Matrices aléatoires et zéros de polynômes: Notes de cours

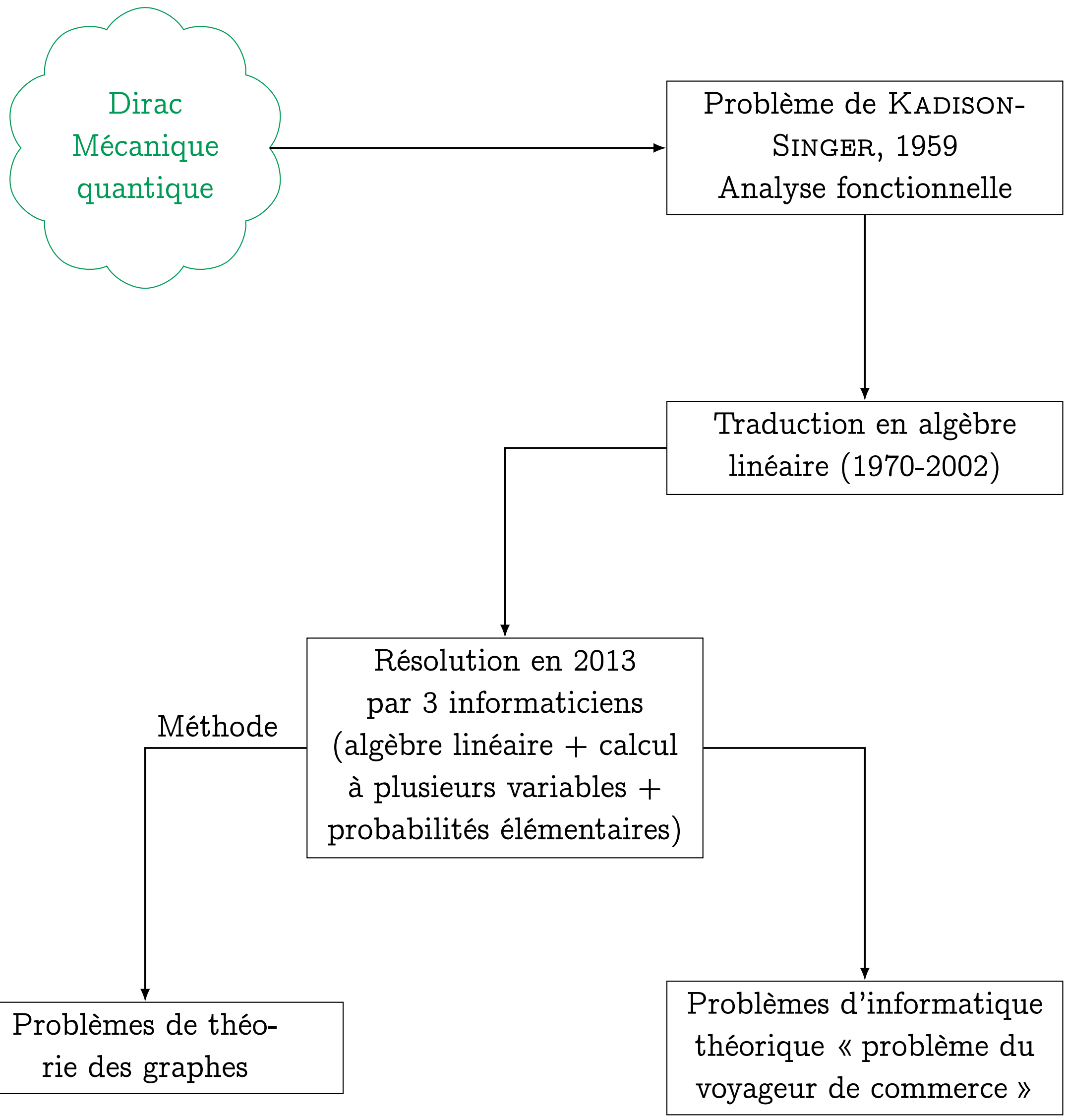
Enseignant: Prof. Alain Valette
Scribe: Laurent Hayez

Année 2018-2019, semestre de printemps
Dernière modification: 4 avril 2019

Table des matières

Chapitre 0

Résumé



Chapitre 1

De la mécanique quantique à l'analyse fonctionnelle

1.1 Mécanique quantique

Paul Adrien Maurice Dirac (1902 - 1984) physicien anglais d'origine valaisanne, a écrit en 1930 les « Principles of quantum mechanics » (plusieurs fois ré-édités).

Principes :

- Les états d'un système physique sont représentés par les vecteurs-unités d'un espace de Hilbert $\mathcal{H}$.

> **Exemples 1.1.** — Une particule libre à une dimension : $\mathcal{H} = L^2(\mathbb{R})$.
> — Une particule libre à deux dimensions : $\mathcal{H} = L^2(\mathbb{R}^2)$.
>
> ★

- Les grandeurs physiques (« observables ») sont des opérateurs auto-adjoints ($T = T^*$) sur $\mathcal{H}$.

> **Exemples 1.2.** — Sur $L^2(\mathbb{R})$, l'**opérateur de position** P est la multiplication par x sur $L^2(\mathbb{R})$.
> — L'**opérateur de moment** (ou impulsion) $Q = \frac{1}{i}\frac{d}{dx}$.
>
> ★

- Ce qu'on peut observer (en laboratoire) est la **probabilité** que la valeur d'une observable sur un état donné, soit comprise entre deux valeurs a et b.

> **Exemples 1.3.** — Soit $\psi \in L^2(\mathbb{R})$, $\|\psi\| = 1$, c'est-à-dire
>
> $$\int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1.$$
>
> La probabilité que la position d'une particule dans l'état ψ, soit entre

a et b est

$$\int_a^b |\psi(x)|^2 dx.$$

En effet, le fait que $\|\psi\| = 1$ nous dit que ψ est une densité de probabilité.

★

Ces trois principes sont parfois appelés les « axiomes de la mécanique quantique ».

Dans le chapitre 18 de son livre, « Probability amplitudes », Dirac donne une recette pour obtenir ces probabilités.

Définition 1.4. Deux observables S et T sont dit **compatibles** si $ST = TS$.

Exemples 1.5.

1. Les observables P et Q précédemment définis ne sont pas compatibles. En effet,

$$PQ - QP = \frac{-1}{i}\mathrm{Id},$$

qui est la **relation d'indétermination de Heisenberg**.

2. Pour une particule libre à deux dimensions, posons P_x l'opérateur de multiplication par la première variable x sur $L^2(\mathbb{R})$ ($P_x f(x,y) = xf(x,y)$) et P_y l'opérateur de multiplication par la deuxième variable y. Ces deux opérateurs sont compatibles,

$$P_x P_y = P_y P_x.$$

★

La recette de Dirac est la suivante :

1. considérer un ensemble maximal d'observables deux à deux compatibles,
2. spécifier les probabilités associées sur les observables dans un état quantique donné,
3. étendre ces probabilités à toutes les observables, même non compatibles.

1.2 C^*-algèbres

Si $\mathcal{H}$ est un espace de Hilbert, on note $\mathcal{B}(\mathcal{H})$ l'espace des opérateurs linéaires bornés (donc continus) de $\mathcal{H}$ vers $\mathcal{H}$, avec la norme opérateur

$$\|T\| = \sup_{\|x\| \leqslant 1} \|Tx\|.$$

L'espace $\mathcal{B}(\mathcal{H})$ est une **algèbre de Banach**

$$\|ST\| \leqslant \|S\|\|T\|.$$

Si $T \in \mathcal{B}(\mathcal{H})$, $T^* \in \mathcal{B}(\mathcal{H})$ est défini par

$$\langle T\xi\,,\,\eta\rangle = \langle \xi\,,\,T^*\eta\rangle \quad \forall \xi,\eta \in \mathcal{H}.$$

Définition 1.6.
- Une **sous-algèbre** est un sous espace vectoriel qui est stable pour la multiplication.
- Une **$*$-sous-algèbre** A a la propriété $T \in A \Rightarrow T^* \in A$ et est fermée pour la norme opérateur.

Définition 1.7. Une **C^*-algèbre** est une $*$-sous-algèbre fermée de $\mathcal{B}(\mathcal{H})$.

Exemples 1.8.
1. $\mathcal{B}(\mathcal{H})$, $\mathbb{C}\mathbb{1} = \{\lambda\mathbb{1},\ \lambda \in \mathbb{C}\}$, $\{0\}$, $\mathcal{K}(\mathcal{H})$ l'ensemble des opérateurs compacts.
2. Soit X un espace topologique compact, et $C(X)$ l'ensemble des fonctions continues de X dans $\mathbb{C}$ muni de la norme

$$\|f\|_\infty = \max_{x\in X} |f(x)|$$

et de l'involution

$$f^*(x) = \overline{f(x)}.$$

On prend sur X une mesure de probabilité μ telle que $\mu(U) > 0$ pour tout ouvert non vide de X. *L'exemple a garder en tête est* $[0,1]$ *avec la mesure de Lebesgue.* Ainsi $\mathcal{H} = L^2(X,\mu)$ est un espace de Hilbert. Si on multiplie une fonction de L^2 par une fonction continue, elle reste dans L^2, ainsi

$$\pi\colon C(X) \to \mathcal{B}(\mathcal{H}),\ \ f \mapsto \text{multiplication par } f \text{ sur } \mathcal{H}.$$

Exercice 1.1. Avec $\|\pi(f)\| = \|f\|_\infty$, $C(X)$ est une C^*-algèbre. ♠

★

Théorème 1.9 (Gelfand, 1940). *Soit* A *une* C^**-algèbre commutative à unité* $(1 \in A)$. *Il existe un espace compact* X*, unique à homéomorphisme près, tel que* $A \simeq C(X)$.

Définition 1.10. Un opérateur $T \in \mathcal{B}(\mathcal{H})$ est **positif** ($T \geqslant 0$) si les conditions équivalentes suivantes sont satisfaites :

1. pour tout $\xi \in \mathcal{H}$, $\langle T\xi\,,\,\xi\rangle \geqslant 0$,
2. il existe $S \in \mathcal{B}(\mathcal{H})$ tel que $T = S^*S$,
3. $T = T^*$ et $\mathrm{Sp}(T) \subset [0, +\infty[$ (Rappel : le spectre de T $\mathrm{Sp}(T) = \{\lambda \in \mathbb{C} : T - \lambda\mathbb{1}$ n'est pas inversible$\}$, $\mathrm{Sp}(T)$ est un compact non vide de $\mathbb{C}$ et $\mathrm{Sp}(T) \subseteq B(0, \|T\|)$).

Définition 1.11. Si A est une C^*-algèbre à unité, un **état** sur A est une forme linéaire $\varphi : A \to \mathbb{C}$ telle que

1. $\varphi(1) = 1$,
2. $\varphi(T^*T) \geqslant 0$ pour tout $T \in A$.

On note $S(A)$ l'ensemble des états sur A.

Exemple 1.12. Soit $\xi \in \mathcal{H}$ tel que $\|\xi\| = 1$, alors pour $T \in \mathcal{B}(\mathcal{H})$

$$\varphi(T) = \langle T\xi\,,\,\xi\rangle$$

est un **état vectoriel** et

$$\varphi(1) = \|\xi\|^2 = 1,\ \ \varphi(T^*T) = \langle T^*T\xi\,,\,\xi\rangle = \langle T\xi\,,\,T\xi\rangle = \|T\xi\|^2 \geqslant 0$$

et ainsi $S(A) \neq \emptyset$. ★

Proposition 1.13. *$S(A)$ est une partie convexe de la boule-unité du dual A^* (ici A^* est l'ensemble des formes linéaires continues sur A).*

Preuve.

1. Soient $\varphi_1, \varphi_2 \in S(A)$, soit $t \in [0, 1]$. On doit montrer que

$$t\varphi_1 + (1-t)\varphi_2 \in S(A).$$

On a

$$(t\varphi_1 + (1-t)\varphi_2)(\mathbb{1}) = 1, \quad (t\varphi_1 + (1-t)\varphi_2)(T^*T) \geqslant 0.$$

2. Si $\varphi \in S(A)$, on doit montrer que $\|\varphi\| \leqslant 1$, c'est-à-dire

$$|\varphi(t)| \leqslant 1 \text{ si } \|T\| \leqslant 1.$$

Si $\|T\| \leqslant 1$, alors $\mathbb{1} - T^*T \geqslant 0$ car

$$\langle (\mathbb{1} - T^*T)\xi\,,\,\xi\rangle = \|\xi\|^2 - \|T\xi\|^2 \geqslant 0$$

puisque $\|T\| \leqslant 1$. On peut encore écrire

$$\mathbb{1} - T^*T = S^*S$$

pour $S \in A$. Ainsi

$$1 - \varphi(T^*T) = \varphi(\mathbb{1} - T^*T) = \varphi(S^*S) \geqslant 0 \implies \varphi(T^*T) \leqslant 1.$$

L'application $A \times A \to \mathbb{C}$, $(x, y) \mapsto \varphi(y^*x)$ vérifie l'inégalité de Cauchy-Schwartz

$$|\varphi(y^*x)|^2 \leqslant \varphi(y^*y)\varphi(x^*x)\,\forall x, y \in A.$$

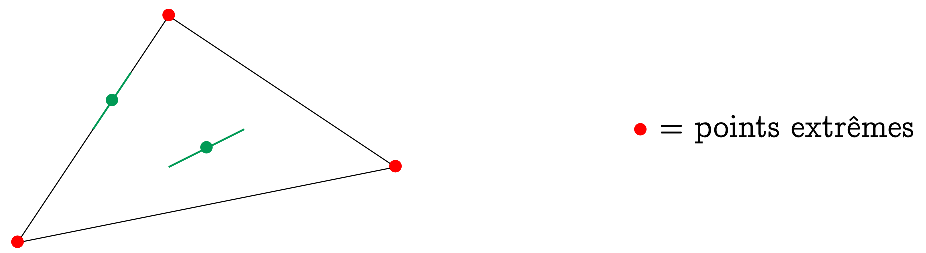
Pour $x = T$, $y = \mathbb{1}$,

$$|\varphi(T)|^2 \leqslant \varphi(\mathbb{1})\varphi(T^*T) \leqslant 1.$$

□

Remarque 1.14. Comme $1 = \varphi(\mathbb{1}) \leqslant \|\varphi\|$, on a que $S(A)$ est contenu dans la sphère unité de A^*. ♣

Définition 1.15. Soit K un convexe dans un espace vectoriel réel ou complexe. Un point K est **extrême** dans K si x n'est pas intérieur à un segment contenu dans K, c'est-à-dire si $x = tx_1 + (1-t)x_2$ avec $0 < t < 1$, $x_1, x_2 \in K$, alors $x = x_1 = x_2$.



Un point extrême de $S(A)$ est appelé un **état pur**.

Exemples 1.16.
1. On montre que tout état vectoriel de $\mathcal{B}(\mathcal{H})$ est pur. On verra qu'il y a des états purs non vectoriels.
2. Soit $A = C(X)$, alors $S(A)$ s'identifie avec $\text{Prob}(X)$ l'espace des mesures de probabilités sur X (un des théorèmes de représentation de Riesz). ★

Les états purs s'identifient aux points extrêmes de $\mathrm{Prob}(X)$, c'est-à-dire aux mesures de Dirac δ_x, définie par (pour $A \subset X$)

$$\delta_x(A) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{sinon.} \end{cases}$$

Théorème 1.17 (Krein-Milman). *Soit K un convexe compact non vide dans un espace vectoriel topologique sur $\mathbb{R}$ ou $\mathbb{C}$. Alors K est l'enveloppe convexe fermée de ses points extrêmes*

$$K = \overline{\mathrm{conv}(\mathrm{ext}(K))}, \quad \mathrm{ext}K = \{\textit{points extrêmes de } K\}.$$

Sur A^*, on a la topologie faible $*$: $\varphi_i \xrightarrow{i\to\infty} \varphi$ si pour tout $x \in A$, $\varphi_i(x) = \varphi(x)$.

Théorème 1.18 (Banach-Alaoglu). *La boule unité de A^* est compacte pour la topologie faible $*$.*

On a vu que $S(A)$ est contenu dans la boule unité de A^*,

$$S(A) = \{\varphi \in A^* \text{ tq } \varphi(\mathbb{1}) = 1,\ \varphi(T^*T) \geqslant 0\ \forall T \in A\},$$

donc $S(A)$ est fermé pour la topologie faible $*$, ainsi $S(A)$ est convexe et compact. Par Krein-Milman, $S(A) = \overline{\mathrm{conv}(\mathrm{ext}(S(A)))}$.

Si $\varphi \in S(A)$, une version du théorème de Hahn-Banach dit que φ s'étend en au moins un état sur $\mathcal{B}(\mathcal{H}) \supset A$.

Soit $K_\varphi = \{\psi \in S(\mathcal{B}(\mathcal{H}))\colon \psi\big|_A = \varphi\}$ l'ensemble des états de $\mathcal{B}(\mathcal{H})$ qui étendent φ. Alors K_φ est convexe et faible $*$-fermé.

cf. pages d'Alain pour la suite du cours du 28 février

Problème : un état pur sur une MASA a-t-il une extension unique à $\mathcal{B}(\mathcal{H})$?

Le résultat principal de Kadison & Singer (1959) est le théorème suivant, qu'on admet.

Théorème 1.19. *Pour la MASA diffuse $L^\infty[0,1]$, la réponse est non.*

L'article laisse ouvert le cas de la MASA discrète $\ell^\infty(\mathbb{N})$.

Problème de Kadison&Singer : la MASA $\ell^\infty(\mathbb{N})$ a-t-elle l'extension unique des états purs ?

Kadison & Singer inclinent à penser que la réponse doit être non (en fait c'est oui).

Proposition 1.20. *Pour* $k \in \mathbb{N}$*, notons* φ_k *l'état pur de* $\ell^\infty(\mathbb{N})$ *donné par* $\varphi_k(\underbrace{(a_n)_{n>0}}_{\in \ell^\infty(\mathbb{N})}) = a_k$*. Alors l'extension unique de* φ_k *à* $\mathcal{B}(\ell^2(\mathbb{N}))$ *est*

$$T \mapsto \langle Te_k \,,\, e_k\rangle = T_{kk}.$$

Preuve. 1. Si ψ est un état pur non vectoriel de $\mathcal{B}(\mathcal{H})$, alors ψ est nul sur l'espace des opérateurs compacts (en particulier sur l'espace des opérateurs de rang fini). Ici,

$$p_k = (0, \dots, 0, \underbrace{1}_{\text{k-ème}}, 0, 0, \dots) \in \ell^\infty(\mathbb{N})$$

et ainsi on a $\varphi_k(p_k) = 1$. Donc une extension de φ_k à $\mathcal{B}(\ell^2(\mathbb{N}))$ est un état vectoriel, disons

$$T \mapsto \langle T\xi \,,\, \xi\rangle \,, \ \|\xi\| = 1.$$

2. À voir : $\xi = \lambda e_k$ avec $|\lambda| = 1$ (où e_k est p_k mais vu dans $\ell^2(\mathbb{N})$ (on le voit une fois comme matrice diagonale, et une fois comme suite)). Si $T = (a_n)_{n>0}$,

$$\langle T\xi \,,\, \xi\rangle = \sum_{n=1}^{\infty} a_n |\xi_n|^2.$$

Mais $T\xi = \varphi_k(T) = a_k$. Ainsi,

$$\sum_{n \neq k} a_n |\xi_n|^2 + a_k(|\xi_k|^2 - 1) = 0 \ \forall (a_n)_{n>0} \in \ell^\infty(\mathbb{N}).$$

Donc $\xi_n = 0$ si $n \neq k$; $|\xi_k| = 1$ et donc $\xi = \lambda e_k$, $|\lambda| = 1$.

□

Définition 1.21. L'application diagonale

$$\mathrm{diag} \colon \mathcal{B}(\ell^2(\mathbb{N})) \longrightarrow \ell^\infty(\mathbb{N}), \ T \mapsto (T_{nn})_{n>0}$$

Intuition : on a un tableau carré, et on garde seulement la diagonale.

Exemples 1.22.

- $\mathrm{diag}(\mathbb{1}) = (1, 1, \dots)$
- Si $T \geqslant 0$, $\mathrm{diag}(T) = \left(\underbrace{\langle Te_n \,,\, e_n\rangle}_{\geqslant 0}\right)_{n>0}$.

Si φ est un état sur $\ell^\infty(\mathbb{N})$, alors $\varphi \circ \mathrm{diag}$ est un état sur $\mathcal{B}(\ell^2(\mathbb{N}))$ qui étend φ. Ceci veut dire que pour étendre φ à $\mathcal{B}(\ell^2(\mathbb{N}))$, on n'a pas besoin du théorème de Hahn-Banach. Grâce à cette observation, le problème de Kadison-Singer se reformule de la façon suivante.

Problème de KS : Si φ est un état pur de $\ell^\infty(\mathbb{N})$, alors $\varphi \circ \mathrm{diag}$ est l'unique extension de φ à $\mathcal{B}(\ell^2(\mathbb{N}))$.

Chapitre 2

De l'analyse fonctionnelle à l'algèbre linéaire

2.1 Conjecture de pavage

En 1979, le mathématicien américain Joel ANDERSON propose la **conjecture de pavage**

Conjecture 2.1. *Pour tout* $\varepsilon > 0$*, il existe* $r \in \mathbb{N}$ *tel que pour tout* $T \in \mathcal{B}(\ell^2(N))$ *(ou* $M_m(\mathbb{C})$*) avec* $\mathrm{diag}(T) = 0$*, il existe* $Q_1, \dots, Q_r$ *des projecteurs diagonaux avec*

$$\sum_{i=1}^{r} Q_i = \mathbb{1}$$

et

$$\|Q_i T Q_i\| \leqslant \varepsilon \|T\|, \ \forall i \in \{1, \dots, r\}.$$

Pour rappel, un projecteur P est tel que

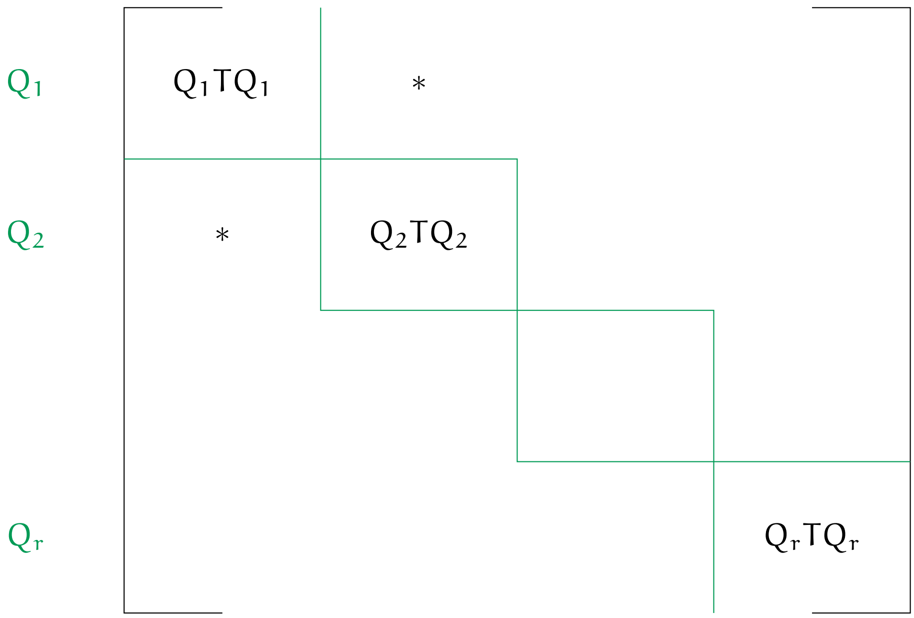
$$P = P^2 = P^*.$$

C'est donc une projection orthogonale sur un certain sous-espace fermé de $\mathcal{H}$. Un projecteur diagonal a pour coefficients $a_{ii} \in 0, 1$ et $a_{ij} = 0$ pour $i \neq j$.

Il y a une bijection entre les projecteurs diagonaux et les parties de $\mathbb{N}$. Ainsi

$$\sum_{i=1}^{r} Q_i = \mathbb{1}$$

signifie qu'on partitionne $\mathbb{N}$ en r parties. Les $Q_1, \dots, Q_r$ donnent une décomposition de $\ell^2(\mathbb{N})$ en r blocs



La conjecture dit que la condition $\operatorname{diag} T = 0$ implique que les blocs sont de norme petite.

Exemple 2.2. Soit $S : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ l'opérateur de décalage unilatéral (« unilateral shift ») défini par $Se_n = e_{n+1}$ ou $S(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots)$. Alors

$$S = \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}$$

On prend la partition de $\mathbb{N} = (2\mathbb{N}) \cup (2\mathbb{N}+1)$. Soit Q_1 la projection sur $\ell^2(2\mathbb{N})$ et Q_2 la projection sur $\ell^2(2\mathbb{N}+1)$. Alors

$$S = \begin{pmatrix} 0 = Q_1 S q_1 & * \\ * & 0 = Q_2 S Q_2 \end{pmatrix}$$

L'opérateur S vérifie la conjecture de pavage (« tiling conjecture ») avec $\varepsilon = 0$ et $r = 2$. ★

Proposition 2.3. *Si la conjecture de pavage est vraie, alors $\ell^\infty(\mathbb{N})$ a l'extension unique des états purs.*

Preuve. Soit φ un état pur de $\ell^\infty(\mathbb{N})$, ψ une extension de φ à $\mathcal{B}(\ell^2(\mathbb{N}))$. On veut montrer que
$$\psi = \varphi \circ \mathrm{diag}.$$
C'est-à-dire, pour tout $T \in \mathcal{B}(\ell^2(N))$,
$$\psi(T) = \varphi(\mathrm{diag}(T)).$$
En remplaçant T par $T - \mathrm{diag}(T)$, on obtient un opérateur de diagonale nulle. On doit donc montrer que si $\mathrm{diag} T = 0$, alors $\psi(T) = 0$, qui est ce qu'on va démontrer.

On montre que pour tout $\varepsilon > 0$, $|\psi(T)| \leqslant \varepsilon\|T\|$. Par la conjecture de pavage, on trouve $r \in \mathbb{N}$ et des projecteurs diagonaux $Q_1, \dots, Q_r$ tels que
$$\sum_{i=1}^{r} Q_i = \mathbb{1}$$
et
$$\|Q_i T Q_i\| < \varepsilon\|T\|.$$
On utilise à présent le fait que φ est un état pur : φ est multiplicatif sur $\ell^\infty(\mathbb{N})$ (c'est-à-dire que $\varphi(ST) = \varphi(S)\varphi(T)$ si $S, T \in \ell^\infty(\mathbb{N})$). La raison est que $\ell^\infty(\mathbb{N}) = C(X)$ l'espace des fonctions continues sur X car ℓ^∞ est une C^*-algèbre commutative à unité (c'est le théorème de Gelfand). Les états purs de $C(X$ sont les évaluations aux points, elles sont multiplicatives (ici $X = B\mathbb{N}$ est le compactifié de Stone-Čech de $\mathbb{N}$, c'est la plus grosse compactification de $\mathbb{N}$). Alors
$$\varphi(Q_i) = \varphi(Q_i^2) = \varphi(Q_i)^2 \implies \varphi(Q_i) \in \{0, 1\}.$$
De plus,
$$1 = \varphi(\mathbb{1}) = \varphi\left(\sum_{i=1}^{r} Q_i\right) = \sum_{i=1}^{r} \varphi(Q_i).$$
Puisque $\varphi(Q_i) \in \{0, 1\}$ et qu'on écrit 1 comme somme d'éléments de $\{0, 1\}$, il existe un unique indice i_0 avec $\varphi(Q_{i_0}) = 1$ et $\varphi(Q_i) = 0$ si $i \neq i_0$.

Alors
$$\psi(T) = \psi\left(\left(\sum_{i=1}^{r} Q_i\right) T \left(\sum_{j=1}^{r} Q_j\right)\right) = \sum_{i,j=1}^{r} \psi(Q_i T Q_j).$$
Si on sait que dans ces r^2 termes, le seul terme non nul est $\psi(Q_{i_0} T Q_{i_0})$, alors
$$|\psi(T)| = |\psi(Q_{i_0} T Q_{i_0})| \leqslant \|Q_{i_0} T Q_{i_0}\| \leqslant \varepsilon\|T\|$$
comme on voulait, où la dernière inégalité suit de la conjecture de pavage, et la première du fait que ψ est un état pur, donc de norme 1 (note : à vérifier).

Par Cauchy-Schwartz,

$$
\begin{aligned}
|\psi(Q_i T Q_j)| &= |\psi\left((T^* Q_i)^* Q_j\right) \\
&\leqslant \psi\left((T^* Q_i)^* (T^* Q_i)\right)^{1/2} \underbrace{\psi(\underbrace{Q_j^* Q_j}_{=Q_j})}_{\varphi(Q_j)}{}^{1/2} \\
&= 0 \text{ si } j \neq i_0.
\end{aligned}
$$

De même,

$$
|\psi(Q_i T Q_j)| = 0
$$

si $i \neq i_0$ par un argument similaire. Le seul terme non nul restant est donc $\psi(Q_{i_0} T Q_{i_0})$, comme on le souhaitait. $\square$

2.2 Conjecture de Weaver (2004)

(Approche de Terry Tao, 2013)

Conjecture 2.4 (conjecture de Weaver). *On fixe des entiers* $d, m, r \geqslant 2$ *et une constante* $c > 0$. *Soient* $A_1, \ldots, A_d \in M_m(\mathbb{C})$, $A_i \geqslant 0$, $\operatorname{rang}(A_i) = 1$ *avec* $\|A_i\| \leqslant C$ *pour tout* $i = 1, \ldots d$ *et*

$$
\sum_{i=1}^{d} A_i = \mathbb{1}_m.
$$

Alors il existe une partition $\{S_1, S_2, \ldots, S_r\}$ *de* $\{1, \ldots, d\}$ *telle que*

$$
\left\| \sum_{i \in S_j} A_i \right\| \leqslant \left(\sqrt{\frac{1}{r}} + \sqrt{C} \right)^2
$$

pour $j = 1, \ldots, r$.

Pour cette conjecture, il faut penser à d et m grands, et puisque les matrices sont de rang 1 et qu'on veut avoir leur somme égale à $\mathbb{1}$, on a besoin d'au moins m telle matrices, i.e., $d \geqslant m$. De plus il faut penser à r petit (cas extrême, $r = 2$).

Si $A \geqslant 0$ et $\operatorname{rang}(A) = 1$ alors A est un multiple positif d'un projecteur orthogonal de rang 1. C'est-à-dire qu'il existe $\xi \in \mathbb{C}^m$ tel que $\|\xi\| = 1$ et $\lambda > 0$ tel que

$$
A(v) = \lambda \langle v\,,\, \xi \rangle\, \xi.
$$

Si les A_i ont des images 2 à 2 orthogonales (sans supposer que leur somme vaut $\mathbb{1}$

comme dans la conjecture), alors la somme se décompose par blocs,

$$\sum A_i = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_k \end{pmatrix} \implies \left\| \sum A_i \right\| = \max \|A_i\| \leqslant C.$$

Ainsi, ce que veut dire la conjecture, c'est que ces matrices sont *presque* orthogonales. En d'autres termes, on peut partitionner l'ensemble des indices de façon à ce que pour chaque classe de la partition, les A_i soient d'images *quasiment* orthogonales.

Proposition 2.5. *La conjecture de* WEAVER *implique la conjecture de pavage.*

Manque le cours du 14 mars

2.3 Matrices aléatoires

Référence classique : M.L. MEHTA, Random matrices, Academic Press 1991.

Définition 2.6. Une **matrice aléatoire** est un tableau carré

$$\begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{pmatrix}$$

où les X_{ij} sont des variables aléatoires indépendantes à valeurs complexes.

Le polynôme caractéristique

$$p_X(z) := \det(z\mathbb{1}_n - X)$$

est un polynôme aléatoire et les valeurs propres sont des variables aléatoires au sens ordinaire.

Typiquement, la théorie s'intéresse au comportement des valeurs propres pour $n \to \infty$.

Exemple 2.7. Un modèle très simple : on choisit au hasard un nombre dans $\{1, \ldots, n\}$ (avec probabilité $1/n$), si le résultat est i, on choisit le projecteur sur le i-ème vecteur de la base canonique de $\mathbb{C}^n$. ★

Historique : Eugene Wigner (1902 - 1995), physicien hongrois naturalisé américain, prix Nobel en physique en 1963 pour ses travaux sur théoriques sur la structure des noyaux atomiques. À la fin des années 1940, il propose l'hypothèse : pour modéliser un gros noyau d'uranium (par exemple U^{239}) au niveau quantique, il faut s'intéresser au spectre d'une grosse matrice aléatoire X, auto-adjointe (i.e. $\overline{X_{ij}} = X_{ji}$) où les X_{ij} sont des variables gaussiennes iid.

Théorèmes de MSS : voir photocopie.

Si A est une matrice, on définit l'espérance de A par

$$\mathbb{E}[A] = \sum_i p_i X_i$$

où les X_i sont les valeurs possibles de A. Pour une loi continue, on aurait $(X, \mathcal{B}, \mu) \xrightarrow{A} M_m(\mathbb{C})$, on a

$$\mathbb{E}[A] = \int_X A(\omega) d\mu(\omega) \in M_m(\mathbb{C}).$$

Preuve (preuve fausse du théorème 1, mais inspirante). On a

$$(\mathbb{E}[p_A])(z) = \mathbb{E}(\det(z\mathbb{1}_m - A)) = \det(z\mathbb{1}_m - \mathbb{E}[A]) = \det((z-1)\mathbb{1}_m) = (z-1)^m.$$

□

On a $\|A\|$ est la plus grande valeur propre de A (car $A \geqslant 0$) est aussi la plus grande racine de p_A. Le théorème 2 se reformule de la façon suivante.

Théorème 2.8. *Pour au moins une réalisation des A_i, la plus grande racine de p_A est inférieure à la plus grande racine de $\mathbb{E}[p_A]$.*

C'est une version non linéaire d'un principe de probabilité : si X est une variable aléatoire à valeurs réelles, alors pour au moins une réalisation, $X \leqslant \mathbb{E}[X]$.

Proposition 2.9. *Les théorèmes 1 et 2 de MSS impliquent la conjecture de* Weaver *2.4.*

Preuve. Soient $A_1, \dots, A_d \in M_m(\mathbb{C})$ telles que

$$\sum_{i=1}^{d} A_i = \mathbb{1}_m,$$

et $A_i \geqslant 0$ de rang 1, avec $\|A_i\| \leqslant C$.

Pour $i = 1, \dots, d$, on définit une variable aléatoire $\tilde{A_i}$ à valeurs dans $M_{mr}(\mathbb{C})$, on écrit

$$\mathbb{C}^{mr} = \mathbb{C} \oplus \mathbb{C}^m \oplus \cdots \oplus \mathbb{C}^m.$$

On choisit $j \in \{1, 2, \ldots, r\}$ avec probabilité $\frac{1}{r}$ et on place rA_i dans le j-ème facteur, i.e. le bloc jj de la matrice $\tilde{A_i}$ vaut rA_i, et les autres blocs valent 0. $\tilde{A_i}$ est une variable aléatoire à valeurs dans les opérateurs positifs de rang 1 dans $M_{mr}(\mathbb{C})$. On pose

$$\tilde{A} = \sum_{i=1}^{d} \tilde{A_i}.$$

On veut appliquer les théorèmes 1 et 2 de MSS à $\tilde{A}$. Vérifions l'hypothèse du théorème 1. On a

$$
\begin{aligned}
\mathbb{E}[\tilde{A}] &= \sum_{i=1}^{d} \mathbb{E}[\tilde{A_i}] \\
&= \sum_{i=1}^{d} \frac{1}{r} \sum_{j=1}^{r} \begin{pmatrix} 0 &&&&&& \\ & \ddots &&&&& \\ && 0 &&&& \\ &&& rA_i &&& \\ &&&& 0 && \\ &&&&& \ddots & \\ &&&&&& 0 \end{pmatrix} \\
&= \sum_{j=1}^{r} \sum_{i=1}^{d} \begin{pmatrix} 0 &&&&&& \\ & \ddots &&&&& \\ && 0 &&&& \\ &&& A_i &&& \\ &&&& 0 && \\ &&&&& \ddots & \\ &&&&&& 0 \end{pmatrix} \\
&= \sum_{j=1}^{r} \begin{pmatrix} 0 &&&&&& \\ & \ddots &&&&& \\ && 0 &&&& \\ &&& \mathbb{1}_m &&& \\ &&&& 0 && \\ &&&&& \ddots & \\ &&&&&& 0 \end{pmatrix} \\
&= \begin{pmatrix} \mathbb{1}_m &&&&&& \\ & \ddots &&&&& \\ && \mathbb{1}_m &&&& \\ &&& \mathbb{1}_m &&& \\ &&&& \mathbb{1}_m && \\ &&&&& \ddots & \\ &&&&&& \mathbb{1}_m \end{pmatrix} \\
&= \mathbb{1}_{mr}.
\end{aligned}
$$

De plus, $\|\tilde{A_i}\| = r\|A_i\| \leqslant rC$. Par les théorèmes 1 et 2, pour au moins une réalisation des $\tilde{A_i}$, on a

$$\|\tilde{A}\| \leqslant \text{plus grande racine de } \mathbb{E}[p_{\tilde{A}}] \leqslant (1+\sqrt{rC})^2.$$

La partition $\{S_1, \ldots, S_r\}$ de $\{1, 2, \ldots, d\}$ est associée à cette réalisation,

$$S_j = \{i \in \{1, 2, \ldots, d\} \cdot j \text{ a été choisi au } i\text{-ème essai}\}.$$

Alors

$$r \left\| \sum_{i \in S_j} A_i \right\| = \left\| \sum_{i \in S_j} \tilde{A_j} \right\| \leqslant \left\| \sum_{i=1}^{d} \tilde{A_i} \right\| = \|\tilde{A}\| \leqslant (1 + \sqrt{rc})^2.$$

On divise par r, et on obtient

$$\left\| \sum_{i \in S_j} A_i \right\| \leqslant \left(\sqrt{\frac{1}{r}} + \sqrt{C} \right)^2.$$

□

Chapitre 3

Preuves des théorèmes 1 et 2

3.1 Polynômes réels stables

Notons $\mathbb{H} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ le demi-plan supérieur de $\mathbb{C}$.

Définition 3.1. Un polynômes en d variables $p(z_1, \ldots, z_d)$ est **réel stable** si

- ses coefficients sont réels ;
- p n'admet aucun zéro dans $\mathbb{H}^d$.

Exemple 3.2. Pour $d = 1$, $p(z)$ est réel stables si et seulement si ses coefficients sont réels et tous les zéros de p sont réels. ★

Proposition 3.3. *Soient* $A_1, \ldots, A_d \geqslant 0$ *dans* $M_m(\mathbb{C})$. *Alors*

$$q(z, z_1, \ldots, z_d) := \det\left(z\mathbb{1}_m + \sum_{i=1}^{d} z_i A_i\right)$$

est réel stable.

Preuve. On a

$$q(\overline{z}, \overline{z_1}, \ldots, \overline{z_d}) = \overline{q(z, z_1, \ldots, z_d)}$$

car $A_i = A_i^*$ (i.e. les matrices sont auto-adjointes), donc les coefficients de q sont réels.

Si $q(z, z_1, \ldots, z_d) = 0$, alors $\det\left(z\mathbb{1}_m + \sum_{i=1}^{d} z_i A_i\right) = 0$. Ainsi la matrice $z\mathbb{1}_m + \sum_{i=1}^{d} z_i A_i$ n'est pas inversible, et donc le noyau n'est pas vide, i.e. il existe $v \in \mathbb{C}^m$ non nul avec $zv + \sum_{i=1}^{d} z_i A_i v = 0$. On fait le produit scalaire avec v,

$$z \underbrace{\|v\|^2}_{>0} + \sum_{i=1}^{d} z_i \underbrace{\langle A_i v\,,\, v\rangle}_{\geqslant 0} = 0$$

Donc on ne peut pas avoir $\mathrm{Im} z > 0$ et $\mathrm{Im} z_i > 0$ pour tout i, donc q ne s'annule pas dans $\mathbb{H}^{d+1}$. □

Une **spécialisation de** p est un polynôme en d variables obtenu en donnant une valeur fixe à une des variables.

Attention : si $p(z_1, z_2) = z_1(z_2 - 1)$, alors $p(z, 1) \equiv 0$.

Proposition 3.4. *Si $p(z_1, \ldots, z_d)$ est réel stable, en spécialisant z_d en une valeur réelle t, on obtient un polynôme réel stable, ou le polynôme 0.*

Preuve. Soit $q(z_1, \ldots, z_{d-1}) := p(z_1, \ldots, z_{d-1}, t)$. Les coefficients de q sont réels. On écrit

$$q(z_1, \ldots, z_{d-1}) = \lim_{n\to\infty} p(z_1, \ldots, z_{d-1}, t + i/n)$$

uniformément sur tout compact de $\mathbb{C}^{d-1}$. Le théorème de Hurwitz de l'analyse complexe s'applique. En effet, $p(z_1, \ldots, z_{d-1}, t + i/n)$ est sans zéro dans $\mathbb{H}^{d-1}$. Donc la limite est sans zéro dans $\mathbb{H}^{d-1}$ ou identiquement nul. □

On note $\partial_i = \frac{\partial}{\partial z_i}$ la dérivée partielle par rapport à z_i.

Proposition 3.5. *Si $p(z_1, \ldots, z_d)$ est réel stable, alors pour tout $t \in \mathbb{R}$,*

$$(1 + t\partial_d)p = p + t\frac{\partial p}{\partial z_d}$$

est réel stable.

Preuve. Si $t = 0$, c'est bon. Supposons $t \neq 0$. On va procéder par l'absurde.

$$((1 + t\partial_d)p)(z_1, \ldots, z_d) = 0 \text{ et } (z_1, \ldots, z_d) \in \mathbb{H}^d.$$

Soit

$$q(z) := p(z_1, \ldots, z_{d-1}, z).$$

Alors q n'a pas de zéro dans $\mathbb{H}$ (sinon p ne serait pas réel stable). En particulier,

$$q(z_d) \neq 0.$$

Si $n = \deg q$, alors

$$q(z) = \prod_{i=1}^{n} (z - \omega_i) \qquad (*)$$

où les ω_i sont les zéros complexes de q. Par ce qui précède, $\mathrm{Im}\,\omega_i < 0$ pour tout i.

Alors,

$$0 = ((1+t\partial_d)p)(z_1,\dots,z_d) = (q+tq')(z_d) = q(z_d)\left(1+t\frac{q'(z_d)}{q(z_d)}\right)$$

où le terme $\frac{q'(z_d)}{q(z_d)}$ est la dérivée logarithmique de q en z_d. Ainsi on obtient

$$0 = 1 + t\frac{q'(z_d)}{q(z_d)}.$$

On prend à présent la dérivée logarithmique de $(*)$. Alors

$$\frac{q'(z_d)}{q(z_d)} = \sum_{i=1}^{n} \frac{1}{z-\omega_i}.$$

On déduit de ceci que (en évaluant en z_d)

$$0 = 1 + t\sum_{i=1}^{n}\frac{1}{z_d-\omega_i} = 1 + t\sum_{i=1}^{n}\frac{\overline{z_d}-\overline{\omega_i}}{|z_d-\omega_i|^2}.$$

On prend à présent les parties imaginaires :

$$0 = t\sum_{i=1}^{n}\frac{\mathrm{Im}(\omega_i)-\mathrm{Im}(z_d)}{|z_d-\omega_i|^2}$$

et $\mathrm{Im}(\omega_i) < 0$ et $\mathrm{Im}(z_d) > 0$, qui est une contradiction car $t \neq 0$ par hypothèse et la somme est strictement négative (donc non nulle). □

Pour $x = (x_1,\dots,x_d) \in \mathbb{R}^d$, l'**orthant** $\{y \geqslant x\}$ est par définition

$$\{y \geqslant d\} := \left\{(y_1,\dots,y_d) \in \mathbb{R}^d \colon y_i \geqslant x_i \ \forall i = 1,\dots,d\right\}$$

(on y pensera comme le quart de plan (en deux dimensions) où les coordonnées de y sont toutes plus grandes que celles de x, ou comme le huitième de plan en trois dimensions).

Pour une fonction $f\colon \mathbb{R}^d \to \mathbb{C}$ de classe $\mathcal{C}^1$, on note $\Phi_f^j = \frac{\partial_j f}{f}$ la dérivée logarithmique de f par rapport à la j-ème variable.

Lemme 3.6 (Lemme 3.7 des notes). *Soit $x = (x_1,\dots,x_d) \in \mathbb{R}^d$, $p(z_1,\dots,z_d)$ réel stable, sans zéro dans l'orthant $\{y \geqslant x\}$. Supposons qu'il existe $j \in \{1,\dots,d\}$ tel qu'il existe $\delta > 0$ (penser à δ grand) avec $\Phi_p^j(x_1,\dots,x_d) \leqslant 1 - \frac{1}{\delta}$. Alors $(1-\partial_j)p$ n'a pas de zéro dans l'orthant $\{y \geqslant x + \delta e_j\}$, et de plus, pour tout $i = 1,\dots,d$,*

$$\Phi^i_{(1-\partial_j)p}(x+\delta e_j) \leqslant \Phi^i_p(x). \tag{3.1}$$

Proposition 3.7. *Soit* $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, $p(z_1, \dots, z_d)$ *un polynôme réel stable sans zéro dans l'orthant* $\{y \geqslant x\}$. *S'il existe* $\delta > 0$ *tel que*

$$\Phi_p^j(x_1, \dots x_d) \leqslant 1 - \frac{1}{\delta} \ \forall j = 1, \dots, d,$$

alors le polynôme

$$\prod_{i=1}^{d}(1 - \partial_i)p$$

est sans zéro dans $\{y \geqslant x + D\}$ *où* $D = (\delta, \delta, \dots, \delta)$.

Preuve. Pour $1 \leqslant k \leqslant d$, soit $x^{(k)} = (x_1 + \delta, x_2 + \delta, \dots, x_k + \delta, x_{k+1}, \dots, x_d)$ et

$$q_k = \prod_{i=1}^{k} (1 - \partial_i)\, p$$

qui est réel stable par la Proposition 3.5. Par récurrence à partir du Lemme 3.6, q_k n'a pas de zéro dans l'orthant $\{y \geqslant x^{(k)}\}$ et

$$\Phi_{q_k}^j(x^{(k)}) \leqslant 1 - \frac{1}{\delta}$$

pour $j = 1, \dots, d$. Pour $k = d$, on obtient la proposition. □

3.2 Polynômes caractéristiques mixtes

Rappel : pour $A \in M_m(\mathbb{C})$, le polynôme caractéristique est

$$p_A(z) = \det(z\mathbb{1}_m - A).$$

Définition 3.8. Pour $A_1, \dots, A_d \in M_m(\mathbb{C})$, le **polynôme caractéristique mixte** est

$$\mu[A_1, \dots, A_d](z) = \left(\prod_{j=1}^{d} (1 - \partial_j) \det \left(z\mathbb{1}_m + \sum_{i=1}^{d} z_i A_i \right) \right) \Bigg|_{z_1 = z_2 = \dots = z_d = 0}.$$

Remarque 3.9. Si $A_1, \dots, A_d \geqslant 0$, par les Propositions 3.3, 3.4 et 3.5, $\mu[A_1, \dots, A_d]$ est réel stable. ♣

Proposition 3.10. *Si* $\operatorname{rang}(A_i) = 1$ *pour* $i = 1, \ldots, d$ *avec* $A = \sum_{i=1}^{f} A_i$*, alors*

$$p_A(z) = \mu[A_1, \ldots, A_d](z).$$

Preuve. La preuve se fait en deux pas.

1. Pour tout $B \in M_m(\mathbb{C})$, le polynôme $(z_1, \ldots, z_d) \mapsto \det\left(B + \sum_{i=1}^{d} z_i A_i\right)$ est affine-multilinéaire, c'est-à-dire un exposant $\geqslant 2$ n'apparaît dans aucun terme. Ainsi chaque terme est de la forme

$$C z_1^{\varepsilon_1} z_2^{\varepsilon_2} \cdots z_d^{\varepsilon_d}$$

   avec $\varepsilon_i \in \{0, 1\}$. Ou encore, en chaque variable, on a un polynôme de degré $\leqslant 1$.

   Voyons-le pour $d = 1$ (puis récurrence facile). Le fait que $\operatorname{rang}(A_1) = 1$ implique que $\dim \operatorname{Im} A_1 = 1$. On prend une base de $\mathbb{C}^m$ dont le premier vecteur est dans $\operatorname{Im} A_1$. Dans cette base,

$$A_1 = \begin{pmatrix} * & * & \cdots & * \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

   En développant $\det(B + z_1 + A_1)$ par rapport à la première ligne, on obtient un polynôme de degré inférieur ou égal à 1 en z_1 (en effet, B n'introduit que des coefficients constants dans la matrice, les z_i étant seulement dans la première ligne, on a la conclusion).

2. Formule de Taylor à d variables. Un polynôme affine-multilinéaire est égal à son développement de Taylor d'ordre $(1, 1, \ldots, 1)$. On a

$$\det\left(B + \sum_{i=1}^{d} t_i A_i\right) = \left(\prod_{i=1}^{d} (1 + t_i \partial_i) \det\left(B + \sum_{i=1}^{d} z_i A_i\right)\right)\Bigg|_{z_1 = \cdots = z_d = 0}$$

   pour tout $t_1, \ldots t_d \in \mathbb{R}$. On fait $t_1 = \cdots = t_d = -1$, $B = z\mathbb{1}_m$ de sorte que

$$\det\left(z\mathbb{1}_m - \sum_{i=1}^{d} A_i\right) = p_A(z) = \mu[A_1, \ldots, A_d](z).$$

□

Exemple 3.11. Soit $p(z_1, z_2) = a_{00} + a_{10} z_1 + a_{01} z_2 + a_{11} z_1 z_2$. Alors

$$((1 + t_1 \partial_1) p)(z_1, z_2) = p(z_1, z_2) + t_1 (a_{10} + a_{11} z_2),$$

$$((1 + t_2\partial_2)(1 + t_1\partial_1)p)(z_1, z_2) = p(z_1, z_2) + t_1(a_{10} + a_{11}z_2) + t_2(a_{01} + a_{11}z_1) + t_2t_1a_{11}.$$

En posant $z_1 = z_2 = 0$, alors on obtient donc dans le développement de Taylor

$$a_{00} + a_{10}z_1 + a_{01}z_2 + a_{11}z_1z_2.$$

★

3.3 Preuve du théorème 2 de MSS

Lemme 3.12. *Soient $A_1, \dots, A_d$ des variables aléatoires indépendantes, à valeurs dans les matrices de rang 1 de $M_n(\mathbb{C})$. Soit $A = \sum_{i=1}^d A_i$. Alors*

$$\mathbb{E}[p_A](z) = \mu[\mathbb{E}[A_1], \mathbb{E}[A_2], \dots, E[A_d]](z).$$

Le rôle de l'indépendance est que si X, Y sont des variables aléatoires ordinaires, indépendantes, alors

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y],$$

et plus généralement, si $X_1, \dots, X_k$ sont des variables aléatoires usuelles indépendantes, alors

$$\mathbb{E}[X_1^{\varepsilon_1}X_2^{\varepsilon_2}\cdots X_k^{\varepsilon_k}] = \mathbb{E}[X_1]^{\varepsilon_1}\cdots\mathbb{E}[X_k]^{\varepsilon_k} \ (\varepsilon_i \in \{0, 1\}).$$

Plus généralement, si $p(z_1, \dots, z_k)$ est un polynôme affine multilinéaire (donc tous les exposants sont 0 ou 1), alors

$$\mathbb{E}[p](X_1, \dots, X_k) = p(\mathbb{E}[X_1], \dots, \mathbb{E}[X_k]).$$

Preuve (du Lemme 3.12)**.** Par la Proposition 3.10, on sait que

$$p_A(z) = \mu[A_1, \dots, A_d](z).$$

Ainsi $\mu[A_1, \dots, A_d](z)$ est un polynôme affine multilinéaire en les coefficients de $A_1, \dots, A_d$ (par la preuve de la Proposition 3.10). Par indépendance et le raisonnement ci-dessus,

$$\mathbb{E}[p_A](z) = \mathbb{E}[\mu[A_1, \dots, A_d](z)] = \mu[\mathbb{E}[A_1], \dots \mathbb{E}[A_d]](z).$$

□

Si $p(z)$ est un polynôme réel stable, on note $ZM(p)$ [1] le plus grand zéro réel de p.

1. La notation ZM est pour « zéro maximal ».

Lemme 3.13. *Soient $A_1, \dots, A_d$ des variables aléatoires indépendantes, à valeurs dans les matrices de rang 1 de $M_n(\mathbb{C})$. Soit $A = \sum_{i=1}^{d} A_i$. Supposons encore que les A_i sont positives et prennent un nombre fini de valeurs. Alors pour tout $j \in \{1, \dots, d\}$, pour tout réalisation de $A_1, \dots, A_{j-1}$ on a pour au moins une réalisation de A_j*

$$\begin{aligned} &\mathsf{ZM}\left(\mu\left[A_1, \dots, A_{j-1}, A_j, \mathbb{E}[A_{j+1}], \dots, \mathbb{E}[A_d]\right]\right) \\ &\quad \leqslant \mathsf{ZM}\left(\mu\left[A_1, \dots A_{j-1}, \mathbb{E}[A_j], \mathbb{E}[A_{j+1}], \dots, \mathbb{E}[A_d]\right]\right). \end{aligned}$$

Preuve. L'espérance $\mathbb{E}[A_j]$ est une moyenne pondérée des valeurs de A_j, donc

$$\mu\left[A_1, \dots, A_{j-1}, \mathbb{E}[A_j], \mathbb{E}[A_{j+1}], \dots, \mathbb{E}[A_d]\right]$$

est une moyenne pondérée des polynômes

$$\mu\left[A_1, \dots, A_{j-1}, A_j, \mathbb{E}[A_{j+1}], \dots, \mathbb{E}[A_d]\right],$$

qui sont réels stables. Par la série 5,

$$\mathsf{ZM}\left(\mu\left[A_1, \dots A_{j-1}, \mathbb{E}[A_j], \mathbb{E}[A_{j+1}], \dots, \mathbb{E}[A_d]\right]\right)$$

est dans l'enveloppe convexe (i.e. le plus petit intervalle contenant les ZM de $\mu\left[A_1, \dots, A_{j-1}, A_j, \mathbb{E}[A_{j+1}], \dots, \mathbb{E}[A_d]\right]$ prises sur les réalisations de A_j). Pour au moins une réalisation de A_j, on aura

$$\begin{aligned} &\mathsf{ZM}\left(\mu\left[A_1, \dots, A_{j-1}, A_j, \mathbb{E}[A_{j+1}], \dots, \mathbb{E}[A_d]\right]\right) \\ &\quad \leqslant \mathsf{ZM}\left(\mu\left[A_1, \dots A_{j-1}, \mathbb{E}[A_j], \mathbb{E}[A_{j+1}], \dots, \mathbb{E}[A_d]\right]\right). \end{aligned}$$

□

Preuve (du Théorème 2). Par le Lemme 3.12,

$$\mathbb{E}[p_A] = \mu\left[\mathbb{E}[A_1], \dots, \mathbb{E}[A_d]\right]$$

$$\implies \mathsf{ZM}(\mathbb{E}[p_A]) = \mathsf{ZM}\left(\mu\left[\mathbb{E}[A_1], \dots, \mathbb{E}[A_d]\right]\right).$$

On utilise les d inégalités du Lemme 3.13. Pour une réalisation de A_1, on aura

$$\begin{aligned} &\mathsf{ZM}\left(\mu\left[A_1, \mathbb{E}[A_2], \dots, \mathbb{E}[A_d]\right]\right) \\ &\quad \leqslant \mathsf{ZM}\left(\mu\left[\mathbb{E}[A_1], \dots, \mathbb{E}[A_d]\right]\right). \end{aligned}$$

Pour au moins une réalisation de A_1 et A_2, on aura

$$\begin{aligned} & \mathsf{ZM}\left(\mu\left[A_1, A_2, \mathbb{E}[A_3], \ldots, \mathbb{E}[A_d]\right]\right) \\ & \qquad \leqslant \mathsf{ZM}\left(\mu\left[A_1, \mathbb{E}[A_2], \ldots, \mathbb{E}[A_d]\right]\right). \end{aligned}$$

En continuant ainsi, pour au moins une réalisation de $A_1, \ldots, A_d$, on aura

$$\mathsf{ZM}\left(\mu\left[A_1, A_2, \ldots, A_d\right]\right) \leqslant \mathsf{ZM}\left(\mu\left[p_A\right]\right).$$

Mais par la Proposition 3.10,

$$\mathsf{ZM}\left(\mu\left[A_1, A_2, \ldots, A_d\right]\right) = \mathsf{ZM}(p_A),$$

qui est la plus grande valeur propre de $A = \|A\|$ (car $A \geqslant 0$). Pour cette réalisation,

$$\|A\| \leqslant \mathsf{ZM}(\mathbb{E}[p_A])$$

□

3.4 Preuve du Théorème 1 de MSS

Proposition 3.14. *Si* $A_1, \ldots, A_d \geqslant 0$ *dans* $M_n(\mathbb{C})$ *sont telles que*

$$\sum_{i=1}^{d} A_i = \mathbb{1}_m$$

et $\mathrm{Tr}(A_i) \leqslant \varepsilon$ *pour* $i = 1, \ldots, d$. *Soit*

$$p(z_1, \ldots, z_d) := \det\left(\sum_{i=1}^{d} z_i A_i\right).$$

Alors

$$\prod_{i=1}^{d}(1 - \partial_i)p$$

n'a pas de zéro dans l'orthant $\{y \geqslant E\}$ *où* $E = (1+\sqrt{\varepsilon})^2(1, 1, \ldots, 1)$.

Preuve. On va utiliser la Proposition 3.7 avec $x = (\varepsilon + \sqrt{\varepsilon})(1, 1, \ldots, 1)$ et $\delta = 1+\sqrt{\varepsilon}$. Soit $t := \varepsilon+\sqrt{\varepsilon}$, alors $t+\delta = (1+\sqrt{\varepsilon})^2$. On doit vérifier plusieurs conditions pour utiliser la proposition :

1. p est réel stable, sans zéro dans $\{y \geqslant x\}$, pour $x = t(1, 1, \ldots, 1)$,
2. $\Phi_p^j(x) \leqslant 1 - \frac{1}{\delta}$.

Vérifions ces deux conditions.

1. On a que
$$\det\left(z\mathbb{1}_m + \sum_{i=1}^{d} z_i A_i\right)$$
est réel stable par la Proposition 3.3, et p est une spécialisation en $z = 0$. Ainsi p est réel stable par la Proposition 3.4. Si $y \geqslant x$, alors
$$\sum_{i=1}^{d} y_i A_i \geqslant \sum_{i=1}^{d} t A_i = t\mathbb{1}_m$$
(puisque $y_i \geqslant t$ pour tout i). Ainsi
$$\sum_{i=1}^{d} y_i A_i$$
est inversible et
$$\det\left(\sum_{i=1}^{d} y_i A_i\right) \neq 0,$$
i.e. $p(y_1, \ldots, y_d) \neq 0$ (ce qui signifie que p n'a pas de zéro dans l'orthant considéré).
2. On utilise la formule de Jacobi. Si $A \colon \mathbb{R} \to M_m(\mathbb{C})$, $t \mapsto A(t)$ est de classe C^1, alors la dérivée logarithmique de $\det(A(t))$ est donnée par
$$\frac{\det(A(t))'}{\det(A(t))} = \mathrm{Tr}(A(t)^{-1} A'(t)).$$
Moralement, la dérivée logarithmique du déterminant est la trace de la dérivée logarithmique.
Prenons
$$A(x_1, \ldots, x_d) := \sum_{i=1}^{d} x_i A_i.$$
Alors par Jacobi, on a
$$\Phi_p^j(x_1, \ldots, x_d) = \mathrm{Tr}\left(\left(\sum_{i=1}^{d} x_i A_i\right)^{-1} A_j\right).$$

En particulier,

$$\begin{aligned}
\Phi_p^j(t,\dots,t) &= \mathrm{Tr}\left(\left(t\sum_{i=1}^d A_i\right)^{-1} A_j\right)\\
&= \mathrm{Tr}\left((t\mathbb{1}_m)^{-1}A_j\right)\\
&= \frac{\mathrm{Tr}(A_j)}{t}\\
&\leqslant \frac{\varepsilon}{t}.
\end{aligned}$$

En exercice, on montre que

$$\frac{\varepsilon}{t}+\frac{1}{\delta}=1 \implies \Phi_p^j(t,\dots,t)\leqslant 1-\frac{1}{\delta}.$$

La Proposition 3.14 résulte alors de la Proposition 3.7.

□

Preuve (du Théorème 1). Par le Lemme 3.12,

$$\begin{aligned}
\mathbb{E}[p_A](z) &= \mu\left[\mathbb{E}[A_1],\dots,\mathbb{E}[A_d]\right](z)\\
&= \left(\prod_{i=1}^d (1-\partial_i)\det\left(z\mathbb{1}_m+\sum_{i=1}^d z_i\mathbb{E}[A_i]\right)\right)\Bigg|_{z_1=\cdots=z_d=0} && \mathbb{1}_m=\mathbb{E}[A]=\sum_{i=1}^d \mathbb{E}[A_i]\\
&= \left(\prod_{i=1}^d (1-\partial_i)\det\left(\sum_{i=1}^d (z+z_i)\mathbb{E}[A_i]\right)\right)\Bigg|_{z_1=\cdots=z_d=0} && t_i:=z+z_i\\
&= \left(\prod_{i=1}^d (1-\partial_i)\det\left(\sum_{i=1}^d t_i\mathbb{E}[A_i]\right)\right)\Bigg|_{t_1=\cdots=t_d=z}.
\end{aligned}$$

Puisqu'on veut utiliser la Proposition 3.14, on doit encore vérifier la condition sur les traces. Mais

$$\mathrm{Tr}(\mathbb{E}[A_i])=\mathbb{E}[\mathrm{Tr}(A_i)]=\mathbb{E}[\|A_i\|]\leqslant\varepsilon$$

(où la dernière égalité vient du fait $A_i\geqslant 0$ de rang 1, et donc on peut trouver une base où A_i est diagonale avec valeur propre λ de multiplicité 1 et des 0 ailleurs, et l'inégalité est une hypothèse). Par la Proposition 3.14, le polynôme

$$\prod_{i=1}^d (1-\partial_i)\det\left(\sum_{i=1}^d t_i\mathbb{E}[A_i]\right)$$

n'a pas de zéro dans l'orthant $\{y\geqslant E\}$ où $E=(1+\sqrt{\varepsilon})^2(1,1,\dots,1)$. Donc

$$ZM(\mathbb{E}[p_A])\leqslant(1+\sqrt{\varepsilon})^2.$$

Index