Mémoire de magistère

Guillaume AUBRUN

24 septembre 2002

Introduction

La première partie de ce texte est une introduction rapide à un problème classique de la théorie locale des espaces de Banach, l'estimation du diamètre du compact de Banach-Mazur. Les espaces aléatoires, intimement liés aux matrices aléatoires, ont joué un rôle fulgurant dans la résolution de ce problème. On peut espérer qu'une meilleure connaissance des matrices aléatoires puisse permettre de résoudre d'autres problèmes ouverts de la théorie locale.

Le reste du mémoire consiste en l'exposition des différentes méthodes qui permettent d'obtenir des résultats de concentration pour la plus grande valeur propre d'une matrice aléatoire de l'Ensemble Gaussien Unitaire : manipulation directe des variables gaussiennes, propriété de concentration de la mesure dans l'espace gaussien, utilisation des opérateurs intégraux, résultats à la Tracy-Widom.

Ce mémoire inclut un certain nombre de démonstrations, qui peuvent parfois sembler un peu trop "techniques" ; mais j'ai délibérément voulu conserver un panorama assez large des différentes méthodes utilisables, en montrant le genre de preuves qu'elles engendrent. Néanmoins, un tel panorama ne saurait être complet, et j'énumère en dernière partie quelques autres idées qui ont eu ou pourraient avoir des conséquences intéressantes pour les questions abordées ici.

1 Les motivations de la théorie locale des espaces de Banach

Le principe de la théorie locale des espaces de Banach est d'étudier ces espaces à travers leurs sous-espaces de dimension finie (mais grande). Le premier outil nécessaire à cette étude est un moyen de mesurer la distance entre deux espaces vectoriels normés de même dimension n. On définit à cet effet la distance de Banach-Mazur :

DÉFINITION 1.0.1 — *Soit E et F deux espaces vectoriels normés de dimension n. On définit*

$$d_{BM}(E,F) := \inf_T ||T|| \cdot ||T^{-1}|| \tag{1}$$

L'infimum est pris sur tous les isomorphismes T de E sur F.

Cette "distance" vérifie une inégalité triangulaire multiplicative

$$d_{BM}(E,G) \leqslant d_{BM}(E,F) d_{BM}(F,G) \tag{2}$$

D'autre part, deux espaces sont à distance 1 si et seulement si ils sont isométriques. En effet, supposons que $d_{BM}(E,F)=1$. On peut alors trouver une suite d'isomorphismes (T_n) de E sur F tels que $||T_n|| \leqslant 1$ et $||T_n^{-1}|| \leqslant 1+\frac{1}{n}$. Par compacité, cette suite admet un point d'accumulation qui est bien une isométrie de E sur F.

Remarquons également que $d_{BM}(E,F) = d_{BM}(E^*,F^*)$ (il suffit de passer à l'isomorphisme adjoint).

Un des théorèmes fondamentaux est le suivant (on note ℓ_p^n l'espace L_p associé à la mesure de comptage sur un espace à n points)

THÉORÈME 1.0.1 (JOHN) — *Tout espace vectoriel normé E de dimension n vérifie*

$$d_{BM}(E,\ell_2^n) \leqslant \sqrt{n} \tag{3}$$

Cette inégalité est optimale, par exemple on peut voir que $d_{BM}(\ell_2^n, \ell_\infty^n) = \sqrt{n}$. Remarquons tout d'abord que la borne $\sqrt{n}$ est atteinte par l'identité naturelle entre ces deux espaces. Considérons maintenant n'importe quel isomorphisme T de norme 1 de ℓ_∞^n sur ℓ_2^n et soit (e_j) la base canonique de ℓ_∞^n. Écrivons l'égalité du parallélogramme dans ℓ_2^n

$$\sum_{j=1}^{n} ||Te_j||^2 = \frac{1}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} ||\sum_{j=1}^{n} \varepsilon_j Te_j||^2 \leqslant \frac{1}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} ||\sum_{j=1}^{n} \varepsilon_j e_j||^2 = 1 \tag{4}$$

Il existe donc un indice j tel que $||Te_j|| \leqslant \frac{1}{\sqrt{n}}$, ce qui prouve que $||T^{-1}|| \geqslant \sqrt{n}$ et donc que $||T|| \cdot ||T^{-1}|| \geqslant \sqrt{n}$.

Définissons maintenant $\mathcal{K}_n$ comme l'espace des classes d'isométrie d'espaces vectoriels normés de dimension n. Les remarques précédentes montrent que $(\mathcal{K}_n, d_{BM})$ est un espace métrique, on peut par ailleurs montrer qu'il est compact. Cet espace est appelé compact de Banach-Mazur, ou compact de Minkowski selon les terminologies.

Un corollaire immédiat du théorème précédent est

Corollaire 1.0.2 — *Le diamètre de (K_n, d_{BM}) est borné par n.*

On peut naturellement se demander si cette estimation est précise ou non. Dans tout les cas où l'on sait calculer précisémenet la distance entre deux espaces vectoriels normés explicites, cette distance n'excède jamais $\sqrt{n}$ (à une constante multiplicative près). Cependant, on peut prouver le théorème suivant

Théorème 1.0.3 (Gluskin) — *Il existe une constante numérique $c > 0$ telle que pour tout n, le diamètre de (K_n, d_{BM}) vaut au moins cn.*

La démonstration de ce théorème (ou tout au moins la seule démonstration connue, voir par exemple [6]) est entièrement probabiliste : on construit deux espaces de dimension n de manière aléatoire et indépendante, et on montre qu'avec probabilité grande leur distance est de l'ordre de n. Ces espaces aléatoires peuvent par exemple être des projections de ℓ_1^{3n} sur $\mathbb{R}^n$. On considère l'image par une matrice aléatoire d'une structure normée fixe.

Ce théorème illustre bien la puissance des espaces générés aléatoirement, qui fournirent des contre-exemples à de nombreux problèmes de ce genre. Pour cela il faut avoir des estimations précises sur la norme (ainsi que sur d'autres paramètres) des matrices aléatoires qui interviennent.

2 Matrices aléatoires

2.1 Présentation de l'ensemble EGU.

Il existe beaucoup de modèles différents de matrices aléatoires : matrices carrées ou non, réelles ou complexes, symétriques ou non, ... On peut également imposer ou non l'indépendance des coefficients et choisir la loi qu'ils suivent.

Dans la suite de cette introduction nous ne considérerons qu'un modèle très particulier de matrices aléatoires, pour lequel les calculs sont particulièrement simples : l'ensemble gaussien unitaire EGU. Il existe des analogues de la plupart des résultats cités dans des situations plus générales ; nous parlerons de ceci dans la dernière partie.

On note $\mathcal{M}_n(\mathbb{C})$ l'ensemble des matrices complexes $n \times n$ et $\mathcal{H}_n$ le sous-ensemble formé des matrices hermitiennes. On note $A = (a_{ij})_{i,j=1\ldots n}$ l'élément générique de $\mathcal{H}_n$. $U(n)$ désigne le groupe des matrices unitaires.

$\mathcal{H}_n$ est muni d'un produit scalaire donné par $\langle A, B\rangle = \mathrm{tr}(AB^*)$. Il possède donc une structure euclidienne et une mesure de Lebesgue associée. Cette mesure de Lebesgue peut s'écrire comme un produit de mesures de Lebesgue 1-dimensionnelles

$$dA = \prod_{i=1}^{n} da_{ii} \prod_{i<j} d\Re a_{ij} \prod_{i<j} d\Im a_{ij} \tag{5}$$

Définition 2.1.1 — *L'ensemble gaussien unitiare (EGU) est l'espace probabilisé $(\mathcal{H}_n, \mathbb{P}_n)$, où $\mathbb{P}_n$ est définie par sa densité par rapport à la mesure de Lebesgue*

$$d\mathbb{P}_n := \frac{1}{c_n} e^{-\frac{n}{2}\operatorname{tr}(A^2)} dA \tag{6}$$

La valeur de la constante c_n est ajustée afin d'obtenir une mesure de probabilité. On note $\mathbb{E}_n$ l'espérance associée à la probabilité $\mathbb{P}_n$.

Remarque : comme la fonctionnelle $A \to \operatorname{tr}(A^2)$ est invariante sous l'action par conjugaison du groupe unitaire $U(n)$, on voit que la mesure $\mathbb{P}_n$ est aussi invariante sous cette même action.

Proposition 2.1.1 — *Soit A une matrice de l'EGU, de taille $n \times n$. Alors*

1. $a_{ij} = \overline{a_{ji}}$
2. *Les variables aléatoires (a_{ii}), $(\Re a_{ij})_{i\leqslant j}$ et $(\Im a_{ij})_{i\leqslant j}$ sont indépendantes.*
3. $\forall\, i, a_{ii}$ *suit une loi gaussienne* $N(0, \frac{1}{n})$.
4. $\forall\, i < j, \Re a_{ij}$ *et* $\Im a_{ij}$ *suivent une loi gaussienne* $N(0, \frac{1}{2n})$.

De plus, les conditions 1–4 déterminent uniquement la mesure $\mathbb{P}_n$.

Démonstration — La propriété 1 est évidente puisque A est hermitienne. Pour les autres points, il suffit d'écrire

$$\frac{1}{c_n} e^{-\frac{n}{2}\operatorname{tr}(A^2)} = \frac{1}{c_n} \prod_{i=1}^{n} e^{-\frac{n}{2}a_{ii}^2} \prod_{i<j} e^{-n(\Re a_{ij})^2} \prod_{i<j} e^{-n(\Im a_{ij})^2} \tag{7}$$

et d'utiliser l'écriture (5) de la mesure de Lebesgue. ■

Cette dernière proposition est particulièrement importante lorsque l'on cherche à effectuer des simulations numériques : c'est une représentation concrète de la mesure $\mathbb{P}_n$.

On peut définir de même l'ensemble gaussien orthogonal (EGO) en munissant l'espace $\mathcal{S}_n$ des matrices symétriques réelles de taille n d'une mesure de probabilité définie de manière analogue, qui sera invariante par l'action du groupe orthogonal $O(n)$. Bien que beaucoup de résultats s'étendent au cas orthogonal, on se contentera de les énoncer dans le cas unitaire, pour lequel les calculs sont souvent plus simples.

Sauf mention contraire, on supposera désormais que l'on travaille avec EGU ; l'expression "matrice aléatoire" devra être comprise comme signifiant "élément de EGU".

Dans la suite, on s'intéressera au comportement des valeurs propres de ces matrices, lorsque la dimension n tend vers l'infini.

2.1.1 Loi du demi-cercle

Soit $A_n \in \mathcal{H}_n$. Les valeurs propres de A_n sont réelles et on les appelle (dans l'ordre décroissant)

$$\lambda_1(A_n) \geqslant \ldots \geqslant \lambda_n(A_n) \tag{8}$$

Définition 2.1.2 — *La mesure empirique associée à A_n est la mesure de probabilité sur $\mathbb{R}$ suivante*

$$N(A_n) := \frac{1}{n}\sum_{j=1}^{n} \delta_{\lambda_j(A_n)} \tag{9}$$

où δ_x désigne la masse de Dirac au point x.

Soit $\mathbb{P}$ la probabilité produit $\prod_{n=1}^{\infty} \mathbb{P}_n$. On peut alors énoncer le résultat suivant

Théorème 2.1.2 (Loi du demi-cercle de Wigner) — *Soit (A_n) une suite de matrices de l'EGU, où A_n est de taille $n \times n$. Alors, $\mathbb{P}$-presque sûrement, la suite de mesures $N(A_n)$ converge étroitement vers $\mu_\cap$.*

Ici, $\mu_\cap$ désigne la la probabilité sur $\mathbb{R}$ dont la densité a la forme d'un demi-cercle

$$d\mu_\cap := \frac{1}{2\pi} 1_{[-2,2]} \sqrt{4-x^2}dx \tag{10}$$

On dit qu'une suite de mesures (μ_n) converge étroitement vers une mesure μ si pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ continue bornée, $\int f d\mu_n$ converge vers $\int f d\mu$.

Démonstration — Nous esquissons ici une preuve, due à L.Pastur (cf [5]), qui utilise la transformée de Stieltjes.

Définition 2.1.3 (Transformée de Stieltjes) — *Soit μ une mesure de probabilité sur $\mathbb{R}$. On appelle tranformée de Stieltjes de μ la fonction S_μ définie pour $z \in \mathbb{C} \setminus \mathbb{R}$ par*

$$S_\mu(z) := \int_{\mathbb{R}} \frac{\mu(d\lambda)}{\lambda - z} \tag{11}$$

Proposition 2.1.3 — *La fonction S_μ ainsi définie vérifie les propriétés suivantes*

1. *Elle est analytique sur $\mathbb{C} \setminus \mathbb{R}$.*
2. *Elle laisse globalement invariant le demi-plan supérieur.*
3. *$S_\mu(\overline{z}) = \overline{S_\mu(z)}$.*
4. *$\lim_{y\to\infty} y|S_\mu(\imath y)| = 1$.*

De plus, une fonction f qui vérifient les conditions 1–4 précédentes (appelée "fonction de Herglotz"). est nécessairement la transformée de Stieltjes d'une probabilité sur $\mathbb{R}$, et on peut retrouver cette dernière grâce à la formule d'inversion de Perron-Frobenius, lorsque I est un intervalle de $\mathbb{R}$ tel que $\mu(\partial I) = 0$

$$\mu(I) = \lim_{\varepsilon\to 0} \frac{1}{\pi} \int_I \Im S_\mu(\lambda + \imath\varepsilon) d\lambda \tag{12}$$

De plus la transformation de Stieltjes est un homéomorphisme de l'espace des mesures de probabilité sur $\mathbb{R}$, muni de la topologie étroite, vers l'espace des fonctions de Herglotz, muni de la topologie de la convergence uniforme sur les compacts de $\mathbb{C} \setminus \mathbb{R}$.

Voyons maintenant comment cet outil intervient dans la démonstration du théorème de Wigner. Soit (A_n) une suite de matrices de l'EGU, où A_n est de taille $n \times n$. On note $g_n = S_{N(A_n)} = \frac{1}{n}\mathrm{tr}(A_n - \bullet \mathrm{Id})^{-1}$, et donc d'après la proposition précédente il suffit de montrer que $\mathbb{P}$-presque sûrement, g_n converge vers $S_\cap$ uniformément sur tout compact. Par ailleurs, ils est plus commode de travailler l'espérance de g_n ; pour cela on définit $f_n(z) := \mathbb{E}_n g_n(z)$.

Notons $B_\eta := \{z \in \mathbb{C}, |\Im z| > \eta\}$. La preuve des lemmes suivants, que nous admettons, est assez calculatoire (voir [5])

Lemme 2.1.4 — $\forall\, z \in B_\eta, \mathbb{V}\mathrm{ar}_n[g_n(z)] \leqslant \frac{C}{n^2\eta^2}$.

Lemme 2.1.5 — *La fonction f_n vérifie l'inégalité, pour $z \in B_\eta$*

$$|f_n(z) + \frac{1}{z} + \frac{1}{z} f_n(z)^2| \leqslant \frac{C(\eta)}{n} \tag{13}$$

Par ailleurs, il est facile de voir que g_n (et donc f_n) est bornée sur B_η par $\frac{1}{\eta}$. Par le théorème de compacité de Montel, on peut extraire une sous-suite $f_{\sigma(n)}$ qui converge uniformément sur tout compact de B_η vers une fonction limite f. En passant à la limite dans (13), f doit nécessairement vérifier l'équation $f(z)^2 + zf(z) + 1 = 0$, autrement dit f est de la forme

$$f(z) = \frac{1}{2}(-z \pm \sqrt{z^2 - 4}) \tag{14}$$

Mais f est une fonction de Herglotz, par conséquent on doit prendre la détermination de la racine carrée sur B_η qui se comporte en $z+O(1)$ quand z tend vers l'infini. Ansin f est bien déterminée, ce qui veut dire qu'en fait toute la suite (f_n) converge vers f.

Pour passer de f_n à g_n, on utilise l'inégalité de Bienaymé-Tchebychev et le lemme 2.1.4 (ici $z \in B_\eta$)

$$\mathbb{P}_n[|f_n(z) - g_n(z)| > \varepsilon] \leqslant \frac{1}{\varepsilon^2}\mathbb{V}\mathrm{ar}[g_n(z)] \leqslant \frac{C}{\varepsilon^2\eta^2 n^2} \tag{15}$$

La série $\sum_{n=1}^{\infty} \mathbb{P}_n[|f_n(z) - g_n(z)| > \varepsilon]$ est donc convergente, et le lemme de Borel-Cantelli nous permet de conclure que $\mathbb{P}$-presque sûrement, $f_n(z) - g_n(z)$ tend vers 0, c'est-à-dire que $g_n(z)$ tend vers $f(z)$. En faisant tendre η vers 0, cette conclusion reste vraie pour tout z dans $\mathbb{C} \setminus \mathbb{R}$.

Soit Q une partie dense de $\mathbb{C} \setminus \mathbb{R}$. On sait que $\mathbb{P}$-presque sûrement, g_n tend vers f sur Q. Or les fonctions g_n et f sont toutes analytiques, et le théorème de Vitali nous permet de conclure qu'en fait g_n tend cers f uniformément sur tout compact de $\mathbb{C} \setminus \mathbb{R}$.

On sait que f est une fonction de Herglotz. Soit μ la fonction dont f est la transformée de Stieltjes. En écrivant la formule de Perron-Frobenius pour un intervalle I de $\mathbb{R}$, on obtient

$$\mu(I) = \lim_{\varepsilon \to 0} \frac{1}{\pi} \int_I \Im f(\lambda + \imath\varepsilon) d\lambda = \frac{1}{2\pi} \int_{I\cap[-2,2]} \sqrt{4-\lambda^2} d\lambda = \mu_\cap(I) \tag{16}$$

Ainsi, $\mathbb{P}$-presque sûrement, $S_{N(A_n)}$ converge uniformément sur tout compact vers $S_\cap$, donc $N(A_n)$ converge étroitement vers $\mu_\cap$.

■

Il existe de nombreuses variantes de cet énoncé, plus ou moins fortes, et l'on en connaît également beaucoup de démonstrations. À la démonstration d'origine, qui reposait sur une analyse combinatoire assez ardue, ont succédé des preuves plus conceptuelles. Signalons par exemple, outre l'utilisation de la transformation de Stieltjes, l'approche des probabilités libres, qui donne à la distribution "en demi-cercle" un rôle qui est l'analogue non-commutatif de celui joué par la loi gaussienne (il existe en particulier un analogue du théorème central limite qui s'inscrit parfaitement dans le cadre présenté ici).

3 Un premier résultat de concentration

Cependant, tous les résultats précédents sont des résultats globaux, au sens où ils ne permettent pas de prévoir le comportement des valeurs propres au voisinage d'un certain point. De même, ils ne précisent rien sur les grandes valeurs propres : par exemple la loi du demi-cercle nous permet d'affirmer que le nombre de valeurs propres plus grandes que 2 est en $o(n)$, mais a priori ce nombre peut tendre vers l'infini.

3.1 Majoration de l'espérance de la norme

En fait, on peut montrer que l'espérance de la norme (= de la plus grande valeur propre) d'une matrice de l'EGU est bornée uniformément en n. C'est relativement facile à voir si on ne cherche pas à avoir la borne optimale.

Proposition 3.1.1 — *On a l'estimation uniforme*

$$\sup_{n\in\mathbb{N}} \mathbb{E}_n \lambda_1(A_n) < \infty \tag{17}$$

Démonstration — Soit $(x_\alpha)_{\alpha\in\mathcal{R}}$ un $\frac{1}{2}$-réseau fini de S_{n-1}, c'est-à-dire tel que $\forall\, x \in S_{n-1}, \exists\, \alpha$ tel que $||x - x_\alpha|| \leqslant \frac{1}{2}$. Alors $||A_n x|| \leqslant \frac{1}{2}||A_n|| + ||A_n x_\alpha||$, donc on prenant le supremum pour x dans la sphère, on obtient

$$||A_n|| \leqslant 2 \max_\alpha ||A_n x_\alpha|| \tag{18}$$

Un raisonnement similaire, en écrivant $||A_n x_\alpha|| = \sup_{x \in S_{n-1}} \langle A_n x_\alpha, x \rangle$, donne

$$||A_n|| \leqslant 4 \max_{\alpha,\beta} \langle A_n x_\alpha, x_\beta \rangle \tag{19}$$

Or $\langle A_n x_\alpha, x_\beta \rangle = \sum_{i,n} a_{ij} x_\alpha^i \overline{x_\beta^j}$, donc d'après la proposition 2.1.1, cette quanitité est une somme de gaussiennes indépendantes, et la loi de $\langle A_n x_\alpha, x_\beta \rangle$ est une loi $N(0, \sigma^2)$ avec $\sigma^2 \leqslant \frac{1}{n}$. Nous allons maintenant utiliser le lemme suivant

Lemme 3.1.2 — *Soit $(X_i)_{1 \leqslant i \leqslant N}$ une suite de N variables aléatoires gaussiennes $N(0,1)$ (non nécessairement indépendantes). Alors il existe une constante $C > 0$ telle que*

$$\mathbb{E} \max_i |X_i| \leqslant C \sqrt{\log N} \tag{20}$$

Démonstration — On note Ψ la queue de la distribution gaussienne standard

$$\Psi(x) := \int_x^\infty e^{-t^2/2} \frac{dt}{\sqrt{2\pi}} \tag{21}$$

On commence par remarquer que la fonction $x \to \frac{1}{\sqrt{2\pi}} e^{-(\Psi^{-1}(x))^2/2}$ est concave, ce qui permet de montrer que dans notre problème, le cas le pire est celui, symétrique, où pour tout i, la probabilité que X_i soit la plus grande des N variables aléatoires vaut $\frac{1}{N}$, et où cela correspond à $X_i \geqslant \Psi^{-1}(\frac{1}{N})$.

Pour régler ce cas, on utilise des encadrements asymptotiquement très précis de Ψ : par exemple, pour x positif, on a

$$\frac{2e^{-x^2/2}}{x + \sqrt{x^2+4}} \leqslant \Psi(x) \leqslant \frac{4e^{-x^2/2}}{3x + \sqrt{x^2+8}} \tag{22}$$

En tournant autour de ces inégalités, on arrive sans embûche à la majoration annoncée en $\sqrt{\log(N)}$. ∎

Revenons à la démonstration de la proposition. On savait grâce à (19) que

$$\mathbb{E}_n ||A_n|| \leqslant \mathbb{E}_n \max_{\alpha,\beta} \langle A_n x_\alpha, x_\beta \rangle \tag{23}$$

Donc le lemme précédent nous permet d'affirmer que $\mathbb{E}_n ||A_n|| \leqslant \frac{1}{\sqrt{n}} \sqrt{\log(N^2)}$, où N est le cardinal du réseau $\mathcal{R}$. Mais on sait bien estimer le cardinal d'un tel réseau :

Lemme 3.1.3 — *On peut trouver un $\frac{1}{2}$-réseau de S_{n-1} de cardinalité inférieure à 5^n.*

Démonstration — C'est très simple : soit $(x_i)_{1 \leqslant i \leqslant N}$ un ensemble $\frac{1}{2}$-séparé maximal dans S_{n-1}. C'est nécessairement un $\frac{1}{2}$-réseau de S_{n-1}. Considérons maintenant la familles de boules n-dimensionnelles $(B(x_i, \frac{1}{4}))_{1 \leqslant i \leqslant N}$. Ces boules sont disjointes et contenues dans $B(0, \frac{5}{4})$, donc la comparaison des volumes donne immédiatement que $N(\frac{1}{4})^n \leqslant (\frac{5}{4})^n$, donc $N \leqslant 5^n$. ∎

Dans le cas qui nous intéresse, on obtient $\mathbb{E}_n ||A_n|| \leqslant C$, ce qui est bien l'estimation annoncée.

∎

On peut en réalité obtenir une majoration exacte pour $E_n ||A_n||$ en utilisant des résultats sur la comparaison de processus gaussiens, comme par exemple le lemme de Slepian (voir [2]). On obtient alors par cette méthode $\mathbb{E}_n ||A_n|| \leqslant 2$.

Ce résultat est bien sûr asymptotiquement optimal, car d'après la loi du demi-cercle on a, $\mathbb{P}$-presque sûrement, $\mathbb{E}_n ||A_n|| > 2 - \varepsilon$ pour n assez grand (à ε fixé).

3.2 Concentration de la mesure

Maintenant que l'on connaît la limite précise de l'espérance de la plus grande valeur propre, on aimerait estimer la probabilité que cette valeur propre dévie beaucoup de son espérance. C'est possible grâce à des résultats de concentration de la mesure.

Regardons d'abord ce phénomène sur la sphère euclidienne S_{n-1}, munie de la distance géodésique. On note σ_{n-1} la mesure de Haar normalisée sur S_{n-1}. Si $A \subset S_{n-1}$, on note A_ε l'ensemble des points qui sont à distance au plus ε de A et $\partial\sigma_{n-1}(A) := \overline{\lim}_{\varepsilon\to 0}\frac{1}{\varepsilon}(\sigma_{n-1}A_\varepsilon - \sigma_{n-1}(A))$. $\partial\sigma_{n-1}(A)$ est en quelque sorte la mesure $n-2$-dimensionnelle de la frontière de A. On peut alors énoncer le résultat suivant

Proposition 3.2.1 (Inégalité isopérimétrique sur la sphère) — *Soit $A \subset S_{n-1}$ et $C \subset S_{n-1}$ une calotte (c'est-à-dire une boule pour la distance géodésique) telle que $\sigma_{n-1}(A) = \sigma_{n-1}(C)$. Alors $\forall\, \varepsilon > 0, \sigma_{n-1}(A_\varepsilon) \geqslant \sigma_{n-1}(C_\varepsilon)$.*

Corollaire 3.2.2 — *Soit $A \subset S_{n-1}$ et C une calotte de même mesure que A. Alors $\partial\sigma_{n-1}(A) \geqslant \partial\sigma_{n-1}(C)$.*

Autrement dit, parmi les parties de S_{n-1} d'un volume donné, ce sont les calottes qui ont la plus "petite" frontière. Ce résultat est bien sûr l'analogue sphérique de l'inégalité isopérimétrique usuelle dans $\mathbb{R}^n$: parmi les parties de $\mathbb{R}^n$ de volume donné, ce sont les boules euclidiennes qui ont la plus petite surface. Ce problème était déjà connu sous l'Antiquité (en dimension 2, c'est le problème de Didon).

Démonstration — Il n'existe pas de démonstration élémentaire de cette inégalité. La méthode traditionnelle pour démontrer l'analogue euclidien est, en partant d'une partie donnée de $\mathbb{R}^n$, de lui appliquer une suite symétrisations bien choisies ("symétrisations de Steiner") pour tendre vers la boule euclidienne et conservant le volume et en diminunant la surface. Dans le cas sphérique on peut appliquer une méthode similaire en remplaçant la symétrisation de Steiner par la symétrisation à deux points ([1]).

■

Corollaire 3.2.3 (Concentration de la mesure sur la sphère) — *Soit $f : S_{n-1} \to \mathbb{R}$ une fonction 1-lipschitizienne (par rapport à la distance géodésique). On note M_f une médiane de f. On a alors l'estimation, pour $\varepsilon > 0$*

$$\sigma_{n-1}(\{x, |f(x) - M_f| > \varepsilon\}) \leqslant \sqrt{\pi/2}e^{-\varepsilon^2(n-2)/2} \tag{24}$$

Cette propriété est extrêmement importante. Elle affirme qu'avec grande probabilité, la valeur d'une fonction suffisamment régulière définie sur la sphère est proche de la valeur médiane. Il est également important de noter que le terme exponentiel tend vers 0 lorsque n augmente ; autrement dit, la propriété de concentration est d'autant meilleure que la dimension est grande.

Démonstration — Si $f(x) > M_f + \varepsilon$, alors par la propriété de Lipschitz $d(x, \{y, f(y) \leqslant M_f\}) > \varepsilon$. Par conséquent $\{x, |f(x) - M_f| \leqslant \varepsilon\} \subset \{x, f(x) \leqslant M_f\}_\varepsilon \cup \{x, f(x) \geqslant M_f\}_\varepsilon$.

Mais $\sigma_{n-1}(\{x, f(x) \geqslant M_f\}) \geqslant \frac{1}{2}$, donc par l'inégalité isopérimétrique, $\sigma_{n-1}(\{x, f(x) \geqslant M_f\}) \geqslant \sigma_{n-1}(C_\varepsilon)$, où C est une demi-sphère.

Un calcul sans grande difficulté, faisant intervenir des intégrales de Wallis, nous donne $\sigma_{n-1}(C_\varepsilon) \geqslant 1 - \sqrt{\pi/8}e^{-\varepsilon^2(n-2)/2}$. On obtient bien le résultat annoncé. ■

La fonction qui nous intéresse ici (la norme d'une matrice de l'EGU) n'est pas définie sur la sphère, mais sur l'espace gaussien. En effet, on peut identifier $\mathcal{H}_n$ à $\mathbb{R}^N$ avec $N = n^2$ en prenant pour base $\{\sqrt{n}a_{ii}\}_{1\leqslant i\leqslant n} \cup \{\sqrt{2n}\Re a_{ij}\}_{i<j} \cup \{\sqrt{2n}\Im a_{ij}\}_{i<j}$. Sous cette identification, l'image de la mesure $\mathbb{P}_n$ est la mesure gaussienne usuelle $d\gamma_N$ sur $\mathbb{R}^N$ ($|\bullet|$ désignant la norme euclidienne)

$$d\gamma_N := \frac{1}{(2\pi)^{N/2}}e^{-|x|^2/2}dx \tag{25}$$

Mais il y a un argument, dû à Poincaré, qui permet de passer de la sphère à l'espace gaussien

Lemme 3.2.4 (Lemme de Poincaré) — *Soit n un entier naturel fixé. Si $N > n$, P_N désigne la projection naturelle de $\mathbb{R}^N$ sur $\mathbb{R}^n$ (projection sur les n premières coordonnées). On note $\hat{\sigma}_{N-1}$ la mesure de Haar normalisée sur la sphère $\sqrt{N}S^{N-1}$. Alors pour tout borélien A de $\mathbb{R}^n$, $P_N * \hat{\sigma}_{N-1}(A)$ converge vers $\gamma_N(A)$.*

On peut alors déduire une inégalité de concentration pour la mesure gaussienne (ce n'est pas totalement évident, voir par exemple [3] pour les détails).

Corollaire 3.2.5 (Concentration pour l'espace gaussien) — *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction L-lipschitzienne (par rapport à la distance euclidienne). On note M_f une médiane de f pour γ_n. Alors pour $\varepsilon > 0$*

$$\gamma_n(\{x, f(x) \geqslant M_f + \varepsilon\}) < e^{-t^2/2L^2} \tag{26}$$

Par ailleurs, le même résultat est vrai si on remplace la médiane par l'espérance.

On peut maintenant appliquer ce résultat à la fonction $A_n \to ||A_n||$ définie sur l'EGU (qui s'identifie à l'espace gaussien via la correspondance précédente). L'image de la norme euclidienne par cette identification est, à une constante $\sqrt{n}$ près, la norme de Hilbert-Schmidt $|| \bullet ||_{HS}$. Or on sait que $||A|| - ||B|| \leqslant ||A - B|| \leqslant ||A - B||_{HS}$, ce qui montre que notre fonction $A_n \to ||A_n||$ est $\frac{1}{\sqrt{n}}$-lipschitzienne par rapport à la métrique euclidienne. Le corollaire précédent nous permet donc de conclure (en utilisant le fait que $\mathbb{E}_n ||A_n|| \leqslant 2$)

$$\mathbb{P}_n(\lambda_1(A_n) \geqslant 2 + \varepsilon) < e^{-n\varepsilon^2/2} \tag{27}$$

C'est un premier résultat de concentration pour la plus grande valeur propre.

4 Tracy-Widom et un second résultat de concentration

4.1 Distribution limite de Tracy-Widom

Nous allons commencer par écrire les probabilités qui interviennent comme des déterminants de certains opérateurs, comme c'est fait dans la théorie classique (voir [4]). Fixons d'abord quelques notations : on note (H_n) les polynômes de Hermite, définis par :

$$H_n(t) := (-1)^n e^{t^2} (\frac{d}{dt})^n e^{-t^2} \tag{28}$$

Ces polynômes sont orthogonaux pour la mesure sur $\mathbb{R}$ de densité e^{-x^2} par rapport à la mesure de Lebesgue. On note ensuite

$$\varphi_n(t) := \frac{1}{\sqrt{d_n}} H_n(t) e^{-t^2/2} \tag{29}$$

où $d_n := \int_{\mathbb{R}} H_n(x)^2 dx$. La famille (φ_n) est donc orthonormale dans $L^2(\mathbb{R})$. On pense enfin

$$k_n(x,y) := \sum_{n=0}^{n-1} \varphi_n(x)\varphi_n(y) \tag{30}$$

On peut associer à k_n un opérateur intégral K_n défini sur l'espace de Hilbert $L^2(\mathbb{R})$ par :

$$(K_n f)(x) := \int_{\mathbb{R}} k_n(x,y) f(y) dy \tag{31}$$

L'opérateur K_n n'est autre que la projection orthogonale dans L^2 sur le sous-espace engendré par $(\varphi_n)_{1 \leqslant n \leqslant n}$.

Cette situation est très générale : à chaque fois que l'on a un espace mesuré (X, μ) et un "noyau" $k \in L^2(X \times X)$, on peut lui associer un opérateur K sur $L^2(X)$ par une formule analogue à (31). Dans la suite on notera toujours les noyaux par des lettres minuscules et les opérateurs intégraux correspondants par la majuscule correspondante.

On peut alors exprimer la fonction de répartition de $\lambda_1^{(n)}$ à l'aide du noyau k_n. Plus précisément, on a le résultat suivant (voir par exemple [4]) :

$$\forall\, t \in \mathbb{R}, \mathbb{P}(\sqrt{n}\lambda_1^{(n)} \leqslant t) = \det_{[t,\infty[}(\mathrm{Id} - K_n) \tag{32}$$

Dans la formule (32), la notation $\det_{[t,\infty[}$ doit être comprise comme le déterminant de l'opérateur K_n agissant sur l'espace $L^2([t,\infty[)$, (ou encore de l'opérateur de noyau $k_{n|[t,\infty[^2}$). On notera cette restriction $K_n^{[t]}$.

Ici l'opérateur K_n est de rang fini, et donc la définition du déterminant ne pose pas de problème. Dans la suite le symbole det pourra désigner le déterminant d'une pertubation nucléaire de l'identité, défini alors par

$$\det(\mathrm{Id}+N) := 1 + \sum_{k=1}^{\infty} \mathrm{tr}(\Lambda^k(N)) \tag{33}$$

Tracy et Widom ont démontré que, sous une certaine renormalisation, ces noyaux convergeaient vers un noyau limite, qui fait intervenir la fonction d'Airy. La bonne renormalisation est la suivante

$$\tau_n(x) := \sqrt{2n} + \frac{x}{\sqrt{2}n^{1/6}} \tag{34}$$

notons également $\tilde{k}_n$ le noyau renormalisé :

$$\tilde{k}_n(x,y) := \frac{1}{\sqrt{2}n^{1/6}} k_n(\tau_n(x), \tau_n(y)) \tag{35}$$

Soit k le noyau défini par :

$$k(x,y) := \frac{\mathrm{Ai}(x)\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\mathrm{Ai}(y)}{x-y} \tag{36}$$

Le noyau k est prolongé par continuité sur la diagonale. La fonction Ai est la fonction d'Airy. C'est une solution de l'équation différentielle

$$\frac{\partial^2}{\partial x^2} y(x) = xy(x) \tag{37}$$

Elle tend vers zéro très rapidement en $+\infty$.

Tracy et Widom ont montré que noyau k est la limite des noyaux $(\tilde{K}_n^{[t]})$, dans le sens suivant ([7])

Proposition 4.1.1 (Tracy-Widom) — *Pour tout réel t, la suite d'opérateurs $(\tilde{K}_n^{[t]})$ converge en norme nucléaire vers l'opérateur $K^{[t]}$.*

La fonction $\det(\mathrm{Id}+\bullet)$ étant continue par rapport à la norme nucléaire, on peut en déduire le résultat suivant grâce à la formule (32)

Corollaire 4.1.2 — *On note $\varphi_{TW}(x) := 1 - \det_{[x,+\infty[}(\mathrm{Id}-K)$. Alors*

$$\lim_{n\to\infty} \mathbb{P}_n(\lambda_1(A_n) \geqslant 2 + xn^{-2/3}) = \varphi_{TW}(x) \tag{38}$$

La partie la plus importante des travaux de Tracy et Widom a été d'écrire cette fonction φ_{TW} en termes de solutions d'équations de Painlevé. Remarquons juste que le comportement asymptotique de φ_{TW} est bien connu : pour x assez grand, on a

$$c'e^{-C'x^{3/2}} \leqslant \varphi_{TW}(x) \leqslant ce^{-Cx^{3/2}} \tag{39}$$

4.2 Une version uniforme

Le résultat précédent (le collaire 4.1.2) permet de voir que l'inégalité (27) que nous avions obtenue précédemment par des méthodes de concentration dans l'espace gaussien n'était pas optimale. Rappelons quelle était cette inégalité :

$$\mathbb{P}_n(\lambda_1(A_n) \geqslant 2 + \varepsilon) < e^{-n\varepsilon^2/2} \tag{40}$$

Or, si l'on choisit $\varepsilon = xn^{-2/3}$, le terme de gauche de cette inégalité tend vers $\varphi_{TW}(x)$ (d'après le résultat de Tracy-Widom), qui est très petit, et le terme de droite tend vers 1.

On aimerait donc raffiner cette inégalité pour qu'elle coïncide avec le comportement de Tracy-Widom. Ceci était l'objet de travaux que j'ai effectués en DEA, où je suis parvenu au résultat suivant

Proposition 4.2.1 — *Il existe des constantes universelles c et C telles que, pour tout ε positif*

$$\mathbb{P}_n(\lambda_1(A_n) \geqslant 2 + \varepsilon) \leqslant Ce^{-cn\varepsilon^{3/2}} \tag{41}$$

La preuve de ce théorème suit pas à pas celle de Tracy et Widom, en prenant suffisamment de soin pour garder à chaque étape une estimation uniforme. Elle est expliquée en détail dans l'article "An inequality about the largest eigenvalue of a random matrix".

Généralisations possibles

Il semble y avoir d'autres méthodes pour parvenir à démontrer la proposition 4.2.1. Par exemple, Michel Ledoux m'a récemment signalé qu'il obtenait des résultats très similaires (à un terme polynomial parasite près) par des méthodes d'hypercontractivité. Il n'est pas exclu qu'une analyse plus précise dans cette voie permette de se débarasser de ces termes.

Par ailleurs, Alexander Soshnikov a prouvé que la distribution limite de Tracy-Widom apparaissait dans un cadre beaucoup plus général que celui de l'EGU : celui des matrices de Wigner, où l'on impose juste que les coefficients de la matrice aléatoire suivent des lois indépendantes, avec un contrôle uniforme sur les moments. On peut naturellement se demander si notre analogue "uniforme" reste vrai dans ce cadre.

Références

[1] Y. Benyamini. Two-point symmetrization, the isoperimetric inequality on the sphere and some applications. *Texas functional analysis seminar 1983–1984*, 39–51, *Longhorn Notes*, 1984.

[2] K.R. Davidson and S.J. Szarek. Local operator theory, random matrices and Banach spaces. In *Handbook of the Geometry of Banach Spaces, vol. 1*, pages 317–366, 2001.

[3] M. Ledoux. *The concentration of measure phenomenon.* Mathematical Surveys and Monographs, **89**, 2001.

[4] M.L. Mehta. *Random Matrices.* Academic Press, Boston, MA, second edition, 1991.

[5] L. Pastur. Matrices aléatoires : Statistique asymptotique des valeurs propres. *Notes de cours données au CMI*, 2000.

[6] N. Tomczak-Jaegermann *Banach-Mazur distances and finite-dimensional operator ideals.* Pitman Monographs and Surveys in Pure and Applied Mathematics, **38**, 1989.

[7] C.A. Tracy and H. Widom. Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.*, 159 :151–174, 1994.