $\varepsilon \in (0,1)$ et $k \geq 1$, pour toute fonction $F : \mathbb{R}^k \to \mathbb{R}$ satisfaisant*

$$|\nabla^j F(x)| \leq N^{c_0}$$

pour tout $0 \leq j \leq 5$ et $x \in \mathbb{R}^k$, nous avons, pour tout $i_1, \ldots, i_k \in [\varepsilon N, (1-\varepsilon)N]$,

$$\left|\mathbb{E}\left[F(\lambda_{i_1}(\sqrt{N}\mathbf{X}_N), \ldots, \lambda_{i_k}(\sqrt{N}\mathbf{X}_N))\right] - \mathbb{E}\left[F(\lambda_{i_1}(\sqrt{N}\mathbf{X}'_N), \ldots, \lambda_{i_k}(\sqrt{N}\mathbf{X}'_N))\right]\right| \leq N^{-c_0}.$$

En particulier, une matrice de Wigner hermitienne (resp. symétrique) dont les lois μ et ν ont au moins trois points dans leur support (resp. ont les mêmes moments d'ordre 4 que la matrice du GOE) a, asymptotiquement à l'intérieur du support, les mêmes fonctions de corrélations et la même distribution d'espacement entre deux valeurs propres que dans le cas gaussien, cf. théorèmes 0.1 et 0.2.

Il est clair que le premier résultat permet de comparer les statistiques locales de la matrice de Wigner à celles de la matrice gaussienne pourvu que les moments d'ordre quatre soient les mêmes, ce qui donne l'énoncé concernant les matrices symétriques. Dans le cas hermitien, on peut néanmoins comparer ces statistiques à celles des matrices (14) étudiées par K. Johansson. On voit alors [**32**, Corollary 30] qu'il suffit que les lois μ et ν aient au moins trois points dans leur support pour pouvoir identifier leurs moments à ceux d'une loi divisible par une gaussienne. Pour supprimer cette condition des trois points, qui en particulier exclut la loi de Bernoulli, il a été nécessaire [**10**] d'utiliser l'universalité des espacements pour des matrices du type (14) avec ε tendant vers zéro suffisamment rapidement avec N obtenue par Erdős, Ramirez, Schlein et Yau [**12**]. Ce résultat a été étendu à de nombreux autres ensembles que les matrices de Wigner hermitiennes, voir la prochaine section, mais malheureusement pas avec une vitesse suffisante pour permettre d'omettre toute condition sur les moments en général.

Nous n'allons pas donner une preuve complète du théorème 4.5, qui est assez longue, mais sa stratégie. Pour simplifier, nous nous plaçons dans le cas de coefficients réels.

On pose $A_N = \sqrt{N}\mathbf{X}_N$ et $A'_N = \sqrt{N}\mathbf{X}'_N$ et on souhaite montrer qu'en remplaçant un à un les coefficients de A_N par ceux de A'_N, on ne va pas faire trop varier les statistiques locales. Soit $F(z) = F(\lambda_{i_1}(A_N(z)), \dots, \lambda_{i_k}(A_N(z)))$ avec $A_N(z)$ la matrice dont les coefficients sont égaux à ceux de A_N sauf en (pq) et (qp) où ils valent z. On voudrait montrer que

$$\mathbb{E}[F(\sqrt{N}X_{pq})] - \mathbb{E}[F(\sqrt{N}X'_{pq})] = O(N^{-2-c_0}), \tag{30}$$

afin de remplacer deux à deux les coefficients de $\mathbf{X}_N$ par ceux de $\mathbf{X}'_N$ et obtenir ainsi une erreur de l'ordre de $N^2 \times N^{-2-c_0} = N^{-c_0}$. Pour cela, il nous suffit de prouver que pour $\ell \leq 5$

$$|\partial_z^\ell F(z)| = O(N^{-\ell+O(c_0)+o(1)}). \tag{31}$$

En effet, le développement de Taylor donne

$$F(z) = \sum_{k=0}^{4} \frac{1}{k!}(\partial_z^\ell F)(0) z^\ell + O(N^{-5+O(c_0)+o(1)})|z|^5,$$

que nous appliquons à $z = \sqrt{N}X_{pq}$ et $z = \sqrt{N}X'_{pq}$. Après avoir pris la moyenne sur ces deux variables et utilisé l'égalité des moments jusqu'à l'ordre quatre, la différence de (30) se trouve bornée par $\sqrt{N}^5 \times N^{-5+O(c_0)+o(1)} = N^{-\frac{5}{2}+O(c_0)+o(1)} \leq N^{-2-c_0}$.

Afin de comprendre comment obtenir (31), regardons le cas où $\ell = 1$. Dans ce cas, la première formule de variation d'Hadamard donne

$$\frac{d\lambda_i(A_N(z))}{dz} = \langle u_i(A_N(z)), \partial_z A_N(z) u_i(A_N(z))\rangle$$

avec $u_i(A_N(z))$ le vecteur propre de $A_N(z)$ pour la valeur propre $\lambda_i(A_N(z))$ et $\partial_z A_N(z) = e_p e_q^* + e_q e_p^*$ avec (e_p, e_q) les p-ième et q-ième vecteurs de la base canonique. Par conséquent, (31) est une conséquence, pour $\ell = 1$, de la délocalisation des vecteurs propres, corollaire 4.4, qui nous dit que $\langle u_i(A_N(z)), e_p\rangle$ est d'ordre au plus $N^{-\frac{1}{2}}(\log N)^4$ avec très grande probabilité. Pour le second ordre, la répulsion des valeurs propres va également jouer un rôle comme on trouve par la seconde formule de variation d'Hadamard que

$$\frac{d^2\lambda_i(A_N(z))}{d(z)^2} = -\sum_{j\neq i} \frac{(\langle u_j(A_N(z)), \partial_z A_N(z) u_i(A_N(z))\rangle)^2}{\lambda_j(A_N(z)) - \lambda_i(A_N(z))}.$$

De nouveau, par la délocalisation des vecteurs propres, on s'attend à ce que le numérateur soit au plus de l'ordre de $N^{-2}(\log N)^8$ alors que les valeurs propres devraient être très proches des quantiles de la loi semi-circulaire (comme indiqué par le théorème 4.1, voir aussi l'hypothèse (43)), ce qui implique que $\lambda_j(A_N(z)) - \lambda_i(A_N(z))$ devrait être de l'ordre de $j - i$. Par conséquent la somme sur j est au plus de l'ordre de $\log N$ et on peut de nouveau espérer que cette dérivée seconde soit au plus de l'ordre de $N^{-2}(\log N)^9$ avec grande probabilité. Les dérivées suivantes sont du même type.

L'idée est donc simple mais la preuve reste très technique dans la mesure où il faut d'une part obtenir ces estimations uniformément sur z, et d'autre part arriver à contrôler uniformément des quantités du type

$$\sum_{j\neq i}\frac{1}{|\lambda_j(A_N(z))-\lambda_i(A_N(z))|^2}.$$

Comme la probabilité que $|\lambda_j(A_N(z))-\lambda_i(A_N(z))|\geq n^{-c_0}$ pour un $c_0>0$ est seulement d'ordre $1-n^{-2c_0}$ (l'ensemble où deux valeurs propres coïncident étant de codimension deux), on comprend que ce type d'événement soit difficile à contrôler uniformément.

Nous allons maintenant décrire la preuve plus récente du théorème des quatre moments donnée par L. Erdős, H.T. Yau et J. Yin qui s'appuie sur une étude précise de la transformée de Cauchy-Stieltjes et qui est de nature plus analytique.

4.3. Méthode des quatre moments d'après Erdős, Yau et Yin

La méthode des quatre moments développée par Erdős, Yau et Yin repose sur un raffinement du théorème 4.1 qui consiste à montrer que les transformées de Cauchy-Stieltjes de deux matrices de Wigner sont asymptotiquement équivalentes quand la partie imaginaire de l'argument η tend vers zéro non plus comme $N^{-1}(\log N)^4$ comme dans (19) mais comme $N^{-1-\epsilon}$ pour un $\varepsilon>0$, ceci sous la condition que les moments d'ordre inférieur ou égal à quatre coïncident. Le théorème s'énonce comme suit ;

THÉORÈME 4.6. — *On effectue les mêmes hypothèses que dans le théorème 4.5, et on choisit k,F,c_0 de même. On note, pour une matrice hermitienne X, $G_X(z)=(z-X)^{-1}$ la résolvante de X en $z\in\mathbb{C}\backslash\mathbb{R}$. Soient $\mathbf{Y}_N=N^{-\frac12}\mathbf{X}_N$ et $\mathbf{Y}'_N=N^{-\frac12}\mathbf{X}'_N$. Alors, pour tout choix d'entiers $(\ell_1,\dots,\ell_k)$, on pose $z_j^m=E_j^m\pm i\eta$ pour $1\leq j\leq \ell_i$, $m=1,\dots,k$, avec $E_j^m\in[-2+\kappa,2-\kappa]$, pour c_0 assez petit, pour tout $\eta\in[N^{-1-\varepsilon},N^{-1}]$, avec un $\varepsilon>0$ assez petit, la différence*

(32)
$$\mathbb{E}\left[F\left(\frac{1}{N}\mathrm{Tr}(\prod_{j=1}^{\ell_p}G_{\mathbf{Y}_N}(z_j^p)),1\leq p\leq k\right)\right]-\mathbb{E}\left[F\left(\frac{1}{N}\mathrm{Tr}(\prod_{j=1}^{\ell_p}G_{\mathbf{Y}'_N}(z_j^p)),1\leq p\leq k\right)\right]$$

tend vers zéro quand N tend vers l'infini.

On peut facilement déduire un résultat du type du théorème 4.5 (mais où les indices $(i_1,\dots,i_k)$ ne sont pas fixés, au contraire du théorème 4.5 où on peut par exemple étudier les fluctuations de la $N/2$-ième valeur propre ordonnée, mais génériques, car on a accès seulement aux fonctions de corrélations à partir des transformées de Cauchy-Stieltjes) de nouveau en utilisant que la partie imaginaire de la transformée de Cauchy-Stieltjes n'est que la régularisation de la loi par une variable de Cauchy dont la taille est grosso modo donnée par la partie imaginaire du nombre complexe auquel elle est évaluée. Le théorème ci-dessus montre qu'on peut la prendre d'ordre négligeable devant N^{-1}, qui est l'espacement typique que nous cherchons à observer entre les valeurs propres. Au moins au sens faible, ceci nous permet d'approximer les fonctions de corrélations par leur convolution par des lois de Cauchy de paramètre négligeable par rapport à N^{-1}, ce qui s'exprime en fonction des différences de (32).

Preuve du Théorème 4.6. — Comme dans la preuve de T. Tao et V. Vu, l'idée est de remplacer progressivement les coefficients de $\mathbf{Y}_N$ par celles de $\mathbf{Y}'_N$. Soit $\mathbf{Y}^p_N$, $1 \leq p \leq N(N+1)/2$ une famille de matrices telle que $\mathbf{Y}^1_N = \mathbf{Y}_N$ et $\mathbf{Y}^{N(N+1)/2}_N = \mathbf{Y}'_N$, $\mathbf{Y}^p_N$ et $\mathbf{Y}^{p+1}_N$ ne différant qu'en deux sites, notés (i_p, j_p) et (j_p, i_p). Notons $\mathbf{Y}^k_N = \mathbf{Q} + \frac{1}{\sqrt{N}} V^k$ pour $k = p$ et $p+1$, avec $V^k = v^k_{i_p j_p} e_{i_p} e^*_{j_p} + v^k_{j_p i_p} e_{j_p} e^*_{i_p}$ pour $v^p_{ij} = X_{ij}$ et $v^{p+1}_{ij} = X'_{ij}$. Nous allons comparer la résolvante $R = (z - \mathbf{Q})^{-1}$ de $\mathbf{Q}$ à celle $S^k = (z - \mathbf{Y}^k_N)^{-1}$ de $\mathbf{Y}^k_N$ par le développement

$$R = S^k - N^{-\frac{1}{2}} S^k V^k S^k + \cdots + N^{-5} (S^k V^k)^{10} R. \tag{33}$$

Pour contrôler ce développement, nous allons utiliser des bornes a priori sur les coefficients de S^k et R. Tout d'abord, (19) peut être un peu améliorée (avec des techniques analogues, cf. [16]) pour montrer que, pour tout $\tau > 0$, tout $y \geq N^{-1+\tau}$,

(34)
$$P\left(\max_{0 \leq p \leq N(N+1)/2} \max_{1 \leq k \leq N} \max_{|E| \leq 2-\kappa} \left| \left(\frac{1}{\mathbf{Y}^p_N - E - iy} \right)_{kk} \right| \geq N^{2\tau} \right) \leq C N^{-\log\log N}.$$

On peut étendre le contrôle de la partie imaginaire à $y \geq N^{-1-\varepsilon}$ pour $\varepsilon > 0$ en notant l'inégalité

$$\left| \Im \left(\frac{1}{\mathbf{Y}^p_N - E - i\eta} \right)_{kk} \right| \leq \frac{y}{\eta} \left| \Im \left(\frac{1}{\mathbf{Y}^p_N - E - iy} \right)_{kk} \right| \quad 0 \leq \eta \leq y \tag{35}$$

qui implique avec $\eta = N^{-1+\tau}$ et $y \geq N^{-1-\varepsilon}$ que

(36)
$$P\left(\max_{0 \leq p \leq N(N+1)/2} \max_{1 \leq k \leq N} \max_{|E| \leq 2-\kappa} \left| \left(\Im \frac{1}{\mathbf{Y}^p_N - E - iy} \right)_{kk} \right| \geq N^{3\tau+\varepsilon} \right) \leq C N^{-\log\log N}.$$

Afin d'en déduire un contrôle sur tous les coefficients de la résolvante S^k, notons que, pour tout $1 \leq j, i \leq N$, si (u_i, λ_i) est une base orthonormée de vecteurs propres et valeurs propres associées de $\mathbf{Y}^k_N$, on a

$$\left| (z - \mathbf{Y}^k_N)^{-1}_{ij} \right| = \left| \sum_{\ell=1}^N \frac{u_\ell(i) u_\ell(j)}{z - \lambda_\ell} \right| \leq \left(\sum_{\ell=1}^N \frac{|u_\ell(i)|^2}{|z - \lambda_\ell|} \right)^{\frac{1}{2}} \left(\sum_{\ell=1}^N \frac{|u_\ell(j)|^2}{|z - \lambda_\ell|} \right)^{\frac{1}{2}}.$$

Pour utiliser notre borne (36), remarquons qu'avec $z = E + iy$, $|z - \lambda_\ell| > y$ et sur $U_m = \{\ell : 2^{m-1}y \leq |\lambda_\ell - E| \leq 2^m y\}$,

$$|\lambda_\ell - z|^{-1} \leq 2\Im(E - i2^m y - \lambda_\ell)^{-1}$$

alors que $|\lambda_\ell - z|^{-1}$ est uniformément bornée sur $|\lambda_\ell - E| \geq 1$. Donc, avec probabilité supérieure à $1 - CN^{-\log\log N}$, on a par (36)

$$\sum_{\ell=1}^{N} \frac{|u_\ell(i)|^2}{|z - \lambda_\ell|} \leq 2 \sum_{m=1}^{C\log N} \Im(\langle e_i, (E - i2^{-m}y - \mathbf{Y}_N^k)^{-1} e_i\rangle) \leq C \log N N^{3\tau+\varepsilon}.$$

Nous avons donc montré que, pour tout $y \geq N^{-1-\varepsilon}$,
(37)

$$P\left(\max_{0\leq p\leq N(N+1)/2}\ \max_{1\leq k,\ell\leq N}\ \max_{|E|\leq 2-\kappa} \left|\left(\Im\frac{1}{\mathbf{Y}_N^p - E - iy}\right)_{k\ell}\right| \geq N^{4\tau+\varepsilon}\right) \leq CN^{-\log\log N}.$$

Grâce à (33), on voit aisément que les coefficients de R satisfont la même borne, pourvu que $5\tau + \varepsilon < 2^{-1}$ puisque les coefficients de V^k sont bornés par N^τ avec grande probabilité et R_{ij} est majoré par $y^{-1} \leq N^{1+\varepsilon}$. On peut inverser R et S^k dans (33) de façon à développer S^k en fonction de R et V^k qui sont indépendants, pour avoir

$$S^k = R + N^{\frac12} RV^kR + N^{-1}(RV^k)^2R + N^{-\frac32}(RV^k)^3R + N^{-2}(RV)^4R + N^{-\frac52}(RV^k)^5S^k.$$

Nous démontrons finalement le théorème avec $k = 1 = \ell_1$ pour simplifier les notations. Posons

$$\zeta_1^k = \frac{1}{N}\mathrm{Tr}\left(S^k - R - N^{-\frac52}(RV^k)^5S^k\right) \quad \text{et} \quad \zeta_2^k = N^{-\frac52}\frac{1}{N}\mathrm{Tr}\left((RV^k)^5S^k\right)$$

si bien qu'avec grande probabilité, par (37), ζ_2^k est au plus d'ordre $N^{-\frac52+\varepsilon'}$ pour un $\varepsilon' = 5(5\tau + \varepsilon)$. La formule de Taylor nous donne

$$\begin{aligned}
\mathbb{E}\left[F\left(\frac{1}{N}\mathrm{Tr}(S^k)\right)\right] &= \mathbb{E}\left[F\left(\frac{1}{N}\mathrm{Tr}(R) + \zeta_1^k\right)\right] + O(\|F'\|_\infty)N^{-\frac52+\varepsilon'} \\
&= \sum_{p=1}^{4} \frac{\mathbb{E}[(\zeta_1^k)^p]}{p!}\mathbb{E}\left[F^{(p)}\left(\frac{1}{N}\mathrm{Tr}(R)\right)\right] \\
&\quad + \mathbb{E}\left[F^{(5)}\left(\frac{1}{N}\mathrm{Tr}(R) + \zeta'\right)\frac{(\zeta_1^k)^5}{5!}\right] + O(\|F'\|_\infty)N^{-\frac52+\varepsilon'}
\end{aligned}$$

où $\zeta' \in [0, \zeta_1]$ et nous avons utilisé l'indépendance de ζ_1 et R. En faisant la différence $\mathbb{E}[F(\frac{1}{N}\mathrm{Tr}(S^p))] - \mathbb{E}[F(\frac{1}{N}\mathrm{Tr}(S^{p+1}))]$ et en utilisant que tous les termes d'ordre inférieur ou égal à quatre du développement disparaissent par la condition des moments, on voit qu'il ne reste que des termes au moins d'ordre 5 en V^k, et ils sont au plus d'ordre $N^{-\frac{5}{2}+\varepsilon'}$, ce qui permet de conclure que, pourvu que les dérivées de F ne croissent pas trop vite,

$$\mathbb{E}\left[F\left(\frac{1}{N}\mathrm{Tr}\left(\frac{1}{z-\mathbf{Y}_N^p}\right)\right)\right]-\mathbb{E}\left[F\left(\frac{1}{N}\mathrm{Tr}\left(\frac{1}{z-\mathbf{Y}_N^{p+1}}\right)\right)\right]=O(N^{-\frac{5}{2}+\varepsilon'})$$

avec $\varepsilon' < 1/2$. Si les hypothèses sur les moments d'ordre trois et quatre sont affaiblies, on voit que cette différence reste de l'ordre de $N^{-\frac{3}{2}}|m_3-m_3|'+N^{-2}|m_4-m_4'|$ et reste donc petite devant N^{-2} sous les hypothèses effectuées. Sommant p de 1 à $N(N+1)/2$ permet de conclure. □

5. UNIVERSALITÉ À L'INTÉRIEUR DU SPECTRE ; UNE APPROCHE DYNAMIQUE

Je décris maintenant l'approche développée par L. Erdős, B. Schlein et H.T. Yau et leurs coauteurs pour étendre l'universalité des statistiques locales obtenue par K. Johansson **[21]** pour l'ensemble des matrices hermitiennes divisibles par une gaussienne (14) à de nombreux modèles (en particulier les matrices symétriques) et au cas où le terme de bruit ε tend vers zéro avec la dimension N. Cette étude permet également d'affaiblir les hypothèses sur les moments d'ordre 3 et 4 du théorème 4.5. Elle s'appuie sur la remarque que l'ensemble des matrices divisibles par une gaussienne peut être vu comme le processus à valeurs matrices engendré en faisant évoluer les coefficients de la matrice $\mathbf{V}_N$ selon un processus de Ornstein-Uhlenbeck. En effet, si $(\mathbf{G}_N(t))_{t\geq 0}$ est un processus à valeurs matricielles construit comme la matrice $\mathbf{G}_N$ du GUE mais avec des coefficients browniens à la place des gaussiennes, le processus de Ornstein-Uhlenbeck solution de

$$d\mathbf{X}_N(t)=d\mathbf{G}_N(t)-\frac{1}{2}\mathbf{X}_N(t)dt \qquad \mathbf{X}_N(0)=\mathbf{V}_N \tag{38}$$

a la même distribution que $\mathbf{X}_N$ de (14) à l'instant $t=\log(1-\varepsilon)^{-1}$. Le théorème d'universalité dit donc que les statistiques locales du spectre de ce processus sont approximativement les mêmes en temps petit qu'en temps t tendant vers l'infini, $\mathbf{X}_N(t)$ convergeant alors vers une matrice du GUE. En d'autres termes, l'équilibre local est atteint très vite. La théorie des processus aléatoires, et en particulier les techniques d'hydrodynamique, doivent donc permettre d'améliorer les résultats de K. Johansson. Par ces méthodes, il a été démontré **[11]** que l'universalité des statistiques locales des valeurs propres est vraie pour la matrice de (14) avec ε tendant vers zéro avec N, pourvu que ce soit plus lentement que $N^{-\frac{3}{4}}$. Nous allons exposer ci-dessous l'approche de **[14]** qui donne un résultat légèrement plus faible mais qui s'étend à un cadre beaucoup plus général, englobant en particulier le cas des matrices symétriques ou symplectiques, et qui ne nécessitent aucune formule explicite des densités utilisées par K. Johansson et dans **[9]**.

Il a été montré par F. Dyson **[8]** que les valeurs propres $(\lambda_1(t),\ldots,\lambda_N(t))_{t\geq 0} = \sqrt{N}(\eta_1(t),\ldots,\eta_N(t))_{t\geq 0}$ du processus de Ornstein-Uhlenbeck (38) sont décrites par le système d'équations différentielles stochastiques suivant :

$$d\eta_i(t) = \frac{\sqrt{2}}{\sqrt{\beta N}} dW_t^i + \frac{1}{N}\sum_{j\neq i}\frac{1}{\eta_i(t)-\eta_j(t)}dt - U'(\eta_i(t))dt\,, \tag{39}$$

avec $\beta = 2$, $U(x) = 2^{-1}x^2$ et N mouvements browniens indépendants $(W^i)_{1\leq i\leq N}$. Le paramètre β est introduit car (39) décrit aussi l'évolution des valeurs propres du processus de Ornstein-Uhlenbeck symétrique (cas $\beta = 1$) et symplectique (cas $\beta = 4$) (décrit par (38) mais où $\mathbf{G}_N$ est un mouvement brownien symétrique ou symplectique et non hermitien). La fonction U est supposée deux fois continuement differentiable et convexe. Dans le cas où U n'est pas quadratique, cette mesure ne correspond plus à la loi de matrices de Wigner, mais à des lois appelées modèles de matrices, également beaucoup étudiées. Le cas où U possède un terme logarithmique décrit le cas des matrices de Wishart **[15]**. Comme une partie de cette section s'étend au cas de fonctions U générales, nous nous placerons dans ce cadre bien que la seule application que nous développerons ici concerne le cas où U est quadratique. Ce système d'équations est bien défini et a une unique solution forte dès que $\beta \geq 1$, ce que nous supposerons par la suite (voir e.g. [**1**, section 4.3]). De façon moins probabiliste, le calcul d'Itô implique que la loi P_t^x de la distribution à l'instant t de la solution de (39) issue de x satisfait l'équation différentielle

$$\partial_t \int f(y)dP_t^x(y) = \int Lf(y)dP_t^x(y)$$

avec, si ∂_i est la dérivée par rapport à la i-ième variable, L le générateur donné par

$$Lf(y) = \frac{1}{\beta N}\sum_{i=1}^N \partial_i^2 f(y) + \sum_{1\leq i\leq N}\partial_i f(y)\left(-U'(y_i) + \frac{1}{N}\sum_{j\neq i}\frac{1}{y_i - y_j}\right)$$

pour toute fonction f deux fois continuement différentiable sur $\mathbb{R}^N$. Par intégration par partie, on vérifie aisément que P_t^x satisfait la propriété de symétrie

$$\int P_t^x(f)g(x)d\mathscr{P}_N(x) = \int f(x)P_t^x(g)d\mathscr{P}_N(x), \tag{40}$$

où $\mathscr{P}_N$ est la mesure de probabilités

$$\mathscr{P}_N(dx_1,\ldots,dx_N) = \frac{1}{Z_N}\prod_{i<j}|x_i - x_j|^\beta e^{-N\beta\sum_{i=1}^N U(x_i)}\prod dx_1\cdots dx_N\,. \tag{41}$$

On retrouve dans le cas $\beta = 2$ (resp. $= 1$, resp. $= 4$) et $U(x) = 4^{-1}\beta x^2$ la loi jointe des valeurs propres du GUE (4) (resp. des matrices de Wigner gaussiennes symétriques, resp. symplectiques), à une homothétie par $\sqrt{N}$ près.

Une conséquence de (40) est que $\mathcal{P}_N$ est une mesure stationnaire de la dynamique (39). La dynamique converge quand le temps tend vers l'infini vers cette mesure, avec une vitesse indépendante de N quand U est strictement convexe. Nous supposerons par la suite que la condition initiale est aléatoire, de loi $P_0(dx) = f_0(x)\mathcal{P}_N(dx)$, et noterons $P_t = \int P_t^x P_0(dx)$ la loi de la dynamique à l'instant t.

Le but de cette section est de montrer que les fonctions de corrélations de P_t et les espacements sous P_t sont approximativement les mêmes que celles de la mesure d'équilibre $\mathcal{P}_N$ sous des hypothèses assez générales sur la loi P_0 et pour des temps t tendant vers zéro avec N. Notons p_t^k (resp. $\rho_{k,N}$) la fonction de corrélations à k points sous P_t (resp. $\mathcal{P}_N$).

THÉORÈME 5.1. — *Soit $\mathbf{X}_N$ une matrice de Wigner hermitienne ou symétrique telle que ν et μ satisfont une inégalité de Sobolev logarithmique. Soit P_t le semi-groupe issu des valeurs propres de $X_N/\sqrt{N}$. Alors, il existe $\varepsilon > 0$ tel que, pour tout $\kappa > 0$ suffisamment petit tel que $[E-\kappa, E+\kappa] \subset\,]-2,2[$, tout $\varepsilon' > 0$, toute fonction test O à support compact,*

$$\lim_{N\to\infty} \int_{E-\kappa}^{E+\kappa} dE' \int_{\mathbb{R}^k} O(\alpha_1,\ldots,\alpha_k)(p^k_{N^{-2\varepsilon+\varepsilon'}} - \rho_{k,N})\left(E' + \frac{\alpha_1}{N},\ldots,E'+\frac{\alpha_k}{N}\right) d\alpha_1\cdots d\alpha_k = 0.$$

L'hypothèse de log-Sobolev est superflue, mais nous n'essayerons pas ici d'obtenir les meilleures conditions. Nous renvoyons à [**1**, section 2.3.2] pour la définition et l'implication de cette inégalité sur les valeurs propres des matrices de Wigner. On notera que le résultat est un peu plus faible que dans le théorème 4.5 comme il faut moyenner sur l'énergie E' et considérer des valeurs propres non ordonnées.

COROLLAIRE 5.2. — *Soit $\mathbf{X}_N$ une matrice de Wigner symétrique ou hermitienne telle que ν et μ satisfont une inégalité de log-Sobolev. Les fonctions de corrélations à l'intérieur du spectre ont les mêmes limites (au sens faible) que dans le cas gaussien si ν et μ satisfont $\int x^4 dP(x) > \left(\int x^3 dP(x)\right)^2 + 1$.*

Dans le cas hermitien, il a été démontré dans [**9**] que le ε du théorème 5.1 peut être pris d'ordre aussi proche de 3/4 que souhaité, si bien que comme les quatre premiers moments des coefficients de la matrice $\mathbf{X}_N(N^{-3/4+\delta})$ satisfont les hypothèses du théorème 4.5, on peut conclure sans aucune hypothèse sur les moments d'ordre supérieur ou égaux à trois. Cette précision est obtenue grâce à la formule explicite des densités. Dans le cadre général, la vitesse n'est pas suffisante et la condition sur les moments assure qu'on puisse utiliser le théorème des quatre moments [**32**, Lemma 28].

5.1. Processus et universalité des statistiques locales

Nous montrons tout d'abord que les statistiques locales de la dynamique sont approximativement les mêmes en un temps $N^{-\varepsilon}$ qu'en temps infini, sous certaines hypothèses sur le processus que nous vérifierons par la suite dans le cas des matrices de Wigner. Tout d'abord, nous supposerons qu'il existe une densité continue à support compact ρ ($\rho = \rho_{sc}$ si $U = \beta x^2/4$) d'une mesure de probabilité telle que, pour tout $a \leq b$

$$\lim_{N\to\infty} \sup_{t\geq 0} \left| \int \frac{1}{N} \sum_{j=1}^{N} \mathbf{1}(y_j \in [a,b]) dP_t(y) - \int_a^b \rho(x)dx \right| = 0 . \tag{42}$$

Par ailleurs, nous devons supposer que les y_j sont très proches des quantiles de la mesure limite, donnés par $\int_{-\infty}^{\gamma_j} \rho(x)dx = j/N$. L'hypothèse est qu'il existe $\varepsilon > 0$ tel que

$$Q_N := \sup_{t\geq 0} \int \sum_{j=1}^{N} (y_j - \gamma_j)^2 P_t(dy) \leq CN^{-2\varepsilon} . \tag{43}$$

Nous allons voir que ces hypothèses permettent de montrer que les espacements des particules sont universelles dans le sens suivant.

THÉORÈME 5.3. — *Supposons que les hypothèses* (42) *et* (43) *sont satisfaites. Supposons par ailleurs que la condition initiale P_0 est absolument continue par rapport à la mesure $\mathscr{P}_N$ et que l'entropie $S_{\mathscr{P}_N}(P_0) = \int \log \frac{dP_0}{d\mathscr{P}_N} dP_0$ croît au plus comme CN^m pour des constantes C, m indépendantes de N. Posons, pour un entier n et une fonction G à support sur $\mathbb{R}^n$, pour $m = (m_1 < \cdots < m_n) \in \mathbb{N}^n$ et $x \in \mathbb{R}^N$*

$$\mathscr{G}_{i,m}(x) := G\left(N(x_{i+m_1} - x_i), N(x_{i+m_2} - x_{i+m_1}), \ldots, N(x_{i+m_n} - x_{i+m_{n-1}})\right) . \tag{44}$$

Alors, pour tout $\delta > 0$, et toute fonction G continuement differentiable à support compact, il existe une constante finie C telle que pour tout ensemble $J \subset \{1, \ldots, N - m_n\}$,

$$\left| \int \frac{1}{N} \sum_{i\in J} \mathscr{G}_{i,m}(x) dP_{N^{-2\varepsilon+\delta}}(x) - \int \frac{1}{N} \sum_{i\in J} \mathscr{G}_{i,m}(x) d\mathscr{P}_N(x) \right| \leq CN^{-\varepsilon+3\delta}.$$

Ce théorème permet de montrer l'universalité des fonctions de corrélations sous une hypothèse supplémentaire sur la densité locale que voilà. Pour tout compact $I_0 \subset \{E : \rho(E) > 0\}$, et pour tout $\delta, \sigma > 0$, il existe une constante C_n telle que, pour tout intervalle $I \subset I_0$, $|I| \geq N^{-1+\sigma}$, pour tout $K \geq 1$,

$$\sup_{\tau \geq N^{-2\varepsilon+\delta}} P_\tau \left(\sharp\{i : x_i \in I\} \geq KN|I| \right) \leq C_n K^{-n} . \tag{45}$$

Si p_t^k (resp. $\rho_{k,N}$) représente la fonction de corrélations de la loi P_t (resp. $\mathscr{P}_N$), nous allons montrer le résultat suivant.

Théorème 5.4. — *Supposons* (42), (43) *et* (45) *vérifiées. Supposons que* $S_{\mathcal{P}_N}(P_0)$ *est au plus de l'ordre de* CN^m *pour des constantes* C *et* m *indépendantes de* N*. Soit* E *un point tel que* $\rho(E) > 0$*. Alors, pour tout* $\delta > 0$ *assez petit, pour tout* $p \geq 1$*, pour toute fonction régulière à support compact* O*, pour tout* $\kappa > 0$*, la limite quand* N *tend vers l'infini de*

$$\sup_{t \geq N^{-2\varepsilon+\delta}} \int_{E-\kappa}^{E+\kappa} dE' \int O(y_1, \dots, y_k)(p_t^k - \rho_{k,N})(E' + \frac{y_1}{N\rho(E)}, \dots, E' + \frac{y_p}{N\rho(E)}) dy_1 \cdots dy_p$$

est nulle.

5.1.1. *Preuve du théorème 5.3.*— Ce théorème ne nécessite pas la convexité du potentiel qui régit la diffusion (en fait $\inf U''(x) > -\infty$ suffirait). Sa preuve est fondée sur une idée très astucieuse de **[14]** qui est d'introduire une dynamique auxiliaire qui converge plus rapidement vers l'équilibre, puis de montrer que ce changement de dynamique affecte peu les statistiques locales. Cette dynamique a pour mesure d'équilibre la loi

$$dQ_N(x) = \frac{e^{-N\sum_{i=1}^N W_j(x_j)}}{Z_N} d\mathcal{P}_N(x)$$

avec $W_j(x) = (2R^2)^{-1}(x_j - \gamma_j)^2$ et satisfait le système d'équations

$$(46) \quad d\tilde{\eta}_i(t) = \frac{\sqrt{2}}{\sqrt{\beta N}} dW_t^i + \frac{1}{N} \sum_{j \neq i} \frac{1}{\tilde{\eta}_i(t) - \tilde{\eta}_j(t)} dt - U'(\tilde{\eta}_i(t))dt - \frac{1}{R^2}(\tilde{\eta}_i(t) - \gamma_i)dt\,.$$

Nous prendrons par la suite R tendant vers zéro avec N, si bien que cette dynamique a tendance à garder les valeurs propres beaucoup plus proches des quantiles de la mesure d'équilibre. Notons Q_t^x la loi de cette dynamique à l'instant t. Le générateur de cette dynamique est donné par

$$(47) \quad \tilde{L}f(y) = \frac{1}{\beta N} \sum_{i=1}^N \partial_i^2 f + \sum_{1 \leq i \leq N} \partial_i f(y)\Big(-U'(y_i) + \frac{1}{N}\sum_{j \neq i} \frac{1}{y_i - y_j} - \frac{1}{R^2}(y_i - \gamma_i)\Big)\,.$$

De nouveau, nous avons la propriété de symétrie

$$\int f(x)(\tilde{L}g)(x) dQ_N(x) = -\frac{1}{\beta N} \sum_{j=1}^N \int \partial_i f(x) \partial_i g(x) dQ_N(x) =: -D(f,g)\,.$$

On note $D(f) = D(f,f)$ la forme de Dirichlet sous Q_N.

Proposition 5.5. — *Soit* q_t *la densité, par rapport à la mesure* Q_N*, de la loi* $Q_t = \int Q_t^x q_0(x) dQ_N(x)$ *de la solution de* (46) *issue de* x *de loi* $q_0(x)dQ_N(x)$ *pour une fonction* $q_0 \in L^\infty(Q_N)$.

1. *Nous avons les estimations suivantes*

$$\partial_t D(\sqrt{q_t}) \leq -\frac{\beta}{R^2} D(\sqrt{q_t}) - \frac{\beta}{N^2} \int \sum_{i,j=1}^{N} \frac{(\partial_i \sqrt{q_t} - \partial_j \sqrt{q_t})^2}{(x_i - x_j)^2} dQ_N \,, \tag{48}$$

$$\frac{1}{N^2} \int_0^\infty dt \int \sum_{i,j=1}^{N} \frac{(\partial_i \sqrt{q_t} - \partial_j \sqrt{q_t})^2}{(x_i - x_j)^2} dQ_N \leq D(\sqrt{q_0}) \,, \tag{49}$$

et le contrôle de l'entropie $S_{Q_N}(f) = \int f \log f dQ_N$*, pour une constante* $C > 0$*,*

$$S_{Q_N}(q_t) \leq e^{-\frac{Ct}{R^2}} S_{Q_N}(q_0) \,. \tag{50}$$

2. *Pour tout n entier et toute fonction différentiable G à support compact, pour tout $m \in \mathbb{N}^m$ et avec la notation (44), pour tout $\tau > 0$, il existe une constante finie $C(G)$*

$$\left|\int \frac{1}{N} \sum_{i \in J} \mathscr{G}_{i,m}(x) d\left(q_0 Q_N - Q_N\right)(x)\right| \leq C(G)\Big(\sqrt{D(\sqrt{q_0})\tau N^{-1}} + \sqrt{S_{Q_N}(q_0)} e^{-C\frac{\tau}{2R^2}}\Big) \,.$$

Preuve. — La preuve du premier point de ce résultat s'inspire des idées de **[3]** qui a montré que les diffusions associées à des mesures d'équilibre de densité strictement log-concave convergent vers ces mesures et qu'on peut estimer la vitesse de cette convergence. L'argument doit cependant être adapté car la mesure est dégénérée aux points où deux particules coïncident et la diffusion vit sur la chambre de Weyl $\Delta = \{(\lambda_i) \in \mathbb{R}^N, \lambda_i \neq \lambda_j, \text{ si } i \neq j\}$ (voir [**1**, Theorem 4.3.2]) plutôt que sur $\mathbb{R}^N$. Nous n'entrerons pas dans les détails de cette adaptation et supposerons que toutes les fonctions manipulées sont bien régulières. Pour borner la forme de Dirichlet de $\sqrt{q_t}$, notons que comme pour toute fonction deux fois continuement dérivable, par symétrie,

$$\partial_t \int f(x) q_t(x) dQ_N(x) = \int \tilde{L} f(x) q_t(x) dQ_N(x) = \int f(x) \tilde{L} q_t(x) dQ_N(x)$$

la continuité de q_t assure que $\partial_t q_t(x) = \tilde{L} q_t(x)$ pour tout x. Posons $h_t = \sqrt{q_t}$ et utilisons

$$\partial_t h_t = \frac{1}{2h_t} \partial_t h_t^2 = \frac{1}{2h_t} \tilde{L} h_t^2 = \tilde{L} h_t + \frac{1}{2h_t} \Gamma_1(h_t, h_t)$$

où Γ_1 est le carré du champ

$$\Gamma_1(f,g) := \tilde{L}(fg) - f\tilde{L}g - g\tilde{L}f = \frac{1}{\beta N} \sum_{i=1}^{N} (\partial_i f)(\partial_i g) \,.$$

Par conséquent, nous trouvons que

$$\begin{aligned} \partial_t \beta D(h_t) &= \frac{1}{N} \int \nabla h_t . \nabla \left(\tilde{L} h_t + \frac{1}{2h_t} \Gamma_1(h_t, h_t)\right) dQ_N(x) \\ &= \frac{1}{N} \int \Gamma_1(h_t, \tilde{L} h_t) dQ_N(x) + \frac{1}{N} \int \nabla h_t \cdot \nabla \left(\frac{1}{2h_t} \Gamma_1(h_t, h_t)\right) dQ_N(x) \,. \end{aligned} \tag{51}$$

Par ailleurs notons que pour tout i

$$\partial_i \tilde{L} h_t - \tilde{L} \partial_i h_t = \sum_{j=1}^{N} \mathrm{Hess}\, H_{i,j} \partial_j h_t$$

où le Hamiltonien H est le logarithme de la densité de Q_N par rapport à la mesure de Lebesgue et $\mathrm{Hess} H_{ij} = \partial_i \partial_j H$. Par conséquent, en injectant cette formule dans le premier terme de (51) et en utilisant de nouveau la formule de symétrie de $\tilde{L}$, on trouve

(52)
$$\partial_t \beta D(h_t) = \frac{1}{N} \int \left(\nabla h_t \cdot \mathrm{Hess}\, H_{i,j} \nabla h_t - \frac{1}{\beta N} \sum_{i,j=1}^{N} (\partial_i \partial_j h_t - \frac{\partial_i h_t \partial_j h_t}{h_t})^2 \right) dQ_N \, .$$

La suite repose sur la stricte concavité du Hamiltonien donnée par la convexité de U qui implique que, pour tout $v \in \mathbb{R}^N$,

$$(53) \qquad - \sum_{i,j=1}^{N} v_i \mathrm{Hess}\, H_{i,j}(x) v_j \geq \frac{1}{R^2} \sum_{i=1}^{N} v_i^2 + \frac{1}{N} \sum_{i<j} \frac{(v_i - v_j)^2}{(x_i - x_j)^2}$$

et donc (52) donne (48). On déduit de (48), en négligeant le second terme de son membre de droite, qu'il existe une constante $C > 0$ telle que

$$(54) \qquad D(\sqrt{q_t}) \leq e^{-\frac{Ct}{R^2}} D(\sqrt{q_0})$$

si bien que $D(\sqrt{q_t})$ tend vers zero quand t tend vers l'infini, ce qui montre en passant la convergence de la dynamique vers la mesure d'équilibre. On obtient (49) en intégrant (48), en négligeant le premier terme de son membre de droite, par rapport au temps et en utilisant que $D(\sqrt{q_t})$ tend vers zéro à l'infini. La stricte concavité du Hamiltonien H, $\mathrm{Hess}\, H(x) \leq -R^{-2} I$ implique que la mesure Q_N satisfait une inégalité de log-Sobolev logarithmique de constante R^2, qui peut se démontrer par des arguments similaires à ceux que nous venons de développer, cf. [**3**] ou [**1**, Theorem 4.4.18]. L'inégalité de log-Sobolev permet de majorer l'entropie par la forme de Dirichlet et donne (50) car

$$(55) \qquad \partial_t S_{Q_N}(q_t) = -D(\sqrt{q_t}) \leq -R^{-2} S_{Q_N}(q_t) \, .$$

Le deuxième point de la proposition est fondé sur l'idée que la mesure Q_N est aussi égale à la limite quand t tend vers l'infini de $q_t Q_N$. Pour simplifier les notations, nous considérons le cas où $n = m = 1$. Notons, pour tout $s \leq t$,

$$\Delta_{t,s}(G) := \left| \int \frac{1}{N} \sum_{i=1}^{N} G(N(x_i - x_{i+1}))(q_t - q_s) dQ_N \right| .$$

Nous devons majorer $\Delta_{0,\infty}(G)$ que nous décomposons en

$$(56) \qquad \Delta_{0,\infty}(G) \leq \Delta_{0,\tau}(G) + \Delta_{\tau,\infty}(G) \, .$$

Pour le second terme, notons que, comme G est uniformément bornée,

$$\Delta_{\tau,\infty}(G) \leq \|G\|_\infty \|q_\tau.\mathcal{Q}_N - \mathcal{Q}_N\|_{\mathrm{TV}} \leq \|G\|_\infty \sqrt{2S_{\mathcal{Q}_N}(q_\tau)}$$

par l'inégalité de Csiszár-Kullback-Pinsker [**35**, p.293], avec $\|.\|_{\mathrm{TV}}$ la norme de variation totale. (50) implique par conséquent que

$$\Delta_{\tau,\infty}(G) \leq \|G\|_\infty \sqrt{2S_{\mathcal{Q}_N}(q_0)} e^{-\frac{C\tau}{2R^2}}\,. \tag{57}$$

Par ailleurs, nous avons

$$\begin{aligned}
\Delta_{0,\tau}(G) &= \frac{1}{\beta N}\int_0^\tau \sum_{j=1}^N \partial_j\Big(\frac{1}{N}\sum_{i=1}^N G(N(x_i - x_{i+1}))\Big)\partial_j q_u d\mathcal{Q}_N \\
&\leq \int_0^\tau du \Big| \int \frac{1}{N}\sum_{i=1}^N G'(N(x_i - x_{i+1}))(\partial_i q_u - \partial_{i+1} q_u) d\mathcal{Q}_N \Big| \\
&\leq \int_0^\tau du \int \frac{1}{N}\sum_{i=1}^N |\sqrt{q_u} G'(N(x_i - x_{i+1}))||x_i - x_{i+1}| \frac{|\partial_i q_u - \partial_{i+1} q_u|}{\sqrt{q_u}|x_i - x_{i+1}|} d\mathcal{Q}_N \\
&\leq \Big(\int_0^\tau du \int \sum_{i=1}^N G'(N(x_i - x_{i+1}))^2 (x_i - x_{i+1})^2 q_u d\mathcal{Q}_N\Big)^{\frac{1}{2}} \\
&\quad \times \Big(\int_0^\tau du \int \frac{1}{N^2}\sum_{i=1}^N \frac{(\partial_i\sqrt{q_u} - \partial_{i+1}\sqrt{q_u})^2}{(x_i - x_{i+1})^2} d\mathcal{Q}_N\Big)^{\frac{1}{2}} \leq C(G)\sqrt{\frac{D(\sqrt{q_0})\tau}{N}}
\end{aligned}$$

où nous avons utilisé l'inégalité de Cauchy-Schwarz, puis (49) et le fait que, comme G est à support compact, seuls les i tels que $x_i - x_{i+1}$ est d'ordre N^{-1} contribuent à la somme ci-dessus. (56) et (57) permettent de conclure. □

Afin de démontrer le théorème 5.3, il nous faut supprimer la régularisation par le potentiel dépendant de R. L'idée est de prendre comme condition initiale $q_0(x) = \frac{dP_t}{d\mathcal{Q}_N}$ si bien que le second point de la proposition 5.5 donne

(58)

$$\Big|\int \frac{1}{N}\sum_{i\in J} \mathcal{G}_{i,m}(x) d\left(P_t - \mathcal{Q}_N\right)(x)\Big| \leq C(G)\sqrt{N^{-1}D(\sqrt{\frac{dP_t}{d\mathcal{Q}_N}})t} + C(G)\sqrt{S_{\mathcal{Q}_N}(\frac{dP_t}{d\mathcal{Q}_N})} e^{-\frac{t}{2R^2}}\,.$$

Afin de conclure, il suffit d'utiliser cette inégalité avec $P_0 = p_0$ et $P_t = P_0 = \mathcal{P}_N$ et de faire la différence pourvu que les entropies et les formes de Dirichlet de $\frac{dP_t}{d\mathcal{Q}_N}$ et $\frac{d\mathcal{P}_N}{d\mathcal{Q}_N}$ soient bien contrôlées. Le lemme suivant fournit les estimations nécessaires.

LEMME 5.6. — *Soit $t = R^2 N^{\varepsilon'}$ avec $\varepsilon' > 0$ aussi petit que souhaité et supposons $S_{\mathcal{Q}_N}(\frac{dP_0}{d\mathcal{Q}_N}) \leq CN^m$. Alors, avec $\Lambda = \sup_{t\geq 0}\sum_j \int (\frac{x_j - \gamma_j}{R^2})^2 dP_t(x)$, on a*

$$D(\sqrt{\frac{dP_t}{d\mathcal{Q}_N}}) \leq CN\Lambda \qquad S_{\mathcal{Q}_N}(\frac{dP_t}{d\mathcal{Q}_N}) \leq CR^2 N\Lambda\,.$$

Pour achever la preuve du théorème 5.3, prenons $R^2 = N^{-\varepsilon+(\delta-\varepsilon')/2}$ tendant vers zéro si bien que (43) implique que $\Lambda \le CN^{2(\delta-\varepsilon')}$. Comme $S_{Q_N}(dq/dQ_N) \le S_{\mathscr{P}_N}(q) + \Lambda$ pour $q = \mathscr{P}_N$ ou $q = P_0$, ces entropies croissent au plus polynomialement sous nos hypothèses et (58) avec $t = R^2 N^{\varepsilon'}$ permet de conclure pour ε' et δ assez petits.

Preuve du Lemme. — De nouveau, on dérive l'entropie pour trouver que

$$\begin{aligned}\partial_t S_{Q_N}(\frac{dP_t}{dQ_N}) &= -D(\sqrt{\frac{dP_t}{dQ_N}}) + \int L\frac{dP_t}{dQ_N}dQ_N \\ &= -D(\sqrt{\frac{dP_t}{dQ_N}}) + \sum_j \int \frac{x_j-\gamma_j}{R^2}\partial_j \frac{dP_t}{dQ_N}dQ_N \\ &\le -D(\sqrt{\frac{dP_t}{dQ_N}}) + \frac{1}{2}D(\sqrt{\frac{dP_t}{dQ_N}}) + 2N\sum_j \int (\frac{x_j-\gamma_j}{R^2})^2 dP_t\end{aligned}$$

où nous avons utilisé que $\tilde{L}$ est invariant dans la seconde ligne, et l'inégalité de Schwarz dans la dernière. En utilisant l'inégalité de log-Sobolev, on déduit que

$$\partial_t S_{Q_N}(\frac{dP_t}{dQ_N}) \le -\frac{C}{R^2}S_{Q_N}(\frac{dP_t}{dQ_N}) + 2N\Lambda\,,$$

ce qui donne la seconde inégalité du lemme après intégration comme $N^m e^{-\frac{C}{R^2}t} \le 2N\Lambda R^2$ avec nos choix. On obtient la seconde inégalité en en intégrant l'inégalité ci-dessus entre τ et $\tau/2$ et en utilisant la décroissance de $t \to D(\sqrt{\frac{dP_t}{dQ_N}})$. □

5.1.2. *Preuve du théorème 5.4.*— Nous montrons maintenant comment déduire du théorème 5.3 concernant les espacements des valeurs propres la convergence des fonctions de corrélations en moyenne, sous l'hypothèse (45) d'absolue continuité locale par rapport à la mesure de Lebesgue. Pour cela, notons que nous pouvons supposer la fonction test O symétrique et que nous avons par définition

$$\begin{aligned}&\int_{E-\kappa}^{E+\kappa} dE' \int O(y_1,\ldots,y_p) f_t^p(E' + \frac{y_1}{N\rho_{sc}(E)},\ldots,E' + \frac{y_p}{N\rho_{sc}(E)})dy_1\cdots dy_p \\ &= p!\int_{E-\kappa}^{E+\kappa} dE' \int \sum_{i_1<i_2\cdots<i_p} O(N(x_{i_1}-E'),N(x_{i_2}-E'),\ldots,N(x_{i_p}-E'))dP_t(x) \\ &= p!\sum_{0<m_1<\cdots<m_{p-1}} \int_{E-\kappa}^{E+\kappa} dE' \int \sum_{i=1}^{N} \tilde{O}\left(N(x_i-E'),N(x_{i+m_k}-x_{i+m_{k-1}})_{1\le k\le p-1}\right) dP_t(x)\end{aligned}$$

avec $m_0 = 0$ et $\tilde{O}(u_1,\ldots,u_p) = O(\rho_{sc}(E)u_1, \rho_{sc}(E)(u_2-u_1),\ldots)$. Nous avons dans le théorème 5.3 estimé des quantités du type de celles ci-dessus, avec deux différences notables. Tout d'abord, nous avons ici une somme a priori infinie de ces quantités ; ce problème est résolu par la compacité de $\tilde{O}$ et l'absolue continuité locale qui permet de montrer que seulement le cas où $\sum m_j$ et i sont bornés contribue à la somme. Par ailleurs, les fonctions ne dépendaient pas du premier terme $N(x_i - E')$ dans le théorème 5.3 ; ce problème disparaît grâce à la moyennisation sur E' (qui peut être étendue à la droite réelle par la remarque ci-dessus). Nous ne détaillons pas plus cet argument, voir [**15**, section 7].

5.2. Applications aux matrices de Wigner

Dans cette partie, nous indiquons comment déduire le théorème 5.1 et le corollaire 5.2 des théorèmes 5.3 et 5.4.

En ce qui concerne le théorème 5.1, il s'agit simplement de montrer que les hypothèses (42) et (43) sont vérifiées puisqu'alors c'est une conséquence directe du théorème 5.4. Mais P_t est la loi d'une matrice de Wigner $\mathbf{X}_N(t)$ dont les coefficients ont des lois ν_t et μ_t dont les queues sont sous-gaussiennes par définition. (42) avec $\rho = \rho_{sc}$ est donc une conséquence de (1), et (45) une conséquence de (25). (43) est un raffinement de (18) qui permet d'inclure les bords ; nous ne le détaillerons pas ici et renvoyons le lecteur à [**15**, section 8]. La preuve de [**16**, Theorem 6.3] nécessite l'hypothèse de log-Sobolev, afin de montrer que les valeurs propres sont proches de leurs moyennes [**1**, Theorem 2.3.5].

Pour le corollaire 5.2, l'idée est de démontrer qu'on peut approcher les lois ν et μ par leur évolution sous la dynamique ν_t et μ_t pour que les statistiques locales soient peu modifiées. La première approche développée dans [**14**] a consisté à construire un « flot inverse » de manière à bien approximer n'importe quelle loi ν par $P_t\tilde{\nu}_t$ pour une certaine loi $\tilde{\nu}_t$ pour laquelle le théorème 5.1 est valide. $\tilde{\nu}_t$ est appelé « flot inverse ». Plus récemment, voir [**16**], l'argument a plutôt été d'utiliser le théorème des quatre moments 4.5. Notons déjà que, si les moments d'ordre 3 de μ et ν s'annulent, la différence des moments d'ordre quatre entre ν_t et ν est au plus d'ordre $t = N^{-\varepsilon}$ pour un $\varepsilon > 0$ si bien que nous pouvons conclure. Sinon, il suffit d'utiliser un mélange de l'argument du flot inverse et du théorème des quatre moments pour construire, cf. [**16**, Lemma 6.5], une loi $\tilde{\nu}_t$, satisfaisant l'inégalité de log-Sobolev, et telle que $\nu_t = P_t\tilde{\nu}_t$ satisfasse les conditions des moments jusqu'à l'ordre 4 quand $t = N^{-\varepsilon}$ pour un $\varepsilon > 0$.

Remerciements

Je remercie vivement G. Aubrun, C. Bernardin, P. Biane, L. Erdős, C. Garban, M. Ledoux, S. Péché et T. Tao pour leurs multiples commentaires concernant une version préliminaire de cet exposé, et le projet ANR GranMa ANR-08-BLAN-0311 pour faciliter ma recherche sur les matrices aléatoires.

RÉFÉRENCES

[1] G. Anderson, A. Guionnet & O. Zeitouni – *An introduction to random matrices*, Studies adv. math., vol. 118, Cambridge Univ. Press, 2009.

[2] A. Auffinger, G. Ben Arous & S. Péché – Poisson convergence for the largest eigenvalues of heavy tailed random matrices, *Ann. Inst. Henri Poincaré Probab. Stat.* **45** (2009), p. 589–610.

[3] D. Bakry & M. Émery – Diffusions hypercontractives, in *Séminaire de probabilités, XIX, 1983/84*, Lecture Notes in Math., vol. 1123, Springer, 1985, p. 177–206.

[4] E. Brézin & S. Hikami – Correlations of nearby levels induced by a random potential, *Nuclear Phys. B* **479** (1996), p. 697–706.

[5] ———, Spectral form factor in a random matrix theory, *Phys. Rev. E* **55** (1997), p. 4067–4083.

[6] P. Deift & D. Gioev – *Random matrix theory : invariant ensembles and universality*, Courant Lecture Notes in Math., vol. 18, Courant Institute of Mathematical Sciences, New York, 2009.

[7] P. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venakides & X. Zhou – Strong asymptotics of orthogonal polynomials with respect to exponential weights, *Comm. Pure Appl. Math.* **52** (1999), p. 1491–1552.

[8] F. J. Dyson – A Brownian-motion model for the eigenvalues of a random matrix, *J. Mathematical Phys.* **3** (1962), p. 1191–1198.

[9] L. Erdős, S. Péché, J. A. Ramírez, B. Schlein & H.-T. Yau – Bulk universality for Wigner matrices, *Comm. Pure Appl. Math.* **63** (2010), p. 895–925.

[10] L. Erdős, J. A. Ramírez, B. Schlein, T. Tao, V. Vu & H.-T. Yau – Bulk universality for Wigner Hermitian matrices with subexponential decay, *Math. Res. Lett.* **17** (2010), p. 667–674.

[11] L. Erdős, J. A. Ramírez, B. Schlein & H.-T. Yau – Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation, *Electron. J. Probab.* **15** (2010), p. 526–603.

[12] L. Erdős, B. Schlein & H.-T. Yau – Local semicircle law and complete delocalization for Wigner random matrices, *Comm. Math. Phys.* **287** (2009), p. 641–655.

[13] ———, Wegner estimate and level repulsion for Wigner random matrices, *Int. Math. Res. Not.* **2010** (2010), p. 436–479.

[14] ———, Universality of random matrices and local relaxation flow, prépublication arXiv :0907.5605.

[15] L. ERDŐS, B. SCHLEIN, H.-T. YAU & J. YIN – The local relaxation flow approach to universality of the local statistics for random matrices, prépublication arXiv :0911.3687.

[16] L. ERDŐS, H.-T. YAU & J. YIN – Bulk universality for generalized Wigner matrices, prépublication arXiv :1001.3453.

[17] P. J. FORRESTER – The spectrum edge of random matrix ensembles, *Nuclear Phys. B* **402** (1993), p. 709–728.

[18] D. L. HANSON & F. T. WRIGHT – A bound on tail probabilities for quadratic forms in independent random variables, *Ann. Math. Statist.* **42** (1971), p. 1079–1083.

[19] HARISH-CHANDRA – Fourier transforms on a semisimple Lie algebra. I, *Amer. J. Math.* **79** (1957), p. 193–257.

[20] C. ITZYKSON & J. B. ZUBER – The planar approximation. II, *J. Math. Phys.* **21** (1980), p. 411–421.

[21] K. JOHANSSON – Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices, *Comm. Math. Phys.* **215** (2001), p. 683–705.

[22] ______, Universality for certain Hermitian Wigner matrices under weak moment conditions, prépublication arXiv :0910.4467.

[23] O. KHORUNZHIY – High moments of large Wigner random matrices and asymptotic properties of the spectral norm, prépublication arXiv :0907.3743.

[24] M. L. MEHTA – *Random matrices*, Pure and Applied Math., vol. 142, Elsevier/Academic Press, 2004.

[25] S. PÉCHÉ – Universality in the bulk of the spectrum for complex sample covariance matrices, prépublication arXiv :0912.2493.

[26] S. PÉCHÉ & A. SOSHNIKOV – Wigner random matrices with non-symmetrically distributed entries, *J. Stat. Phys.* **129** (2007), p. 857–884.

[27] A. RUZMAIKINA – Universality of the edge distribution of eigenvalues of Wigner random matrices with polynomially decaying distributions of entries, *Comm. Math. Phys.* **261** (2006), p. 277–296.

[28] Y. SINAI & A. SOSHNIKOV – Central limit theorem for traces of large random symmetric matrices with independent matrix elements, *Bol. Soc. Brasil. Mat. (N.S.)* **29** (1998), p. 1–24.

[29] A. SOSHNIKOV – Universality at the edge of the spectrum in Wigner random matrices, *Comm. Math. Phys.* **207** (1999), p. 697–733.

[30] T. TAO & V. VU – Random matrices : universality of local eigenvalue statistics up to the edge, *Comm. Math. Phys.* **298** (2010), p. 549–572.

[31] ———, Random covariance matrices : Universality of local statistics of eigenvalues, prépublication arXiv :0912.0966.

[32] ———, Random matrices : Universality of local eigenvalue statistics, prépublication arXiv :0906.0510.

[33] C. A. TRACY & H. WIDOM – Level-spacing distributions and the Airy kernel, *Comm. Math. Phys.* **159** (1994), p. 151–174.

[34] ———, The distribution of the largest eigenvalue in the Gaussian ensembles : $\beta = 1, 2, 4$, in *Calogero-Moser-Sutherland models (Montréal, QC, 1997)*, CRM Ser. Math. Phys., Springer, 2000, p. 461–472.

[35] C. VILLANI – *Topics in optimal transportation*, Graduate Studies in Math., vol. 58, Amer. Math. Soc., 2003.

[36] E. P. WIGNER – Characteristic vectors of bordered matrices with infinite dimensions, *Ann. of Math.* **62** (1955), p. 548–564.

[37] J. WISHART – The generalized product moment distribution in samples from a normal multivariate population, *Biometrika* **20A** (1928), p. 32–52.

Alice GUIONNET

École normale supérieure de Lyon
U.M.P.A.
UMR 5669 du CNRS
46, allée d'Italie
F–69364 LYON Cedex 07
E-mail : `Alice.Guionnet@umpa.ens-lyon.fr`