PROBABILITÉS NON-COMMUTATIVES ET ENTROPIE LIBRE

NOÉ BLASSEL

1. $\mathbb{C}$-algèbres

Pour le probabiliste, la théorie des probabilités non-commutatives peut s'envisager comme une opportunité de s'affranchir l'encodage ensembliste des évènements, pour ne considérer les espaces de probabilités que comme des objets algébriques, identiques à l'algèbre de leurs variables aléatoires. Dans le cas classique, cette algèbre est commutative, mais en relâchant la contrainte de commutativité, on peut définir des espaces de probabilités, dits *non commutatifs* et découvrir une théorie parallèle, différente de celle des probabilités classiques, mais cependant propice à de nombreuses analogies avec celle-ci. Avant de s'y aventurer, on pourrait craindre que ce qu'on gagne en simplicité et en synthèse à considérer les espace de probabilités comme de simples algèbres, on le paye en abandonnant le point de vue de la théorie de la mesure et ses techniques puissantes. L'objet de cette section est de faire sentir pourquoi ce n'est pas le cas. On y introduit d'abord les premiers exemples de $\mathbb{C}$-algèbres, qui se trouveront être omniprésentes, dans les différents contextes que nous présenteront. Puis, on montrera comment, en imposant certaines conditions sur nos algèbres, on peut peu à peu récupérer notre capacité à faire de l'analyse, en passant des $\mathbb{C}$-algèbres aux $*$-algèbres, puis aux C*-algèbres, pour enfin arriver à la notion d'algèbre de von Neumann, qui, grâce à la synthèse qu'elle permet entre propriétés analytiques et propriétés algébriques, constitue un cadre privilégié pour les probabilités non-commutatives.

1.1. Structures de $\mathbb{C}$-algèbres.

Définition 1.1. *Soit $\mathcal{A}$ un anneau. Une **$\mathcal{A}$-algèbre unitaire**, est un $\mathcal{A}$-module $\mathcal{M}$ muni d'une structure d'anneau, telle que la multiplication soit $\mathcal{A}$-bilinéaire. Dans notre contexte $\mathcal{A}$ est souvent un corps, $\mathbb{C}$, si bien que $\mathcal{M}$ est un $\mathcal{A}$-espace vectoriel. La distributivité de la multiplication implique alors qu'une multiplication définie sur $B \times B$, où B est une base de $\mathcal{M}$ s'étend de manière unique à $\mathcal{M} \times \mathcal{M}$.*

On appelle morphisme de $\mathcal{A}$-algèbres unitaires tout morphisme de $\mathcal{A}$-modules telle que l'application sous-jacente soit aussi un morphisme pour la structure d'anneau, et plus généralement morphisme de $\mathcal{A}$-algèbres tout morphisme φ de $\mathcal{A}$-modules vérifiant

$$\varphi(ab) = \varphi(a)\varphi(b) \ \forall a, b \in \mathcal{M}$$

Exemple 1. Des exemples familiers d'algèbres sont fournis par les algèbres commutatives, dont le prototype est l'algèbre des polynômes $\mathbb{C}[X_1, \cdots, X_n]$. Une autre famille d'exemples est celle des algèbres de fonctions, dont la nature pourra varier en fonction de la structure du domaine de définition X.

Si X est un ensemble, l'algèbre $\mathcal{F}(X, \mathbb{C})$ de toutes les fonctions à valeurs complexes.

Si X est (l'ensemble sous-jacent à) un espace topologique, l'algèbre $\mathcal{C}(X, \mathbb{C}) \leqslant \mathcal{F}(X, \mathbb{C})$ des fonctions continues

Si $X \subset \mathbb{C}^m$, l'algèbre $P(X, \mathbb{C}) \leqslant \mathcal{C}(X, \mathbb{C})$ des fonctions polynomiales, image de $\mathbb{C}[X_1, \cdots, X_n]$ par le morphisme d'évaluation $\varphi : \mathbb{C}[X_1, \cdots, X_m] \mapsto \mathcal{F}(X, \mathbb{C})$ envoyant X_i sur $\pi_i : X \mapsto \mathbb{C}$ la projection sur la i-ème coordonnée.

Si $(X, \mathcal{G})$ est un espace mesurable, l'algèbre $M(X, \mathcal{G}, \mathbb{C}, \mathcal{B}(\mathbb{C})) \leqslant \mathcal{F}(X, \mathbb{C})$ des fonctions boréliennes à valeurs complexes.

Si $(X, \mathcal{G}, \mu)$ est un espace mesuré, l'algèbre $L^\infty(X, \mathcal{G}, \mu) \leqslant M(X, \mathcal{G}, \mathbb{C}, \mathcal{B}(\mathbb{C}))$ des fonctions mesurables essentiellement bornées en module, où encore, de manière plus adaptée à notre étude, avec μ une mesure de probabilités, l'algèbre $L^{\infty-}(X, \mathcal{G}, \mu) = \bigcap_{p \geqslant 1} L^p(X, \mathcal{G}, \mu)$ des fonctions mesurables ayant des moments de tout ordre. Le seul point est de voir que si $f, g \in L^{\infty-}(X, \mathcal{G}, \mu)$ alors $fg \in L^{\infty-}(X, \mathcal{G}, \mu)$ car

$$\forall p > 1, \ \int_X |fg|^p \mathrm{d}\mu \leqslant \sqrt{\int_X |f|^{2p} \mathrm{d}\mu \int_X |g|^{2p} \mathrm{d}\mu} < \infty$$

par Cauchy-Schwartz, et comme $f, g \in L^{2p}$.

Exemple 2. L'exemple fondamental d'algèbre non-commutative est celle des polynômes non-commutatifs en n indéterminées, $\mathbb{C}\langle X_1, \cdots, X_n \rangle$: il s'agit du $\mathbb{C}$-espace vectoriel ayant pour base les monômes unitaires non-commutatifs :

$$\left\{ X_{i_1}^{p_1} \cdots X_{i_k}^{p_k} |\ k \in \mathbb{N},\ p_j \in \mathbb{N}_+ \ \forall j,\ i_1 \neq i_2, i_2 \neq i_3, \cdots, i_{k-1} \neq i_k \right\}$$

qui sont simplement les mots sur l'alphabet $\{X_1, \cdots, X_n\}$. On définit alors la multiplication sur cette base par concaténation des monômes :

$$\left(z X_{i_1}^{p_1} \cdots X_{i_k}^{p_k} \right) \cdot \left(w X_{j_1}^{q_1} \cdots X_{j_l}^{q_l} \right) = \begin{cases} z w X_{i_1}^{p_1} \cdots X_{i_k}^{p_k} X_{j_1}^{q_1} \cdots X_{j_l}^{q_l} & i_k \neq j_1 \\ z w X_{i_1}^{p_1} \cdots X_{i_k}^{p_k + q_1} X_{j_2}^{q_2} \cdots X_{j_l}^{q_l} & i_k = j_1 \end{cases}$$

où $z, w \in \mathbb{C}$.

Cette définition s'étend par distributivité aux combinaisons $\mathbb{C}$-linéaires de monômes, et le mot vide, qu'on note 1 (le cas $k = 0$), apparaît comme l'élément neutre pour cette multiplication. Étant donné une $\mathbb{C}$-algèbre $\mathcal{M}$ munie de n éléments $(a_1, \cdots, a_n)$, on notera $\langle a_1, \cdots, a_n \rangle$ la sous-$\mathbb{C}$-algèbre de $\mathcal{M}$ générée par les $(a_1, \cdots, a_n)$, l'image de $\mathbb{C}\langle X_1, \cdots, X_n \rangle$ par le morphisme φ de $\mathbb{C}$-algèbres uniquement déterminé par $\varphi(X_i) = a_i \ \forall\ 1 \leqslant i \leqslant n$.

Exemple 3. Soit G un groupe au plus dénombrable d'élément neutre e. On peut définir l'algèbre de G, $\mathbb{C}G$ comme le $\mathbb{C}$-espace vectoriel de base G, $\mathbb{C}^{(G)}$ muni d'une loi de multiplication interne, définie par :

$$\left(\sum_{g \in G} \alpha_g g \right) \cdot \left(\sum_{g' \in G} \beta_{g'} g' \right) = \sum_{h \in G} \left(\sum_{gg' = h} \alpha_g \beta_{g'} \right) h = \sum_{(g,g') \in G^2} \alpha_g \beta_{g'} g g',\ \alpha_g, \beta_g \in \mathbb{C}\ \forall g \in G$$

Il s'agit donc des combinaisons $\mathbb{C}$-linéaires d'éléments de G, qu'on peut voir comme l'ensemble des suites complexes indexées par G à support fini, et cette opération (où on abuse en écrivant $g = (\delta_{gh})_{h \in G}$) admet bien pour élément neutre e , et est clairement $\mathbb{C}$-linéaire par rapport à chaque opérande.

En outre on a bien l'associativité :

$$\left(\left(\sum_{g\in G}\alpha_g g\right)\left(\sum_{h\in G}\beta_h h\right)\right)\left(\sum_{k\in G}\gamma_k k\right) = \sum_{l\in G}\left(\sum_{g'k=l}\left(\sum_{gh=g'}\alpha_g\beta_h\right)\gamma_k\right) l$$
$$= \sum_{l\in G}\left(\sum_{ghk=l}\alpha_g\beta_h\gamma_k\right) l = \sum_{l\in G}\left(\sum_{gg'=l}\alpha_g\left(\sum_{hk=g'}\alpha_g\beta_h\right)\right) l$$
$$= \left(\sum_{g\in G}\alpha_g g\right)\left(\left(\sum_{h\in G}\beta_h h\right)\left(\sum_{k\in G}\gamma_k k\right)\right)$$

La première chose qu'on puisse espérer d'une algèbre complexe est qu'elle possède une opération analogue à la conjugaison dans $\mathbb{C}$. C'est cette idée qui motive la définition suivante :

Définition 1.2. *Une* $*$***-algèbre*** $\mathcal{M}$ *est une* $\mathcal{A}$*-algèbre unitaire munie d'une application* $* : \mathcal{M} \mapsto \mathcal{M}$ *telle que :*

*(*1)* $x^{**} = x\ \forall x \in \mathcal{M}$

*(*2)* $(x+y)^* = x^* + y^*\ \forall x, y \in \mathcal{M}$

*(*3)* $1^*_{\mathcal{M}} = 1_{\mathcal{M}}$

*(*4)* $(x\cdot y)^* = y^*\cdot x^*\ \forall x, y \in \mathcal{M}$

Un morphisme entre deux $*$*-algèbres* φ *vérifiant* $\varphi(x^*) = \varphi(x)^*\ \forall x$ *est un* $*$*-morphisme.On dira enfin d'un élément* $x \in \mathcal{M}$ *qu'il est :*

auto-adjoint *si* $x^* = x$

normal *si* $xx^* = x^*x$

unitaire *si* $xx^* = x^*x = 1_{\mathcal{M}}$

positif *si* $x = y^*y$ *pour un certain* $y \in \mathcal{M}$

isométrique *si* $x^*x = 1_{\mathcal{M}}$

On a clairement unitaire $\Longrightarrow$ isométrique et auto-adjoint, positif $\Longrightarrow$ auto-adjoint $\Longrightarrow$ normal. Dans le cas où $\mathcal{M}$ est de surcroît une $\mathbb{C}$-algèbre, on imposera de plus la condition :

$$(*\mathbb{C})(\lambda x)^* = \overline{\lambda}x^* \forall x \in \mathcal{M}\ \forall \lambda \in \mathbb{C}$$

Dans ce cas, $*$ joue un rôle analogue à l'involution sur $\mathbb{C}$ qui à un nombre complexe z associe son conjugué $\overline{z}$, et on a un analogue de la décomposition en partie réelle et imaginaire : $\forall x \in \mathcal{M}$, il existe une unique décomposition sous la forme $x = a + ib$ où a, b sont auto-adjoints, donnée par $x = \frac{x+x^*}{2} + i\frac{x-x^*}{2i}$. Si de plus x est normal, on observe que ces deux parties commutent.

Exemple 4. Soit $\mathcal{H}$ un espace de Hilbert quelconque. L'espace $B(\mathcal{H}) = \mathcal{L}(\mathcal{H}, \mathcal{H})$ des endomorphismes bornés de $\mathcal{H}$ forment une algèbre pour l'opération de composition. On rappelle aussi que le théorème de représentation de Riesz fournit un isomorphisme isométrique $\varphi : \mathcal{H} \mapsto \mathcal{H}'$, où $\mathcal{H}'$ est le dual topologique de $\mathcal{H}$, l'espace des formes linéaires continues sur $\mathcal{H}$.

On peut alors définir une involution $*$ qui à un opérateur A associe son adjoint A^*, caractérisé par $\langle Au, v\rangle = \langle u, A^*v\rangle\ \forall u, v \in \mathcal{H}$, et défini par $A^* : v \mapsto \varphi^{-1}(u \mapsto \langle Au, v\rangle)$, qui munit l'ensemble des endomorphismes continus sur $\mathcal{H}$ d'une structure de $*$-algèbre.

Exemple 5. Considérons la $\mathbb{C}$-algèbre des polynômes non-commutatifs à $2n$ indéterminées,

$$\mathcal{M} = \mathbb{C}\langle X_1, \cdots, X_n, X_1^*, \cdots, X_n^* \rangle$$

On peut définir $*$ sur les monômes par la formule :

$$\left(zX_{i_1}^{\epsilon_1} \cdots X_{i_k}^{\epsilon_k}\right)^* = \overline{z} X_{i_k}^{*\epsilon_k} \cdots X_{i_1}^{*\epsilon_1},\ z \in \mathbb{C},\ k \in \mathbb{N}, \epsilon_j \in \{1, *\}\ \forall j,\ *1 = *,\ ** = 1$$

que l'on étend linéairement à $\mathcal{M}$. On a alors

$$\left(zX_{i_1}^{\epsilon_1} \cdots X_{i_k}^{\epsilon_k}\right)^{**} = \left(\overline{z} X_{i_k}^{*\epsilon_1} \cdots X_{i_1}^{*\epsilon_k}\right)^* = \overline{\overline{z}} X_{i_1}^{**\epsilon_1} \cdots X_{i_k}^{**\epsilon_k} = zX_{i_1}^{\epsilon_1} \cdots X_{i_k}^{\epsilon_k}$$

$$\left(zX_{i_1}^{\epsilon_1} \cdots X_{i_k}^{\epsilon_k} \cdot wX_{j_1}^{\nu_1} \cdots X_{j_l}^{\nu_l}\right)^* = \overline{zw} X_{j_l}^{*\nu_l} \cdots X_{j_1}^{*\nu_1} X_{i_k}^{*\epsilon_k} \cdots X_1^{*\epsilon_1}$$

$$= \left(\overline{w} X_{j_l}^{*\nu_l} \cdots X_{j_1}^{*\nu_1}\right) \cdot \left(\overline{z} X_{i_k}^{*\epsilon_k} \cdots X_1^{*\epsilon_1}\right) = \left(wX_{j_1}^{\nu_1} \cdots X_{j_l}^{\nu_l}\right)^* \cdot \left(zX_{i_1}^{\epsilon_1} \cdots X_{i_k}^{\epsilon_k}\right)^*$$

Autrement dit $(*1)$ et $(*4)$ sont vérifiées sur les monômes, puis sur tout $\mathcal{M}$ par linéarité. $(*2)$ et $(*3)$ sont vérifiées par construction.

Si $\mathcal{M}$ est une $*$-algèbre complexe, et $(a_1, \cdots, a_n) \in \mathcal{M}^n$, on parlera de la sous-$*$-algèbre $*\langle a_1, \cdots, a_n \rangle$ générée par les a_i, comme l'image de $\mathbb{C}\langle X_1, \cdots, X_n, X_1^*, \cdots, X_n^* \rangle$ par le $*$-morphisme φ déterminé par $\varphi(X_i) = a_i\ \forall 1 \leqslant i \leqslant n$ (et donc également $\varphi(X_i^*) = a_i^*\ \forall 1 \leqslant i \leqslant n$)

Exemple 6. On revient à l'algèbre de groupe $\mathbb{C}G$, en posant

$$\left(\sum_{g \in G} \alpha_g g\right)^* = \sum_{g \in G} \overline{\alpha_g} g^{-1}$$

$(*1)$,$(*2)$,$(*3)$ et $(*\mathbb{C})$ sont claires, il sufift de vérifier $(*4)$:

$$\left(\left(\sum_{g \in G} \alpha_g g\right)\left(\sum_{h \in G} \beta_h h\right)\right)^* = \sum_{(g,h) \in G^2} \overline{\alpha_g \beta_h}\, (gh)^{-1}$$

$$= \sum_{(g,h) \in G^2} \overline{\beta_h} \overline{\alpha_g} h^{-1} g^{-1} = \left(\sum_{g \in G} \overline{\beta_h} h^{-1}\right)\left(\sum_{g \in G} \overline{\alpha_g} g^{-1}\right)$$

Ainsi $*$ définit une structure de $*$-algèbre sur $\mathbb{C}G$. Notons en particulier que comme $g^* = g^{-1}$, chaque $g \in G$, vu comme vecteur de $\mathbb{C}G$, est unitaire.

Définition 1.3. *Une **algèbre de Banach** sur un corps $\mathbb{K}$ est une $\mathbb{K}$-algèbre $\mathcal{M}$ telle que $\mathcal{M}$ soit un espace de Banach en tant qu'espace vectoriel, c'est-à-dire qu'il existe une norme sur $\mathcal{M}$ $||\cdot||$, telle que $\mathcal{M}$ soit complet pour la distance induite, et vérifiant :*

$$\textbf{(BA1)}\ :\|a \cdot b\| \leqslant \|a\| \cdot \|b\|\ \forall a, b \in \mathcal{M}$$

De plus, si $\mathcal{M}$ est unitaire, on requiert :

$$\textbf{(BA2)}\ :\|1_{\mathcal{M}}\| = 1$$

*Une **C*-algèbre** $\mathcal{M}$ est une $\mathbb{C}$-algèbre de Banach involutive (en particulier elle satisfait $*\mathbb{C}$), satisfaisant la propriété suivante :*

$$\textbf{(C*1)}\ :\|x^*x\| = \|x\|\|x^*\|\ \forall x \in \mathcal{M}$$

On appelera morphisme de C-algèbres un $*$-morphisme entre deux C*-algèbres. L'exemple fondamental de C*-algèbre est fourni par celle des opérateurs bornés sur un espace de Hilbert complexe :*

Exemple 7. L'espace $\mathcal{L}(E,E)$ muni de la norme d'opérateur

$$\|A\| = \sup_{\|x\|=1} \|Ax\|$$

est une algèbre de Banach pour tout espace de Banach E. En particulier si $\mathcal{H}$ est un espace de Hilbert complexe, $B(\mathcal{H})$ est une $*$-algèbre de Banach complexe où $*\mathbb{C}$ est garantie par sesquilinéarité du produit scalaire :

$$\langle \lambda Au, v\rangle = \lambda\langle Au, v\rangle = \lambda\langle u, A^*v\rangle = \langle u, \overline{\lambda}A^*v\rangle \ \forall u, v \in \mathcal{H} \ \forall \lambda \in \mathbb{C}$$

Pour C*1, faisons la remarque suivante : pour $A \in \mathcal{B}(\mathcal{H})$, on a

$$\|A\| = \sup_{\|u\|,\|v\|=1} |\langle Au, v\rangle|$$

En effet,on a par Cauchy-Schwartz

$$|\langle Au, v\rangle| \leqslant \|A\|\|u\|\|v\|$$

$$\sup_{\|u\|,\|v\|=1} \langle Au, v\rangle \leqslant \|A\|$$

Pour l'autre sens, on peut choisir une suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de norme 1 telle que

$$\|Au_n\| \xrightarrow{n\to\infty} \|A\|$$

et on pose $v_n = \frac{Au_n}{\|Au_n\|}$ qui vérifie $\|v\|_n = 1 \ \forall n$.

Ceci donne

$$|\langle Au_n, v_n\rangle| = \|Au_n\| \xrightarrow{n\to\infty} \|A\|$$

$$\sup_{\|u\|,\|v\|=1} |\langle Au, v\rangle| \geqslant \|A\|$$

En particulier,

$$\|A\| = \sup|\langle Au, v\rangle| = \sup|\langle u, A^*v\rangle| = \sup|\overline{\langle A^*v, u\rangle}| = \sup|\langle A^*v, u\rangle| = \|A^*\|$$

donc $*$ est une isométrie. D'autre part,

$$\|A^*A\| = \sup|\langle A^*Au, v\rangle| = \sup|\langle Au, Av\rangle| = \|A\|^2$$

d'où $\|A^*A\| = \|A\|\|A\| = \|A\|\|A^*\|$.

De manière plus générale, toute sous-algèbre de $\mathcal{B}(\mathcal{H})$ stable par passage à l'adjoint est une $*$-algèbre complexe normée vérifiant C*1. Pour en faire une C*-algèbre, il suffit donc qu'elle soit de Banach, c'est-à-dire complète, ce qui équivaut à dire qu'elle constitue une partie fermée de $\mathcal{B}(\mathcal{H})$ pour la topologie induite par la norme d'opérateurs. Bien que toutes les C*-algèbres ne sont pas nécessairement décrites commes sous-algèbres de $\mathcal{B}(\mathcal{H})$ (un exemple est fourni par les fonctions continues sur un compact de$\mathbb{C}$), elles constituent un exemple suffisament important pour noter cette quasi-définition plus concrète, qui donnent leur noms aux C*-algèbres (pour *closed $*$-algebra*), et qui trouvera écho dans l'introduction des algèbres de von Neumann. Il se trouve, par un théorème de Gelfand et Neimark[Arv76, p. 34] que toute C*-algèbre est $*$-isomorphe à une C*-algèbre d'opérateurs de manière isométrique, si bien que cette quasi-définition peut aussi servir de définition.

En outre, cette observation justifie la terminologie suivante : pour $(a_1, \cdots, a_n) \in \mathcal{B}(\mathcal{H})$, on écrira

$$\mathrm{C}^*\langle a_1, \cdots, a_n\rangle = \overline{\varphi(\mathbb{C}\langle X_1, \cdots X_n, X_1^*, \cdots X_n^*\rangle)}$$

la C*-algèbre engendrée par les $(a_1, \cdots, a_n)$, où φ est le $*$-morphisme d'évaluation envoyant X_i sur a_i, et où l'adhérence est prise pour la topologie de la norme d'opérateur. Il suffit de vérifier que l'adhérence est bien une $*$-algèbre, ce qui est garanti par la continuité des applications $(a,b) \mapsto a+b, (a,b) \mapsto ab$ et $a \mapsto a^*$.

On est passé d'une structure purement algébrique, celle de $*$-algèbre, à la structure de C*-algèbre qui hérite de ses propriétés algébriques, mais qui remplit également une condition d'ordre topologique, celle d'être complète (où fermée du point de vue des algèbres d'opérateurs). Il existe une troisième structure, celle d'algèbre de von Neumann, qui, vue du point de vue des sous-algèbres de $\mathcal{B}(\mathcal{H})$, correspond à un affaiblissement de topologies.

Définition 1.4. *On se place dans l'algèbre d'opérateurs $\mathcal{B}(\mathcal{H})$.*

*La topologie **uniforme** ou de la **norme d'opérateur** est la topologie induite par la norme. Une suite $(A_n)_{n\geqslant 1}$ converge vers 0 au sens de cette topologie si et seulement si $\|A_n\| \to 0$.*

*La topologie **forte d'opérateurs** est la topologie engendrée par la famille d'applications $(A \mapsto Ax)_{x\in\mathcal{H}}$. de $\mathcal{B}(\mathcal{H})$ dans $\mathcal{H}$, c'est-à-dire la topologie la moins fine rendant cette famille continue. Une suite $(A_n)_{n\geqslant 1}$ converge vers 0 si et seulement si $\|A_n x\| \to 0\ \forall x \in \mathcal{H}$.*

*La topologie **faible d'opérateurs** est la topologie engendrée par la famille d'applications $(A \mapsto y(Ax))_{x\in\mathcal{H}, y\in\mathcal{H}'}$, ou, de manière équivalente par représentation de Riesz, celle des $(A \mapsto \langle Ax, y\rangle)_{x,y\in\mathcal{H}}$. Une suite $(A_n)_{n\geqslant 1}$ converge vers 0 si et seulement si $\langle A_n x, y\rangle \to 0\ \forall x, y \in \mathcal{H}$.*

Explicitement, ces topologies admettent les prébases formées des cylindres :
Pour la topologie forte, les ouverts de la forme $U^s(A, x, \epsilon) = \{B \in \mathcal{B}(\mathcal{H}) : \|Ax - Bx\| < \epsilon\}$
Pour la topologie faible, ceux de la forme $U^w(A, x, y, \epsilon) = \{B \in \mathcal{B}(\mathcal{H}) : |\langle Ax - Bx, y\rangle| < \epsilon\}$
En conséquence, tout ouvert s'écrit comme une union d'intersections finies de tels cylindres.
Les inégalités $\|Ax\| \leqslant \|A\|\|x\|, |\langle Ax, y\rangle| \leqslant \|Ax\|\|y\|$ montrent que la convergence en norme implique la convergence forte, qui à son tour implique la convergence faible.
Inversement, les exemples $A_n : f \mapsto \chi_{[n,n+1]} f$ dans $L^2(\mathbb{R})$ et $B_n : (u_k)_k \mapsto (u_{k-n})_k$ dans $\ell^2(\mathbb{Z})$ montrent que la topologie uniforme est en général strictement plus fine que la forte d'opérateurs, elle-même en général strictement plus fine que la faible d'opérateurs. On introduit à présent une notion algébrique à la base d'une des définitions possibles des algèbres de von Neumann.

Définition 1.5. *Soit $\mathcal{A}$ une $\mathcal{R}$-algèbre unitaire, et $\mathcal{M} \subset \mathcal{A}$ une sous-algèbre. On définit le **commutant** de $\mathcal{M}$ comme l'ensemble*

$$\mathcal{M}' = \{x \in \mathcal{A} : xm = mx\ \forall m \in \mathcal{M}\} = \bigcap_{m\in\mathcal{M}} \ker \varphi_m$$

où $\varphi_m : x \mapsto xm - mx$ est un endomorphisme $\mathcal{R}$-linéaire.

C'est donc un sous-module de $\mathcal{A}$, et $a, b \in \mathcal{M}' \implies abm = amb = mab\ \forall m \in \mathcal{M} \implies ab \in \mathcal{M}', 1_{\mathcal{A}} \in \mathcal{M}'$ montre que c'est une sous-algèbre unitaire de $\mathcal{A}$.

*On définit alors $\mathcal{M}'' = (\mathcal{M}')', \mathcal{M}^{(n+1)} = (\mathcal{M}^{(n)})'$. On nomme en particulier $\mathcal{M}''$ le **bicommutant** de $\mathcal{M}$.*

On voit alors facilement les propriétés suivantes :

$$\forall t \in \mathcal{M}'\ \forall m \in \mathcal{M}\ mt = tm \implies \mathcal{M} \subseteq \mathcal{M}''$$

$$\mathcal{N} \subseteq \mathcal{M} \implies \bigcap_{\mathcal{M}} \ker \varphi_m = \mathcal{M}' \subseteq \mathcal{N}' = \bigcap_{\mathcal{N}} \ker \varphi_n$$

$$\mathcal{M} \subseteq \mathcal{M}'', \mathcal{M}' \subseteq \mathcal{M}^{(3)} \implies \mathcal{M}^{(3)} \subseteq \mathcal{M}', \mathcal{M}^{(4)} \subseteq \mathcal{M}''$$

$$\implies \mathcal{M}' = \mathcal{M}^{(2n-1)}, \mathcal{M}'' = \mathcal{M}^{(2n)}\ \forall n \geqslant 1$$

Nous arrivons au théorème du bicommutant, qui permet une caractérisation à la fois topologique et algébrique de certaines sous-algèbres de $\mathcal{B}(\mathcal{H})$, et qui est à la source de la définition des algèbres de von Neumann. Pour ce faire, notons le lemme suivant, qui fait un premier lien entre la géométrie de $\mathcal{H}$ et les projecteurs de $\mathcal{B}(\mathcal{H})$. Comme nous le rappellerons plus loin, un opérateur $P \in \mathcal{B}(\mathcal{H})$ est appelé projecteur si il est auto-adjoint et idempotent, soit $P = P^2 = P^*$. Rappelons qu'à tout sous-espace fermé de $\mathcal{H}$ on peut associer son projecteur, dit projecteur orthogonal.

Lemme 1.6. *Soit $M \subset \mathcal{H}$ un sous-espace fermé, P son projecteur orthogonal et $T \in \mathcal{B}(\mathcal{H})$. On a :*

i) $TM \subseteq M \iff PTP = TP$

ii) $TM \subseteq M$ *et* $T^*M \subseteq M \iff TP = PT$

Démonstration. i) $\implies$: Soit $x \in \mathcal{H}$. $Px \in M \implies TPx \in M \implies PTPx = TPx \implies PTP = TP$. Réciproquement, $PTP = TP, x \in M \implies Px = x \implies TPx = Tx = PTPx \implies Tx \in M$. ii) $\implies$: Par *i)*, on obtient $PT^*P = T^*P$ et $PTP = TP$ en prenant l'adjoint de la première égalité, on obtient $(PT^*P)^* = PTP = PT$, donc $PT = TP$. Réciproquement, $PT = TP \iff T^*P = PT^*$. Or $PT = TP \implies PTP = TP \implies TM \subseteq M$, et donc également $T^*M \subseteq M$. □

Théorème 1.7. *Du bicommutant (von Neummann, 1930)*[Sun87, p. 12]

Soit $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ une sous-$*$-algèbre de $\mathcal{H}$. Alors les adhérences de $\mathcal{M}$ pour les topologies fortes et faibles coïncident et sont égales au bicommutant $\mathcal{M}''$. Autrement dit, il y a équivalence entre les conditions :

i) $\mathcal{M} = \mathcal{M}''$

ii) $\mathcal{M}$ *est fermée pour la topologie faible d'opérateurs.*

iii) $\mathcal{M}$ *est fermée pour la topologie forte d'opérateurs.*

Démonstration. Commençons par observer que $\forall S \subset \mathcal{B}(\mathcal{H})$, S' est fermé pour la topologie faible d'opérateurs.

Soit $A \notin S'$. Alors $\exists B \in S$ tel que $AB - BA \neq 0$, et $\exists x, y \in \mathcal{H} : \langle (AB - BA)x, y\rangle = \langle ABx, y\rangle - \langle Ax, B^*y\rangle \neq 0$. Par continuité faible de $f : A \mapsto \langle ABx, y\rangle - \langle Ax, B^*y\rangle$, il existe un voisinage faible $\mathcal{V}$ de A tel que $0 \notin f(\mathcal{V})$. En particulier $\mathcal{V} \cap S' = \varnothing \implies \mathcal{B}(\mathcal{H}) \setminus S'$ est ouvert pour la topologie faible, et S' fermé.

On a donc $i \implies ii$.

$ii \implies iii$ est immédiate car la topologie forte d'opérateurs est plus fine que la faible.

Pour montrer $iii \implies i$, supposons que $\mathcal{M}$ soit fermé pour la topologie forte. On va montrer que $\mathcal{M}$ est dense dans $\mathcal{M}''$ pour cette dernière.

Soit donc $A'' \in \mathcal{M}''$, et $\mathcal{U}$ un ouvert contenant A''. Par définition de la topologie forte, $\mathcal{U}$ contient une intersection finie de cylindres $U^s(A'', x_i, \epsilon_i), 1 \leqslant i \leqslant n$.

Il suffit donc de montrer que $\forall \epsilon > 0, \forall x_1, \cdots x_n \in \mathcal{H},\ \exists A \in \mathcal{M} :\ \forall 1 \leqslant i \leqslant n,\ \|A''x_i - Ax_i\| < \epsilon$ pour avoir $\mathcal{U} \cap \mathcal{M} \neq \varnothing$ et conclure à la densité de $\mathcal{M}$.

La première étape consiste à montrer le cas $n = 1$. Soit donc $x \in \mathcal{H}$ et $\epsilon > 0$. On considère le sous-espace fermé $E = \overline{\{Bx, B \in \mathcal{M}\}}$, et soit P son projecteur. Pour $T, A \in \mathcal{M}$, on a clairement $TAx = (TA)x \in E$, et il s'ensuit que $TE \subseteq E$.

En outre, $T^*Ax \in E$ (comme on a supposé $\mathcal{M}$ une $*$-algèbre), donc on a aussi $T^*E \subseteq E$.

Par le lemme, $PT = TP$. Comme T était arbitraire, ceci montre $P \in \mathcal{M}'$. Par définition, on a $A''P = PA''$, donc on a encore $A''E \subseteq E$.

Comme $\mathcal{M}$ est unitaire, on a aussi $x \in E$, ce qui implique $A''x \in E = \overline{\{Ax, A \in \mathcal{M}\}}$, c'est-à-dire, $\exists A \in \mathcal{M} : \|A''x - Ax\| < \epsilon$.

Pour le cas général ($n \geqslant 2$), on considère $\hat{\mathcal{H}} = \mathcal{H} \oplus \cdots \oplus \mathcal{H}$ la somme directe de n copies de $\mathcal{H}$, et $\hat{\mathcal{M}} = \iota(\mathcal{M})$ l'injection diagonale de $\mathcal{M}$ dans $\mathcal{B}(\mathcal{H})$.

Un élément $\mathcal{M} \ni A \overset{\iota}{\mapsto} \hat{A}$ opère sur $\hat{x} = (x_1, \cdots, x_n)$ via $\hat{A}\hat{x} = (Ax_1, \cdots, Ax_n)$. Il est immédiat que $\iota(\mathcal{M})$ est une sous-$*$-algèbre de $\mathcal{B}(\hat{\mathcal{H}})$, et de plus on a $\iota(\mathcal{M})'' \subseteq \iota(\mathcal{M}'')$.

En effet, soit $S = (S_{ij})_{1 \leqslant i,j \leqslant n} \in \iota(\mathcal{M})'$, où on identifie $\mathcal{B}(\hat{\mathcal{H}})$ à $M_n(\mathcal{B}(\mathcal{H}))$. La condition $S\hat{A} = \hat{A}S \ \forall \hat{A} \in \iota(\mathcal{M})$ s'écrit $S_{ij}A = AS_{ij} \ \forall i,j \ \forall A \in \mathcal{M} \iff S_{ij} \in \mathcal{M}' \ \forall i,j$.

Soit à présent $T \in \iota(\mathcal{M})''$, c'est-à-dire $TS = ST \ \forall S \in \iota(\mathcal{M})' = M_n(\mathcal{M}')$. En considérant les $S = E_{ij}$ nuls en toutes leur coordonnées sauf en $S_{ij} = \mathrm{id}_{\mathcal{H}}$, la condition $TE_{ij} = E_{ij}T \ \forall i,j$ impose que T est nulle en dehors de sa diagonale et constante sur celle ci, T est donc de la forme $T = \iota(S)$ avec $S \in \mathcal{B}(\mathcal{H})$. Mais $\hat{S}$ doit commuter avec tous les $\iota(A),\ A \in \mathcal{M}'$, ce qui impose $S \in \mathcal{M}''$.

Soit $\hat{x} \in \hat{\mathcal{H}}$. Par le cas $n = 1$ appliqué à $\hat{\mathcal{H}}$,

$$\forall \epsilon > 0, \forall \hat{x} = (x_1, \cdots, x_n) \in \hat{\mathcal{H}} \forall \hat{A}'' \in \iota(\mathcal{M})'' \subseteq \iota(\mathcal{M}''), \exists \hat{A} = \iota(A) \in \iota(\mathcal{M}) : \ \left\| \hat{A}''\hat{x} - \hat{A}\hat{x} \right\|_{\hat{\mathcal{H}}} < \epsilon$$

On obtient le résultat escompté en écrivant la norme.

$$\left\| \hat{A}''\hat{x} - \hat{A}\hat{x} \right\|_{\hat{\mathcal{H}}}^2 = \sum_{i=1}^{n} \left\| A''x_i - Ax_i \right\|^2 < \epsilon^2 \implies \left\| A''x_i - Ax_i \right\| < \epsilon \ \forall 1 \leqslant i \leqslant n$$

□

Définition 1.8. *Soit $\mathcal{H}$ un espace de Hilbert complexe et $\mathcal{M}$ une sous-$*$-algèbre de $\mathcal{B}(\mathcal{H})$. Si $\mathcal{M} = \mathcal{M}''$, on dit que c'est une* ***algèbre de von Neumann****. Par le théorème du bicommutant, cette définiion peut être remplacée par une des deux définitions équivalentes :*

$\mathcal{M}$ est fermée pour la topologie faible d'opérateurs

$\mathcal{M}$ est fermée pour la topologie forte d'opérateurs

Notons que comme ces topologies sont plus grossières que la topologie uniforme, tout algèbre de von Neumann est fermée métriquement, donc une C*-algèbre.

Comme pour les C*-algèbres, il existe une définition purement abstraite des algèbres de von Neumann, qui portent alors le nom de W*-algèbres. Cependant, toutes les algèbres que nous rencontrerons seront décrites explicitement à l'aide d'une représentation concrète comme une sous-algèbre de $\mathcal{B}(\mathcal{H})$.

Si $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ est une sous-$*$-algèbre, son adhérence faible $\overline{\mathcal{M}}^w$ est une algèbre de von Neumann, et si $A_1, \cdots, A_n \in \mathcal{B}(\mathcal{H})$, on peut en particulier parler de l'algèbre de von Neumann engendrée, $\mathrm{W}^*(A_1, \cdots, A_n) = \overline{*\langle A_1, \cdots, A_n \rangle}^w = \overline{*\langle A_1, \cdots, A_n \rangle}^s = *\langle A_1, \cdots, A_n \rangle''$.

Exemple 8. $1\mathbb{C}$, $\mathcal{B}(\mathcal{H})$ sont des algèbres de von Neumann. En particulier les algèbres de dimension finie $M_n(\mathbb{C})$ de matrices carrées $n \times n$ sont de von Neumann. Si $(\mathcal{B}_i)_{i \in I}$ sont des algèbres de von Neumann, $\bigcap_{i \in I} \mathcal{B}_i$ également, par la caractérisation topologique donnée par le théorème du bicommutant.

En particulier, si $\mathcal{B}$ est une algèbre de von Neumann, $Z(\mathcal{B}) := \mathcal{B} \cap \mathcal{B}'$ est une algèbre de von Neumann. Si $Z(\mathcal{B}) = 1\mathbb{C}$, on dit que $\mathcal{B}$ est un **facteur**.

Si $\mathcal{H} = L^2(X, \mathcal{F}, \mu)$, où μ est une mesure finie, on identifie $L^\infty(X, \mathcal{F}, \mu)$ à la sous-$*$-algèbre $\mathcal{M}$ de $\mathcal{B}(\mathcal{H})$ consistant des opérateurs de multiplication $\{T_f : g \mapsto gf,\ f \in L^\infty(X, \mathcal{F}, \mu)\}$. C'est une algèbre de von-Neumann commutative. Il suffit de montrer $\mathcal{M}' = \mathcal{M} : \mathcal{M} \subset \mathcal{M}'$, puisque c'est une algèbre commutative, et de plus, si $T \in \mathcal{M}'$, en posant $f = T1$, pour chaque $g \in L^\infty(X, \mathcal{F}, \mu)$, on a $Tg = TT_g1 = T_gT1 = T_gf = gf$. On a donc $\|f\|_\infty \leqslant \|T\| \implies f \in L^\infty(X, \mathcal{F}, \mu)$, et $T = T_f$, donc $\mathcal{M}' \subset \mathcal{M}$.

Soit G un groupe au plus dénombrable. On pose $\mathcal{H} = \ell^2(G)$. On peut alors plonger $\mathbb{C}[G]$ linéairement dans $\mathcal{B}(\mathcal{H})$ via le morphisme $\varphi : g \mapsto \lambda_g$ où $\lambda_g : (x_h)_{h\in G} \mapsto (x_{gh})_{h\in G}$, qui est clairement un morphisme d'algèbres injectif. Chaque λ_g ne fait que permuter les termes d'une suite : c'est donc un opérateur borné, en particulier une isométrie de $\mathcal{H}$. De plus, on vérifie que, comme $\langle \lambda_g x, y\rangle = \sum_{h\in G} x_{gh}\overline{y_h} = \overline{\sum_{h\in G} y_{g^{-1}h}\overline{x_h}} = \overline{\langle \lambda_{g^{-1}} y, x\rangle} = \langle x, \lambda_{g^{-1}} y\rangle$ que $\lambda_g^* = \lambda_{g^{-1}}$, ce plongement est donc un $*$-morphisme. En particulier $\varphi(\mathbb{C}[G])$ est une sous-$*$-algèbre de $\mathcal{B}(\mathcal{H})$, on peut donc former l'algèbre de von Neumann $L(G) = \varphi(\mathbb{C}[G])''$.

Un problème proéminent, non résolu, et une des motivations premières pour le développement des notions d'entropie libre, est le suivant :
Étant donné $n \neq m$ des entiers, les algèbres $L(\mathbb{F}_n)$ et $L(\mathbb{F}_m)$ sont-elles isomorphes ? (On écrit $\mathbb{F}_n$ pour le groupe libre engendré par n éléments). La propriété suivante illustre comment une propriété de G peut en principe incider sur la structure de son algèbre $L(G)$. Toute la question est alors de savoir combien l'algèbre $L(G)$ retient de la structure de G, et combien elle oublie.

Proposition 1.9. [Pop10] *Soit G un groupe dénombrable. Alors $L(G)$ est un facteur si pour chaque $G \ni g \neq e$, la classe de conjugaison $Cl(g) = \{h^{-1}gh, h \in G\}$ de g est infinie.*

Démonstration. Notons $(\delta_g)_{g\in G}$ la base canonique de $\ell^2(G)$. Soit A de la forme $\sum_{g\in G} \alpha_g \lambda_g$. Alors pour chaque $h \in G$, on a $\langle A\delta_{g^{-1}h}, \delta_h\rangle = \alpha_g$, et cette valeur ne dépend pas de h. Comme $A \mapsto \langle A\delta_{g^{-1}h}, \delta_g\rangle$ est continue pour la topologie faible d'opérateur, il s'ensuit que, pour tout $A \in L(G)$, l'application $h \mapsto \langle A\delta_{g^{-1}h}, \delta_g\rangle; G \mapsto \mathbb{C}$ est constante. (Autrement en prenant $A_n \to A$, et h_1, h_2 tels que $\langle A\delta_{g^{-1}h_1}, \delta_{h_1}\rangle \neq \langle A\delta_{g^{-1}h_2}, \delta_{h_2}\rangle$, la suite nulle $\langle A_n\delta_{g^{-1}h_1}, \delta_{h_1}\rangle - \langle A_n\delta_{g^{-1}h_2}, \delta_{h_2}\rangle$ ne converge pas vers 0). Notons $c_g(A)$ la valeur de cette application constante.
Soit maintenant $T \in Z(L(G))$. On a en particulier que $T\lambda_k = \lambda_k T\ \forall k \in G \implies \lambda_{k^{-1}} T\lambda_k = T\ \forall k \in G$. Ceci implique que

$$\begin{aligned} c_g(T) &= \langle T\delta_{g^{-1}h}, \delta_h\rangle = \langle \lambda_{k^{-1}} T\lambda_k \delta_{g^{-1}h}, \delta_h\rangle \\ &= \langle T\lambda_k\delta_{g^{-1}h}, \lambda_k\delta_h\rangle = \langle T\delta_{kg^{-1}h}, \delta_{kh}\rangle \\ &= \langle T\delta_{kg^{-1}k^{-1}kh}, \delta_{kh}\rangle = c_{kgk^{-1}}(T) \end{aligned}$$

Autrement dit $g \mapsto c_g(T)$ est constante sur les classes de conjugaisons. D'autre part on a $\langle T\delta_e, \delta_g\rangle = \langle T\delta_{g^{-1}g}, \delta_g\rangle = c_g(T) \implies T\delta_e = \sum_{g\in G} c_g(T)\delta_g$. En particulier, on a

$$\|T\delta_e\|^2 = \sum_{g\in G} |c_g(T)|^2 = |c_e(T)|^2 + \sum_{C \text{classe non-triviale}} |c_C(T)|^2 |C| < \infty$$

Comme chaque classe de conjugaison non-triviale est infinie, il s'ensuit que $T = c_e(T)1 \in \mathbb{C}1$, donc $L(G)$ est un facteur. □

Pour $n \geqslant 2$, l'algèbre $L(\mathbb{F}_n)$ est un facteur.
Il suffit de montrer que pour tout $x \in \mathbb{F}_n \setminus e$, $\mathrm{Cl}_{\mathbb{F}_n}(x)$ est infini. Notons $\mathbb{F}_n = \langle g_1, \cdots g_n\rangle$, et $x = g^p x', p \neq 0$, où l'écriture réduite de x' ne commence pas par g. Soit $k \in \{g_1, \cdots g_n\}, k \neq g$. Alors les $k^m x k^{-m},\ m \geqslant 1$ sont deux à deux distincts, car leurs écritures réduites respectives sont de la forme $k^m h$ où h ne commence pas par k.

Nous sommes à présent en mesure d'expliquer une des raisons d'être de la structure de C*-algèbre, la possibilité de définir un calcul fonctionnel continu

1.2. **Spectre et calculs fonctionnels.**

Un des avantage d'une C*-algèbre sur une $*$-algèbre est qu'elle contient toutes les limites en norme de ses éléments, et à son tour, un des avantages d'une algèbre de von Neumann sur une C*-algèbre est qu'elle contient toutes ses limites faibles d'opérateur. Une des conséquences de cet enrichissement topologique est qu'il est possible sous certaines conditions de définir des calculs fonctionnels plus riches que ce que permet une simple $\mathbb{C}$-algèbre. C'est l'objet de cette partie, qui nécessite avant un exposé de quelques notions de théorie spectrale.

Définition 1.10. *Soit $\mathcal{M}$ une C*-algèbre unitaire d'opérateurs de $\mathcal{B}(\mathcal{H})$. Pour $A \in \mathcal{M}$, on définit le* ***spectre*** *de A comme l'ensemble*

$$\sigma(A) = \{\lambda \in \mathbb{C} : \lambda 1_{\mathcal{M}} - A \text{ n'est pas inversible dans } \mathcal{B}(\mathcal{H})\}$$

Il est immédiat que $\sigma(A^*) = \overline{\sigma(A)}$. On notera aussi R_A ***l'application résolvante*** *:*

$$R_A : \begin{Bmatrix} \mathbb{C} \setminus \sigma(A) & \mapsto & \mathcal{B}(\mathcal{H}) \\ \lambda & \mapsto & (\lambda 1_{\mathcal{B}(\mathcal{H})} - A)^{-1} \end{Bmatrix}$$

On peut remarquer qu'on a l'équation résolvante :

$$\lambda, \mu \in \mathbb{C} \setminus \sigma(A) \implies R_A(\lambda) = R_A(\lambda)(\mu - A)R_A(\mu) = R_A(\lambda)(\lambda - A + \mu - \lambda)R_A(\mu)$$

$$= R_A(\lambda)(\lambda - A)R_A(\mu) + (\mu - \lambda)R_A(\lambda)R_A(\mu)$$

$$\implies R_A(\lambda) - R_A(\mu) = (\mu - \lambda)R_A(\lambda)R_A(\mu)$$

Le théorème suivant évoque le théorème fondamental de l'algèbre, et la preuve est essentiellement la même celle utilisant le théorème de Liouville.

Proposition 1.11. *Soit $A \in \mathcal{B}(\mathcal{H})$. Alors $\sigma(A)$ est un compact non-vide de $\mathbb{C}$.*

Démonstration. Si $\|A\| < 1$, alors $1 - A$ est inversible, d'inverse $\sum_{n=0}^{\infty} A^n$ (il s'agit d'une série normalement convergente dans un Banach). En conséquence, si $\lambda > \|A\|, \lambda - A = \lambda(1 - \frac{A}{\lambda})$ est inversible, donc $\sigma(A) \subset B(0, \|A\|)$.
Si A est inversible, et $B \in B(A, \frac{1}{\|A^{-1}\|})$, on a $\left\|1 - A^{-1}B\right\| = \left\|A^{-1}(A - B)\right\| \leqslant \left\|A^{-1}\right\|\|A - B\| < \left\|A^{-1}\right\|\frac{1}{\|A^{-1}\|} = 1$, donc $1 - (1 - A^{-1}B) = A^{-1}B$ est inversible, et donc également B. L'ensemble $\mathcal{B}(\mathcal{H})^{\times}$ des opérateurs inversibles est donc un ouvert de $\mathcal{B}(\mathcal{H})$. Soit $\varphi : \lambda \mapsto \lambda 1_{\mathcal{B}(\mathcal{H})} - A$. C'est une application continue de $\mathbb{C}$ dans $\mathcal{B}(\mathcal{H})$, on a $\varphi^{-1}(\mathcal{B}(\mathcal{H})^{\times}) = \mathbb{C} \setminus \sigma(A)$, est ouvert, et $\sigma(A)$ est fermé. Supposons pour la contradiction que $\sigma(A) = \varnothing$. R_A est alors définie sur $\mathbb{C}$ tout entier. Montrons d'abord que R_A est une fonction entière. Soit $\lambda_0 \in \mathbb{C}$. On a

$$\forall\ \lambda \in \mathbb{C} : \lambda - A = \lambda_0 - A - (\lambda_0 - \lambda) = (\lambda_0 - A)(1 - (\lambda_0 - \lambda)R_A(\lambda_0))$$

On voit que pour $|\lambda_0 - \lambda| < \frac{1}{\|R_A(\lambda_0)\|}$, $1 - (\lambda_0 - \lambda)R_A(\lambda_0)$ est inversible, avec

$$(1 - (\lambda_0 - \lambda)R_A(\lambda_0))^{-1} = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R_A(\lambda_0)^n$$

On obtient finalement

$$R_A(\lambda) = (\lambda - A)^{-1} = (1 - (\lambda_0 - \lambda)R_A(\lambda_0))^{-1}(\lambda_0 - A)^{-1} = \left(\sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R_A(\lambda_0)^n\right) R_A(\lambda_0)$$

Pour tout $\lambda \in B(\lambda_0, \|R_A(\lambda_0)\|^{-1})$. R_A est donc analytique sur $\mathbb{C}$. En particulier elle est continue, donc bornée sur $B_f(0, \|A\|)$, et de plus, pour $|\lambda| > \|A\|$

$$\left\|(\lambda - A)^{-1}\right\| = \frac{1}{|\lambda|}\left\|\sum_{n=0}^{\infty}\left(\frac{A}{\lambda}\right)^n\right\| \leqslant \frac{1}{|\lambda|}\sum_{n=0}^{\infty}\left\|\frac{A}{\lambda}\right\| = \frac{1}{|\lambda| - \|A\|} \xrightarrow{|\lambda|\to\infty} 0$$

R_A est donc bornée sur $\mathbb{C}$, et constante par une version du théorème de Liouville pour les espaces de Banach[Die60, p. 227]. Comme sa limite est 0, elle est identiquement nulle, ce qui constitue une contradiction, car on a défini R_A comme l'inverse d'un opérateur. □

Si $\mathcal{M} = M_n(\mathbb{C})$, $A \in M_n(\mathbb{C})$, on sait que A est non-inversible si et seulement si $\det(A) = 0$, et l'ensemble $\sigma(A) = \{\lambda : \det(\lambda I_n - A) = 0\}$, est l'ensemble des valeurs propres de A, et on retrouve la notion usuelle de spectre.

Si $\mathcal{M} = L^\infty(\Omega, \mathcal{F}, \mu)$ est l'ensemble des opérateurs de multiplication par une fonction essentiellement bornée sur $L^2(\Omega, \mathcal{F}, \mu)$, μ finie, et $T_f \in \mathcal{M}$, alors $\sigma(T_f)$ est l'*image essentielle* de f, l'ensemble

$$\text{im-ess}(f) = \{\lambda \in \mathbb{C} : \mu(f \in B(\lambda, \epsilon)) > 0 \forall \epsilon > 0\}$$

En effet, si $\lambda \notin \text{im-ess}(f)$, en posant $g = \frac{1}{\lambda - f}$. Prenons $\epsilon : \mu(f^{-1}(B(\lambda, \epsilon)) = 0 \iff |f - \lambda| > \epsilon \mu$-p.p. Alors on a $\|g\|_\infty \leqslant \frac{1}{\epsilon}$, et on a clairement $T_g \cdot T_f = T_f \cdot T_g = 1$. Réciproquement, si on prend $\lambda \in \text{im-ess}(f)$, en posant $g_\epsilon = \mathbb{1}_{f^{-1}(B(\lambda,\epsilon))}$, $g_\epsilon \neq 0\ \forall \epsilon$, $g \in L^2$, et on a

$$\|(\lambda - T_f)g_\epsilon\|^2 = \int_{\mathbb{C}} |(\lambda - f)g_\epsilon|^2 \mathrm{d}\mu \leqslant \epsilon^2 \mu(f^{-1}(B(\lambda, \epsilon))) = \epsilon^2 \|g_\epsilon\|^2$$

Il s'ensuit que T_f ne peut pas être inversible dans L^∞.

De même, on vérifie facilement que si $\mathcal{M} = \mathcal{C}(X, \mathbb{C})$, X compact, $\sigma(f) = f(X)$.

La structure de $\mathbb{C}$-algèbre est suffisante pour définir un **calcul fonctionnel** polynomial, c'est-à-dire, étant donné $A \in \mathcal{M}$, qu'il existe un morphisme d'algèbres $\Phi_A : \mathbb{C}[X] \mapsto \mathcal{M}$ qui envoie le polynôme X sur A, et qui est simplement le morphisme d'évaluation en A. Étant donné $\mathbb{C}[X] \ni P = \sum_{0 \leqslant k \leqslant n} a_k X^k$, $\lambda \in \mathbb{C}$, la factorisation :

$$P(X) - \lambda = a_n \prod_{k=0}^{n}(X - \lambda_k) \implies P(A) - \lambda = a_n \prod_{k=0}^{n}(A - \lambda_k)$$

implique que

$$\lambda \notin \sigma(P(A)) \iff \lambda_k \notin \sigma(A)\ \forall k \iff \lambda \notin P(\sigma(A))$$

En notant que chacun des termes commutent deux à deux, soit $P(\sigma(A)) = \sigma(P(A))$. Ce fait simple est un premier avatar du **théorème de l'image spectrale** dans sa version polynomiale.
On a également que si A est inversible, alors $\sigma(A^{-1}) = \{\frac{1}{\lambda}, \lambda \in \sigma(A)\}$. En effet, $\lambda - A^{-1} = \lambda A A^{-1} - A^{-1} = (A - \frac{1}{\lambda})\lambda A^{-1}$ est inversible exactement lorsque $(\frac{1}{\lambda} - A)$ l'est. Une conséquence de cette observation est la formule suivante :

Proposition 1.12. *Formule de Gelfand*[Con90, p. 197]
On considère la suite $(\|A^n\|)_{n \geqslant 1}$. *L'identité BA1 (sous-multiplicativité de la norme) implique, par le lemme de Fekete, que la limite* $\lim_{n\to\infty} \|A^n\|^{\frac{1}{n}} = \inf_{n \geqslant 1} \|A^n\|^{\frac{1}{n}}$ *est bien définie. De plus, on a*

$$\rho(A) := \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}} = \sup_{\lambda \in \sigma(A)} |\lambda|$$

Démonstration. La preuve de la proposition 1.2.1 montre que $\sup_{\lambda\in\sigma(A)}|\lambda| \leqslant \|A\|$. En applicant la remarque qui précède avec le polynôme X^n, on obtient

$$\sup|\sigma(A^n)| = \sup|\sigma(A)|^n \leqslant \|A^n\| \implies \sup|\sigma(A)| \leqslant \|A^n\|^{\frac{1}{n}} \implies \sup|\sigma(A)| \leqslant \rho(A)$$

Pour l'inégalité inverse, la preuve de la proposition 1.2.1 montre que R_A est une fonction holomorphe de λ sur $\mathbb{C}\setminus\sigma(A)$, et admet donc en particulier une série de Laurent convergant en norme d'opérateurs pour tout $|\lambda| > \sup|\sigma(A)|$,et qui doit coïncider avec l'expansion pour $|\lambda| \geqslant \|A\|$: $R_A(\lambda) = \sum_{n=1}^{\infty}\lambda^{-n}A^{n-1}$. En particulier on doit avoir :

$$\lim_{n\to\infty}\left\|\lambda^{-n}A^n\right\| = 0$$

pour $|\lambda| > \sup|\sigma(A)|$. Soit $\epsilon > 0$ alors on a $\|A^n\| \leqslant (\epsilon + \sup|\sigma(A)|)^n$ pour n suffisament grand, soit $\rho(A) \leqslant \epsilon + \sup|\sigma(A)|\ \forall\epsilon > 0$ □

Un des intérêts de la formule de Gelfand est qu'elle permet de faire le lien entre la structure métrique (la norme) et la structure algébrique (via l'ensemble des éléments inversibles) des algèbres d'opérateurs.

En effet, si $A \in \mathcal{A}$ est normal, l'identité

$$\|A\|^4 = \|A^*A\|^2 = \|(A^*A)^*(A^*A)\| = \left\|(A^*)^2A^2\right\| = \left\|(A^2)^*A^2\right\| = \left\|A^2\right\|^2 \implies \left\|A^2\right\| = \|A\|^2$$

implique par récurrence que

$$\left\|A^{2^n}\right\| = \|A\|^{2^n} \implies \rho(A) = \|A\| = \sup\{|\lambda| : \lambda - A \text{ n'est pas inversible}\}$$

Puis pour un élément quelconque on a

$$\|A\| = \sqrt{\|A^*A\|} = \sqrt{\sup\{|\lambda| : \lambda - A^*A \text{ n'est pas inversible}\}}$$

Une autre conséquence de cette formule est le fait que tout $*$-morphisme de sous-C*-algèbres unitaires d'opérateurs $\varphi : \mathcal{M} \mapsto \mathcal{N}$ est contractif. En effet, soit $A \in \mathcal{M}$. On a $\sigma(\varphi(A)) \subset \sigma(A)$, car l'image d'un inversible par un morphisme d'algèbres unitaires est inversible, et donc $\rho(\varphi(A)) \leqslant \rho(A)$.
Il s'ensuit que $\|\varphi(A)\|^2 = \|\varphi(A^*A)\| = \rho(\varphi(A^*A)) \leqslant \rho(A^*A) = \|A\|^2$

Montrons enfin les propriétés suivantes, qui montrent comment les propriétés algébriques des opérateurs se reflètent dans leur spectre :

Proposition 1.13. *i) A est auto-adjoint $\implies \sigma(A) \subset \mathbb{R}$*

ii) U est unitaire, $\lambda \in \sigma(U) \implies |\lambda| = 1$.

Démonstration. i) Soit A auto-adjoint, et $x + iy \in \sigma(A)$. Supposons que $y \neq 0$.
Alors l'élément $(x + iy - A) = y(i - (\frac{A-x}{y}))$ n'est pas inversible, et donc $i \in \sigma(\frac{A-x}{y})$, qui est auto-adjoint. Il suffit donc de montrer que pour B autoadjoint, $i \neq \sigma(B)$. Supposons que $i - B$ n'est pas inversible, et soit $n \in \mathbb{N}$. Alors on a $(n+1) - (n - iB) = 1 + iB = -i(i - B)$ est non-inversible, d'où $n + 1 \in \sigma(n - iB)$. Ceci donne :

$$(n+1)^2 \leqslant \|n - iB\|^2 = \|(n - iB)^*(n - iB)\| = \left\|n^2 + B^2\right\| \leqslant n^2 + \|B\|^2 \implies \|B\| \geqslant \sqrt{2n+1}$$

Contradiction, puisque B est borné et que n est arbitraire.

ii) Soit U unitaire. On a $\|U\| = 1$, donc $\rho(U) \leqslant 1$ et $\sigma(U) \subset B_f(0,1)$. Or on a aussi $\|U^*\| = 1$, et $\rho(U^*) \leqslant 1$. Comme U, U^* sont mutuellement inversibles, par image spectrale de l'inverse on a $\sigma(U) = \frac{1}{\sigma(U^*)}$ donc $|\sigma(U)| = \{1\}$.

□

Nous pouvons maintenant montrer le théorème de Gelfand, qui permet la définition du calcul fonctionnel continu. Celui-ci s'appuie sur la notion abstraite de spectre d'une C*-algèbre :

Définition 1.14. *Soit $\mathcal{A}$ une $\mathbb{C}$-algèbre de Banach. On appelle* ***spectre*** *de $\mathcal{A}$ l'ensemble $\hat{\mathcal{A}}$ des homorphismes continus d'algèbres à valeurs dans $\mathbb{C}$ non identiquement nuls. Autrement dit, $\hat{\mathcal{A}}$ consiste des formes linéaires continues ϕ telles que $\phi(xy) = \phi(x)\phi(y)$. Si $\mathcal{A}$ est unitaire $\phi \neq 0 \iff \phi(1) = 1$.*
$\hat{A}$ est un sous-ensemble du dual $\mathcal{A}'$, et en conséquence hérite de la topologie faible-$*$, telle que $\phi_n \to \phi \iff \phi_n(a) \to \phi(a)\ \forall a \in \mathcal{A}$.

Théorème 1.15. [Con90, p. 236] *Soit $\mathcal{A}$ une C*-algèbre unitaire commutative d'opérateurs. Alors $\hat{A}$ est compact, et $\mathcal{A} \cong \mathcal{C}(\hat{A})$ comme C*-algèbre, c'est-à-dire qu'il existe un $*$-isomorphisme isométrique de $\mathcal{A}$ dans $\mathcal{C}(\hat{A})$.*

Démonstration. Soit $\phi \in \hat{\mathcal{A}}$. Alors $\|\phi\| \leqslant 1$. En effet, si $a \in \mathcal{A} : \|a\| \leqslant 1$, on a $\phi(a)^n = \phi(a^n) \to 0$ par continuité de ϕ, d'où $|\phi(a)| < 1$, donc $\|\phi\| \leqslant 1$. D'autre part, $\hat{\mathcal{A}}$ est faible-$*$ fermé : si $\phi_n \xrightarrow{\text{faible}-*} \phi$, on a nécessairement $\phi(1) = \lim_n \phi(1) = 1$, et $\phi(xy) = \lim_n \phi_n(xy) = \lim_n \phi_n(x)\phi_n(y) = \phi(x)\phi(y)$. Donc $\hat{\mathcal{A}}$ est un fermé de la boule unité de $\mathcal{A}'$, qui est compact pour la topologie faible-$*$ par le théorème de Banach-Alaoglu, c'est donc un compact.

On considère, pour $a \in \mathcal{A}$, l'application $\hat{a} : \hat{\mathcal{A}} \mapsto \mathbb{C}$, définie par $\hat{a}(\phi) = \phi(a)$. Par définition de la topologie faible-$*$, c'est une application continue de $\hat{\mathcal{A}}$ dans $\mathbb{C}$. Soit $\gamma : \mathcal{A} \mapsto \mathcal{C}(\hat{\mathcal{A}}, \mathbb{C})$ l'application qui à a associe $\hat{a}$. Il est tout à fait manifeste que γ est un morphisme d'algèbres. Supposons que l'on aie, pour tout $a \in \mathcal{A}$, $\sigma(a) = \{\phi(a) : \phi \in \hat{\mathcal{A}}\}$. Alors on a en outre que $\phi(a^*) = \overline{\phi(a)}$ pour tout ϕ. En effet, en écrivant la décomposition de a en parties réelles et imaginaires auto-adjointes, on se ramène au cas où a est auto-adjoint, et à montrer $\phi(a) \in \mathbb{R}$, ce qui est immédiat, puisque $\phi(a) \in \sigma(a) \subset \mathbb{R}$.

Il s'ensuit que γ est un $*$-morphisme. De plus, c'est une isométrie : en effet,

$$\|\gamma(a)\|_\infty = \sup_{\phi \in \hat{\mathcal{A}}} |\phi(a)| = \rho(a) = \|a\|$$

puisque a est normal, comme tous les éléments d'une algèbre commutative.

$\gamma(\mathcal{A})$ est donc une sous-$*$-algèbre de $\mathcal{C}(\hat{\mathcal{A}}, \mathbb{C})$, contenant les constantes, car $\gamma(1) = 1$, et séparant les points, car si $\phi \neq \psi \in \hat{A} \implies \exists a \in \mathcal{A} : \phi(a) \neq \psi(a) \implies \gamma(a)(\phi) \neq \gamma(a)(\psi)$.

Le théorème de Stone-Weierstrass implique donc que $\gamma(\mathcal{A})$ est dense dans $\mathcal{C}(\hat{\mathcal{A}}, \mathbb{C})$. Comme l'image d'un fermé par une isométrie vers un espace complet est fermée, on conclut que γ est une isométrie surjective, donc bijective, donc un C*-isomorphisme.

Le lemme suivant achève donc la preuve du théorème. □

Lemme 1.16. *Soit $\mathcal{A}$ une C*-algèbre unitaire et commutative d'opérateurs. On a $\sigma(a) = \{\phi(a) : \phi \in \hat{\mathcal{A}}\}$.*

Démonstration. Soit $\phi \in \hat{\mathcal{A}}$. $\ker \phi$ est un idéal propre de $\mathcal{A}$ (puisque $\phi \neq 0$), et ne contient donc pas d'élément inversible. En particulier, comme $\phi(a) - a \in \ker \phi$, $\phi(a) - a$ est non-inversible et $\phi(a) \in \sigma(a)$, on a donc $\{\phi(a) : \phi \in \hat{\mathcal{A}}\} \subseteq \sigma(a)$.

Le point délicat est l'inclusion réciproque. Soit $\lambda \in \sigma(a)$. On considère l'idéal principal $\mathcal{I} = \mathcal{A}(\lambda - a)$, qui est propre, puisque $\lambda - a$ est non-inversible. Le cas commutatif du théorème de Krull affirme donc que $\mathcal{I}$ est inclus dans un idéal maximal (propre) $\mathcal{M}$ de $\mathcal{A}$.

De plus, observons que tout pour tout idéal propre $\mathcal{P}$, $\mathcal{P} \subset \mathcal{A} \setminus B(1_\mathcal{A}, 1)$, puisque tous les éléments de cette boule sont inversibles. En particulier, $\mathcal{A} \setminus B(1_\mathcal{A}, 1)$ étant fermé, l'adhérence $\overline{\mathcal{P}}$ est contenue

dans $\mathcal{A} \setminus B(1_{\mathcal{A}}, 1)$. Comme c'est un idéal, il est également propre, et on a en particulier que $\mathcal{M}$ est fermé, par maximalité.

On pose $\mathcal{B} = \mathcal{A}/\mathcal{M}$ c'est une $\mathbb{C}$-algèbre. De plus, par maximalité de $\mathcal{M}$, pour tout $a \notin \mathcal{M}$, l'idéal $(a) + \mathcal{M}$ est égal à $\mathcal{A}$, et en particulier contient l'élément 1, c'est-à-dire qu'il existe $b \in \mathcal{A}, m \in \mathcal{M}$: $ab + m = 1$, soit $[a][b] = [1]$ dans $\mathcal{B}$: tout élément non-nul de $\mathcal{B}$ est inversible. Autrement dit, $\mathcal{B}$ est un corps.

Posons $\|[a]\| = \|a + \mathcal{M}\| = \inf_{m\in\mathcal{M}} \|a - m\|$. C'est une norme sur $\mathcal{B}$, qui en fait une algèbre de Banach :

En effet, on a manifestement $\|\lambda(a + \mathcal{M})\| = \|\lambda a + \mathcal{M}\| = |\lambda|\|a + \mathcal{M}\|$, et

$$\|a + b + \mathcal{M}\| = \inf_{m\in\mathcal{M}} \|a + b - m\| \leqslant \inf_{m,m'\in\mathcal{M}} \big\|a + b - (m + m')\big\|$$
$$\leqslant \inf_{m,m'\in\mathcal{M}} \|a - m\| + \big\|b - m'\big\| = \|a + \mathcal{M}\| + \|b + \mathcal{M}\|$$

Supposons que $\|a + \mathcal{M}\| = 0$. Alors on a $\inf_{m\in\mathcal{M}} \|a - m\| = 0$, donc $a \in \overline{\mathcal{M}} = \mathcal{M}$, c'est-à-dire $a + \mathcal{M} = 0 + \mathcal{M}$.

De plus,

$$\|ab + \mathcal{M}\| = \inf_{m\in\mathcal{M}} \|ab - m\| \leqslant \inf_{m,m'\in\mathcal{M}} \big\|ab - mb - m'a + mm'\big\| = \inf_{m,m'\in\mathcal{M}} \big\|(a - m)(b - m')\big\|$$
$$\leqslant \inf_{m,m'\in\mathcal{M}} \|a - m\|\big\|b - m'\big\| = \|a + \mathcal{M}\|\|b + \mathcal{M}\|$$

Il reste à prouver que $\mathcal{B}$ est complet.

Soit donc $(a_n + \mathcal{M})_n$ une suite de Cauchy de $\mathcal{B}$. On peut extraire une sous-suite $(a_{n_k} + \mathcal{M})_k$ telle que

$$\big\|(a_{n_{k+1}} + \mathcal{M}) - (a_{n_k} + \mathcal{M})\big\| = \big\|(a_{n_{k+1}} - a_{n_k}) + \mathcal{M}\big\| \leqslant \frac{1}{2^k}$$

On pose $m_0 = 0 \in \mathcal{M}$. Pour chaque $k \geqslant 1$,

$$\exists p_k \in \mathcal{M} : \ \big\|(a_{n_{k+1}} - a_{n_k} - p_k)\big\| \leqslant \big\|a_{n_{k+1}} - a_{n_k} + \mathcal{M}\big\| + \frac{1}{2^k} < \frac{1}{2^{k-1}}$$

Puis, en posant $m_{k+1} = p_k + m_k$, on définit une suite d'éléments de $\mathcal{M}$ tels que

$$\forall k \geqslant 0, \big\|(a_{n_{k+1}} - m_{k+1}) - (a_{n_k} - m_k)\big\| < \frac{1}{2^{k-1}}$$

En particulier la suite $(a_{n_k} - m_k)$ est de Cauchy dans $\mathcal{A}$, donc converge vers une limite $x \in \mathcal{A}$. Par continuité de $x \mapsto x + \mathcal{M}$ ($\|x + \mathcal{M}\| \leqslant \|x - 0\|$), on a $a_{n_k} - m_k + \mathcal{M} = a_{n_k} + \mathcal{M} \to x + \mathcal{M}$. On a montré que toute suite de Cauchy dans $\mathcal{B}$ a une sous-suite convergente, ce qui montre que $\mathcal{B}$ est complet, et finalement une algèbre de Banach.

Soit $b \in \mathcal{B}$. Alors $\sigma \neq \varnothing \implies \exists \eta(b) \in \mathbb{C}$ tel que $\eta(b)1_{\mathcal{B}} - b$ est non-inversible dans $\mathcal{B}$. Comme ce dernier est un corps, on a forcément $\eta(b)1_{\mathcal{B}} - b = 0$, et $\sigma(b) = \{\eta(b)\}$. L'application $\eta : b \mapsto \eta(b)$ est un morphisme d'algèbres de $\mathcal{B}$ dans $\mathbb{C}$, qui plus est injectif comme tout morphisme de corps : en effet, si $b, b' \in \mathcal{B}, \mu \in \mathbb{C}$, on a

$$\mu b - \mu\eta(b)1_{\mathcal{B}} = 0, b' - \eta(b')1_{\mathcal{B}} = 0 \implies (\mu b + b') - (\mu\eta(b) + \eta(b'))1_{\mathcal{B}} = 0$$
$$\implies \mu\eta(b) + \eta(b') \in \sigma(\mu b + b') = \{\eta(\mu b + b'\}$$

De même

$$b = \eta(b)1_{\mathcal{B}}, b' = \eta(b')1_{\mathcal{B}} \implies bb' - \eta(b)\eta(b')1_{\mathcal{B}} = 0 \implies \eta(bb') = \eta(b)\eta(b')$$

Le premier théorème d'isomorphisme assure que $\mathcal{B}$ est isomorphe à son image, c'est-à-dire $\mathbb{C}$, puisque $\eta(1_{\mathcal{B}}) = 1$.

Enfin on considère le morphisme d'algèbres $\phi : \mathcal{A} \mapsto \mathbb{C}$ défini par $\phi(a) = \eta(a + \mathcal{M})$, de noyau $\mathcal{M} \neq \mathcal{A}$, donc $\phi \in \hat{\mathcal{A}}$, et d'autre part, $\lambda - a \in \mathcal{I} \subset \mathcal{M}$, donc on $\phi(\lambda - a) = 0 \implies \phi(a) = \lambda$. On a finalement $\sigma(a) \subset \{\phi(a), \phi \in \hat{\mathcal{A}}\}$, comme voulu. □

Dans le cas où $\mathcal{A} = \mathrm{C}^*\langle 1, A\rangle$ est la C*-algèbre unitaire engendrée par un élément normal A, qui est commutative, on a un **calcul fonctionnel continu**, c'est-à-dire un $*$-morphisme unitaire et isométrique $\mathcal{C}(\sigma(A), \mathbb{C}) \xrightarrow{\Phi_A} \mathcal{A}$ tel que $\nu(\mathrm{id}_{\sigma(A)}) = A$.

Le théorème de Gelfand fournit un $*$-morphisme unitaire et isométrique entre $\mathcal{A}$ et $\mathcal{C}(\hat{A}, \mathbb{C})$. On considère l'application $\hat{A} : \hat{\mathcal{A}} \mapsto \sigma(A), \phi \mapsto \phi(A)$. C'est une application continue, surjective par le lemme précédent. C'est en fait un homéomorphisme : comme toute bijection continue $E \mapsto F$ où E est compact et F est séparé est un homéomorphisme (puisque c'est une bijection fermée), il suffit de montrer l'injectivité. Supposons que $\phi(A) = \psi(A)$. Alors pour tout polynôme $P \in \mathbb{C}\langle X, X^*\rangle$, on a $\phi(P(A)) = \psi(P(A))$ or $\{P(A), P \in \mathbb{C}\langle X, X^*\rangle\}$ est dense dans $\mathrm{C}^*\langle 1, A\rangle$, donc on a $\phi = \psi$ par prolongement.

Cet homéomorphisme induit à son tour un $*$-isomorphisme unitaire $\mathcal{C}(\sigma(A), \mathbb{C}) \xrightarrow{\sim} \mathcal{C}(\hat{\mathcal{A}}, \mathbb{C})$, donné par $f \mapsto f \circ \hat{A}$. Par composition, on obtient bien un $*$-isomorphisme Φ_A comme voulu. On note $f(A)$ l'élément $\Phi_A(f)$, et on a donc les relations $f(A) + \mu g(A) = (f + \mu g)(A), f(A)g(A) = (fg)(A), \overline{f}(A) = f(A)^*$. De plus, on a $f(A)^* f(A) = \overline{f}(A)f(A) = |f|^2(A) = f(A)f(A)^*$, donc $f(A)$ est normal.

On a encore la version continue du théorème de l'application spectrale : si $\mathcal{A}$ est une C*-algèbre unitaire d'opérateurs, et $A \in \mathcal{A}$ est normal, alors $\forall f \in \mathcal{C}(\sigma(A), \mathbb{C})$ on a $f(\sigma(A)) = \sigma(f(A))$.

En appliquant pour chaque $\lambda \in \mathbb{C}$ la translation affine $\lambda - X$, on peut se ramener par image spectrale polynomiale à montrer que $f(A)$ est non-inversible ($0 \in \sigma(f(A))$) si et seulement si $0 \in f(\sigma(A))$.

Si $0 \notin f(\sigma(A))$, $g : z \mapsto \frac{1}{f(z)}; \sigma(A) \mapsto \mathbb{C}$ est continue, donc $f(A)g(A) = (fg)(A) = 1$ et $f(A)$ est inversible.

Réciproquement, procédons par l'absurde en supposant que $f(A)$ soit inversible, mais $0 \in f(\sigma(A))$, disons $f(\lambda_0) = 0$. Soit $\alpha > \left\|f(A)^{-1}\right\|$. On peut choisir une fonction continue h vérifiant, $\|h\|_\infty = \alpha$ et $\|hf\|_\infty \leqslant 1$. En effet soit δ tel que $f(B(\lambda_0, \delta)) \subset]-\frac{1}{\alpha}, \frac{1}{\alpha}[$, et considérons la fonction en tipi $h(\lambda) = \alpha(1 - \frac{2|\lambda - \lambda_0|}{\delta})^+$. On a bien $\|h\|_\infty = \alpha$ car $\lambda_0 \in \sigma(A)$). Comme le calcul fonctionnel est isométrique, on a : $\alpha = \|h\|_\infty = \|h(A)\| = \left\|f^{-1}(A)f(A)h(A)\right\| \leqslant \left\|f^{-1}(A)\right\|\|f(A)h(A)\| < \alpha$: une contradiction.

Le calcul fonctionnel est source de toutes sortes de décompositions. On peut montrer que tout élément s'écrit comme combinaison linéaire de quatres éléments de spectre positif, en remarquant que tout élément auto-adjoint a s'écrit

$$a = \left(\frac{1+a}{2}\right)^* \left(\frac{1+a}{2}\right) - \left(\frac{1-a}{2}\right)^* \left(\frac{1-a}{2}\right)$$

Puis en écrivant la décomposition en parties réelles et imaginaires, et en utilisant la proposition suivante :

Proposition 1.17. *Éléments positifs*

Soit $\mathcal{A}^+ = \{A \in \mathcal{A} : \sigma(A) \subset \mathbb{R}_+,\ A = A^*\}$. Alors $\forall \lambda, \mu > 0, A, B \in \mathcal{A}^+$, on a $\lambda A + \mu B \in \mathcal{A}^+$ ($\mathcal{A}^+$ est un cône convexe). De plus on a l'égalité $\mathcal{A}^+ = \{A^* A, A \in \mathcal{A}\}$, si bien que la propriété de positivité correspond exactement à celle d'être auto-adjoint à spectre positif.

Démonstration. Notons d'abord que pour $A = A^*$ et $\lambda \geqslant \|A\| = \sup|\sigma(A)|$, on a $\|\lambda - A\| = \sup\sigma(\lambda - A) = \lambda - \inf\sigma(A)$ par image spectrale polynomiale.

Puis, en notant que pour $A, B \in \mathcal{A}$ auto-adjoints et $\lambda \geqslant 2(\|A\| \vee \|B\|)$, on a :

$$\left\|\frac{\lambda}{2} - A\right\| = \frac{\lambda}{2} - \inf\sigma(A), \left\|\frac{\lambda}{2} - B\right\| = \frac{\lambda}{2} - \inf\sigma(B), \|\lambda - (A+B)\| = \lambda - \inf\sigma(A+B)$$

$$\implies \inf\sigma(A+B) - \inf\sigma(A) - \inf\sigma(B) = \left\|\frac{\lambda}{2} - A\right\| + \left\|\frac{\lambda}{2} - B\right\| - \|\lambda - (A+B)\| \geqslant 0$$

Or $A, B \in \mathcal{A}^+, \lambda, \mu > 0 \implies \inf\sigma(\lambda A), \inf\sigma(\mu B) \geqslant 0 \implies \inf\sigma(\lambda A + \mu B) \geqslant 0 \implies \sigma(\lambda A + \mu B) \subset \mathbb{R}_+$. Comme de plus $\lambda A + \mu B$ est auto-adjoint, il s'ensuit que $\mathcal{A}^+$ est un cône convexe.

Notons également la formule suivante lorsque $\lambda - AB$ est inversible, pour $\lambda \neq 0$, alors $\lambda - BA$ est inversible et on a

$$(\lambda - BA)^{-1} = \frac{1}{\lambda}(1 + B(\lambda - AB)^{-1}A)$$

Ceci implique en particulier que $\sigma(AB) \cup \{0\} = \sigma(BA) \cup \{0\}$.

Soit $A \in \mathcal{A}$, et montrons que $A^*A \in \mathcal{A}^+$. Clairement, A^*A est auto-adjoint. D'autre part, en considérant les fonctions continues sur $\mathbb{R}$ $x \mapsto x^+ \geqslant 0, x \mapsto x^- \geqslant 0$, et en notant $x^+x^- = 0 = x^-x^+, x^+ - x^- = x$, on obtient par calcul fonctionnel continu une décomposition $A^*A = (A^*A)^+ - (A^*A)^- := X - Y$, $X, Y \in \mathcal{A}^+$. Posons $B = AY$. On a $B^*B = YA^*AY = Y(X-Y)Y = -Y^3 \in -\mathcal{A}^+$

En particulier, $\sigma(B^*B) \subset \mathbb{R}_- \implies \sigma(BB^*) \subset \mathbb{R}_-$. Puis, $B^*B + BB^* \in -\mathcal{A}^+$, comme ce dernier est un cône. En remarquant que $B^*B + BB^* = (F - iG)(F + iG) + (F + iG)(F - iG) = 2(F^2 + G^2)$, où F, G sont les parties réelles et imaginaires auto-adjointes de B^*B, on a $F^2, G^2 \in \mathcal{A}^+$ par calcul fonctionnel, ce qui implique que

$$B^*B \in \mathcal{A}^+ \cup -\mathcal{A}^+ \implies \sigma(B^*B) \subset \{0\} \implies B = 0 \implies Y = 0 \implies A = X \in \mathcal{A}^+$$

Réciproquement, si $A \in \mathcal{A}^+$, A est auto-adjoint et par calcul fonctionnel continu, il existe un élément auto-adjoint $\sqrt{A}$ tel que $\sqrt{A}^2 = \sqrt{A}^*\sqrt{A} = A$. □

La structure de C*-algèbre est donc propice à la construction d'un calcul fonctionnel continu. On va voir à présent que la structure d'algèbre de von Neumann est propice à la construction d'un calcul fonctionnel Borélien. Pour satisfaire à cette fin, recueillons d'abord quelques propriétés des opérateurs de $\mathcal{B}(\mathcal{H})$, qui motivent par ailleurs la terminologie.

Définition 1.18. *Un **projecteur** est un opérateur de $\mathcal{B}(\mathcal{H})$ auto-adjoint et idempotent. Comme $P = P^2$, $\ker(1-P) = P\mathcal{H}$, l'image d'un projecteur est donc un sous-espace fermé, et on a toujours $\mathcal{H} \cong P\mathcal{H} \oplus (1-P)\mathcal{H}$.*

Comme P est auto-adjoint et en particulier normal, par image spectrale polynomiale, on a $\sigma(P) = \sigma(P)^2 \implies \sigma(P) \subset \{0, 1\}$ Réciproquement, pour tout sous-espace fermé $E \subset \mathcal{H}$, il existe un projecteur P_E tel que $P_E\mathcal{H} = E$. Cette correspondance établit une bijection entre les sous-espaces fermés de $\mathcal{H}$ et les projecteurs de $\mathcal{B}(\mathcal{H})$.

*Une **isométrie partielle** A est un opérateur tel que A^*A et AA^* sont des projecteurs.*

Proposition 1.19. [Pet13] *Soit $A \in \mathcal{B}(\mathcal{H})$.* **a)** *$A$ est normal $\iff \|Ax\| = \|A^*x\|$ $\forall x$*
b) *A est auto-adjoint $\iff \langle Ax, x\rangle \in \mathbb{R}$ $\forall x$*
c) *A est positif $\iff \langle Ax, x\rangle \geqslant 0$ $\forall x$*
d) *A est isométrique (au sens de la définition 1.1.2) $\iff \|Ax\| = \|x\|$ $\forall x$*
e) *A^*A et AA^* sont des projecteurs $\iff$ il existe $E \subset \mathcal{H}$ un sous-espace fermé tel que $A_{|E}$ est une isométrie et $A_{|E^\perp} = 0$, d'où l'appelation d'isométrie partielle.*

Démonstration. Rappelons l'identité de polarisation pour la forme sesquilinéaire

$$(x,y) \mapsto \langle Ax, y\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \langle A(x+i^k y), x+i^k y\rangle$$

a)Si A est normal $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle = \langle AA^*x, x\rangle = \langle A^*x, A^*x\rangle = \|A^*\|^2$. Réciproquement, si $\|Ax\| = \|A^*x\|\ \forall x$, on a $\langle A^*Ax, x\rangle = \langle A^*Ax, x\rangle \implies \langle (A^*A - AA^*)x, x\rangle = 0$ et par polarisation, on a $\langle (A^*A-AA^*)x, y\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \langle (A^*A-AA^*)(x+i^k y), x+i^k y\rangle = 0\ \forall x, y \implies A^*A = AA^*$.
a) et **d)** se déduisent de manière tout à fait analogue.
c) Si $A = B^*B$, $\langle Ax, x\rangle = \langle B^*Bx, x\rangle = \|Bx\|^2 \geqslant 0$. Réciproquement, si $\langle Ax, x\rangle \geqslant 0 \forall x$, on montre que $\sigma(A) \subset [0, \infty[\iff \forall \lambda > 0, A + \lambda$ a un inverse dans $\mathcal{B}(\mathcal{H})$. Or, on a $\lambda\|x\|^2 \leqslant \langle (A+\lambda)x, x\rangle \leqslant \|(A+\lambda)x\|\|x\| \implies \lambda\|x\| \leqslant \|(A+\lambda)x\|$, par hypothèse et Cauchy-Schwartz. L'inégalité $\langle (A+\lambda)x, x\rangle \geqslant a\|x\|^2$ montre que $A+\lambda$ est injective, d'image dense, et $\lambda\|x\| \leqslant \|(A+\lambda)x\|$ montre que cette image est fermée. Donc $A+\lambda$ est bijective. En remplacant x par $(\lambda + A)^{-1}x$, on obtient $\lambda\|(A+\lambda)^{-1}x\| \leqslant \|x\|$ donc $(A+\lambda)^{-1} \in \mathcal{B}(\mathcal{H})$, comme voulu.
e) Si AA^* et A^*A sont des projecteurs, en notant $E = A^*A\mathcal{H}$, on a $\langle Ax, Ax\rangle = \langle AA^*Ax, Ax\rangle = \langle A^*Ax, A^*Ax\rangle = \langle x, x\rangle\ \forall x \in E$, donc $A_{|E}$ est une isométrie. Pour $x \in E^\perp$, on a aussi $\langle Ax, Ax\rangle = \langle A^*Ax, x\rangle = 0$, donc $A_{|E^\perp} = 0$. Réciproquement, si $A_{|E}$ est une isométrie et $A_{|E^\perp} = 0$, soit $x \in E^\perp$. Alors on a $\langle A^*Ax, A^*Ax\rangle = \langle Ax, AA^*Ax\rangle = 0$, et si $x \in E$, on a, en notant $x_E, x_{E^\perp}$ les projections orthogonales sur $E, E^\perp$, $\langle Ax, Ax\rangle = \langle A(x_E + x_{E^\perp}), A(x_E + x_{E^\perp})\rangle = \langle Ax_E, Ax_E\rangle = \langle x_E, x_E\rangle$ comme $A_{|E}$ est une isométrie et $A_{|E^\perp} = 0$, ceci veut dire $\langle A^*Ax, x\rangle = \langle x_E, x_E\rangle\ \forall x$ et on conclut par polarisation que $A^*Ax = x_E\ \forall x$. □

Définition 1.20. *Soit X un espace topologique compact (qu'on munit de sa tribu borélienne $B(X)$). Une **mesure spectrale** μ sur X est une application $B(X) \mapsto \mathcal{B}(\mathcal{H})$ telle que :*
$\mu(A)$ est un projecteur $\forall A \in B(X)$
$\mu(\varnothing) = 0, \mu(X) = 1$
$\mu(A \cap B) = \mu(A)\mu(B)\ \forall A, B \in B(X)$
$\forall\ x, y \in \mathcal{H}$, $A \mapsto \langle \mu(A)x, y\rangle$ est une mesure de Radon (complexe) finie sur X, qu'on notera $\mu_{x,y}$. On voit en particulier que $\mu_{x,x}$ est une mesure de Radon réelle, et que $|\mu_{x,y}|(X) \leqslant \|x\|\|y\|$.

On observe que $(x,y) \mapsto \langle \mu(A)x, y\rangle$ est une forme sesqui-linéaire continue. En prenant f bornée et borélienne comme limite croissante de fonctions simples, on a

$$(x,y) \mapsto \int_X f \mathrm{d}\mu_{x,y}$$

Est une forme sesquilinéaire continue, et par représentation de Riesz, il existe un unique opérateur $T \in \mathcal{B}(\mathcal{H})$ tel que $\int_X f\mathrm{d}\mu_{x,y} = \langle Tx, y\rangle\ \forall x, y$. On notera

$$T = \int_X f \mathrm{d}\mu$$

On peut prouver, grâce au lemme suivant, un analogue du théorème de convergence monotone :

Lemme 1.21. *Soit $(A_n)_n$ une suite croissante d'opérateurs positifs, bornée uniformément : $M := \sup_n \|A_n\| < \infty$. Alors il existe un opérateur borné A tel que $A_n \to A$ pour la topologie faible d'opérateurs.*

Démonstration. On considère la fonction $x \mapsto \lim_n \langle A_n x, x\rangle$, qui est ponctuellement bien définie puisque cette suite est croissante et bornée par $M\|x\|\|y\|$. Par polarisation, on a aussi $(x,y) \mapsto \lim_n \langle A_n x, y\rangle$ qui est bien définie, et qui est de plus une forme sesquilinéaire. Par le théorème de représentation de Riesz pour les formes sesquilinéaires, il existe un opérateur $A \in \mathcal{B}(\mathcal{H})$ tel que $\lim_n \langle A_n x, y\rangle = \langle Ax, y\rangle \ \forall x, y \in \mathcal{H}$, donc A est la limite faible de A_n. □

Proposition 1.22. *L'application $f \mapsto \int_X f\, d\mu$ est un $*$-morphisme continu de la C*-algèbre des fonctions boréliennes bornées sur X vers $\mathcal{B}(\mathcal{H})$.*

De plus, on a un résultat de convergence monotone : si f_n est une suite croissante de fonctions bornées telle que $\sup_n f_n$ est bornée, alors on a la convergence faible-opérateur :

$$\int_X f_n \, d\mu \xrightarrow{w} \int_X f \, d\mu$$

On arrive enfin qui s'appuie sur le théorème suivant, généralisation du théorème spectral bien connu pour les matrices normales de $M_n(\mathbb{C})$.

Théorème 1.23. [Con90, p. 259]

Soit $\mathcal{A}$ une C-algèbre commutative et unitaire d'opérateurs. Alors il existe une unique mesure spectrale μ_A sur $\sigma(\mathcal{A}) = \hat{\mathcal{A}}$ telle que*

$$\forall A \in \mathcal{A}, A = \int_{\sigma(\mathcal{A})} \gamma(A) \, d\mu_A$$

Où γ est la transformée de Gelfand $A \mapsto \hat{A} : \phi \mapsto \phi(A)$.

La preuve est essentiellement formelle, il s'agit de jouer sur les différentes représentations des formes linéaires et sesquilinéaires fournies par les théorèmes de Riesz.

Démonstration. Pour $x, y \in \mathcal{H}$ fixés, l'application $f \mapsto \langle \gamma^{-1}(f)x, y\rangle$ de $\mathcal{C}(\hat{\mathcal{A}})$ dans $\mathbb{C}$ est une forme linéaire continue (car γ^{-1} est une isométrie linéaire, et par Cauchy-Schwartz). Par le théorème de Riesz-Markov, il existe une unique mesure de Radon complexe $\mu_{x,y}$ sur $\hat{\mathcal{A}}$ telle que $\forall f \in \mathcal{C}(\hat{\mathcal{A}}, \mathbb{C}), \langle \gamma^{-1}(f)x, y\rangle = \int_{\hat{A}} f \mathrm{d}\mu_{x,y}$.

Notons qu'en appliquant le théorème de Riesz à la forme positive $f \mapsto \langle \gamma^{-1}(f)x, x\rangle$, on obtient que $\mu_{x,x}$ est une mesure réelle.

Soit à présent B un borélien de $\hat{\mathcal{A}}$, et considérons l'application

$$(x,y) \mapsto \mu_{x,y}(B)$$

Le théorème de Riesz-Markov assure par ailleurs que $|\mu_{x,y}|(\hat{\mathcal{A}}) = \left\| f \mapsto \langle \gamma^{-1}(f)x, y\rangle \right\| \leqslant \|x\|\|y\|$. L'application ci-dessus est donc une forme sesquilinéaire (par unicité de $\mu_{x,y}$) continue, et le théorème de représentation de Riesz donne l'existence d'un unique opérateur borné $\mu(B)$ vérifiant

$$\langle \mu(B)x, y\rangle = \mu_{x,y}(B) \ \forall x, y \in \mathcal{H}$$

Il est clair que $\mu(\varnothing) = 0$ et $\langle \mu(\hat{\mathcal{A}})x, y\rangle = \mu_{x,y}(\hat{\mathcal{A}}) = \langle \gamma^{-1}(1)x, y\rangle = \langle x, y\rangle$.
D'autre part, on a $\langle \mu(B)x, x\rangle = \mu_{x,x}(B) \in \mathbb{R}$, donc $\mu(B)$ est auto-adjoint, et de plus, l'égalité $\langle \gamma^{-1}(g)\mu(B)x, y\rangle = \int_X g \mathbb{1}_B \mathrm{d}\mu_{x,y}$ montre que $\mathrm{d}\mu_{\mu(B)x,y} = \mathbb{1}_B \mathrm{d}\mu_{x,y}$, donc on a

$$\langle \mu(C)\mu(B)x, y\rangle = \int_{\hat{\mathcal{A}}} \mathbb{1}_C \mathbb{1}_B \mathrm{d}\mu_{x,y} = \int_{\hat{\mathcal{A}}} \mathbb{1}_{C\cap B} \mathrm{d}\mu_{x,y} = \langle \mu(C \cap B)x, y\rangle \ \forall x, y \in \mathcal{H}$$

Soit $\mu(C)\mu(B) = \mu(C \cap B)$. En particulier $\mu(B)^2 = \mu(B)$, donc $\mu(B)$ est un projecteur et μ est une mesure spectrale, dont l'unicité est garantie par sa construction.

Enfin, on a par définition que $\tilde{A} := \int_{\sigma(\mathcal{A})} \gamma(A)\mathrm{d}\mu$ est l'unique opérateur vérifiant $\langle \tilde{A}x, y\rangle = \int_{\hat{\mathcal{A}}} \gamma(A)\mathrm{d}\mu_{x,y} = \langle \gamma^{-1}(\gamma(A))x, y\rangle = \langle Ax, y\rangle \ \forall x, y$. On appelle μ la **mesure spectrale** associée à A. □

Ce théorème est à la base du *calcul fonctionnel borélien*. Si A est un opérateur normal, alors on peut montrer[Con90, p. 264] que l'application $f \mapsto \int_{\sigma(A)} f\mathrm{d}\mu_A$ est un $*$-morphisme continu de la C*-algèbre des fonctions boréliennes bornées sur $\sigma(A)$ vers $\mathcal{B}(\mathcal{H})$, continu pour la topologie faible d'opérateurs. Ceci montre en particulier que toute algèbre de von Neumann contient les fonctions boréliennes de ses éléments normaux.

Regardons le cas où $A \in \mathcal{B}(\mathbb{C}^n) \cong M_n(\mathbb{C})$ est un opérateur normal. Dans ce cas, $\sigma(A)$ est l'ensemble discret de ses valeurs propres, et μ_A est l'unique mesure spectrale sur $\sigma(A)$ vérifiant

$$A = \sum_{\lambda\in\sigma(A)} \lambda\mu_A(\{\lambda\})$$

En prenant $\mu_A(\{\lambda\}) = P(V_\lambda)$ l'opérateur de projection orthogonale sur l'espace propre V_λ associé à λ, on voit que c'est bien une mesure spectrale vérifiant cette équation. Pour f une fonction borélienne définie sur $\sigma(A)$, on définit donc $f(A)$ comme l'opérateur

$$f(A) = \sum_{\lambda\in\sigma(A)} f(\lambda)P(V_\lambda)$$

Du point de vue calculatoire, ceci correspond à l'opération de diagonaliser la matrice A dans une base orthogonale, appliquer f sur la diagonale, puis appliquer la conjugaison inverse.

Si A est un opérateur auto-adjoint d'une algèbre de von Neumann $\mathcal{M}$, en approchant par en-dessous la fonction $\lambda \mapsto \lambda$ sur $\sigma(A)$ par une suite croissante de fonctions simples, qui correspondent à une combinaison linéaire finie de projecteurs, on voit que A est contenue dans l'adhérence faible du sous-espace engendré par les projecteurs de $\mathcal{M}$. La même chose reste vraie pour n'importe quel opérateur B, en le décomposant en deux parties auto-adjointes. Il s'ensuit que la donnée des projecteurs contenus dans $\mathcal{M}$ suffit à décrire entièrement $\mathcal{M}$. Ce fait, et l'interprétation géométrique qui suit, nous pousse à concentrer notre regard sur l'ensemble des projecteurs de $\mathcal{M}$.

Définition 1.24. *Soit $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ une algèbre de von Neumann. On note $P(\mathcal{M})$ l'ensemble des projecteurs de $\mathcal{M}$.*

Pour A, B auto-adjoints, on définit la relation $\leqslant$ par $A \leqslant B \iff B - A$ est positif Si $(P_\alpha)_{\alpha\in I}$ est une famille de projections, on note

$$\bigvee_{\alpha\in I} P_\alpha = P\left(\overline{\sum_{\alpha\in I} E_\alpha}\right)$$

$$\bigwedge_{\alpha\in I} P_\alpha = P(\bigcap_{\alpha\in I} E_\alpha)$$

Où E_α est le sous-espace fermé tel que P_α soit le projecteur orthogonal sur E_α, et $P(E)$ désigne la projection sur E pour un sous-espace fermé E.

Comme on a $\left(\overline{\sum_{\alpha\in I} E_\alpha}\right)^\perp = \bigcap_{\alpha\in I} E_\alpha^\perp$, il s'ensuit que

$$P\left(\overline{\sum_{\alpha\in I} E_\alpha}^\perp\right) = 1 - \bigvee_{\alpha\in I} P_\alpha = P\left(\bigcap_{\alpha\in I} E_\alpha^\perp\right) = \bigwedge_{\alpha\in I} (1 - P_\alpha)$$

Si $\mathcal{M} = \mathcal{B}(\mathcal{H})$, l'ordre sur les projecteurs est exactement l'ordre induit par la correspondance entre sous-espaces fermés et projecteurs, et $P(\mathcal{B}(\mathcal{H}))$ est un **treillis**, c'est-à-dire un ensemble partiellement ordonné tel que chaque paire d'éléments a un plus grand minorant et un plus petit minorant. C'est même un **treillis complet** tel que chaque partie a une borne supérieure et une inférieure, et les opérateurs $\wedge, \vee$ associent précisément une famille $(P_\alpha)_\alpha \xrightarrow{\sim} (E_\alpha)_\alpha$ à leurs bornes, respectivement inférieures et supérieures.

Si $\mathcal{M}$ est une sous-algèbre propre de $\mathcal{B}(\mathcal{H})$, il n'est pas dit *a priori* que ces bornes appartiennent encore à $\mathcal{M}$. C'est une conséquence du lemme 1.20 qui implique que si $(P_n)_n$ est une suite de projecteurs deux-à-deux orthogonaux, la somme $\sum_n P_n$ est bien définie par la limite de ses sommes partielles pour la topologie faible d'opérateur, et que c'est un projecteur, puisque $P(\mathcal{M})$ est fermé pour la topologie faible. La même preuve s'applique en prenant des suites généralisées dans le lemme et comme définition de la topologie faible d'opérateurs, ce qui montre, comme la suite $(\vee_J P_\alpha)_{J\subset I}$ est croissante, que $P(\mathcal{M})$ est un treillis complet (la borne inférieure suit par passage à l'orthogonal).

Pour voir que $P(\mathcal{M})$ est faiblement fermé, il suffit de remarquer que

$$P(\mathcal{M}) = \mathcal{M} \cap \bigcap_{x\in\mathcal{H}} f_x^{-1}(\{0\}) \cap g_x^{-1}(\{0\})$$

où $f_x : A \mapsto \langle (A^* - A)x, x\rangle, g_x : A \mapsto \langle Ax, (A-1)x\rangle$ sont continues pour la topologie faible d'opérateurs.

Si $\mathcal{H}$ est séparable, dans $\mathcal{B}(\mathcal{H})$, on peut toujours, pour calculer les bornes, se ramener au cas où I est dénombrable. En effet, si $\{x_n, n \in \mathbb{N}\}$ est dense dans $\mathcal{H}$, $\{x_n, n \in \mathbb{N}\} \cap \overline{\sum_I E_\alpha}$ est dense dans $\overline{\sum_I E_\alpha}$, qui est donc égal à $\overline{\sum_n \mathbb{C}x_n}$. Le cas de la borne inférieure suit en considérant les supplémentaires orthogonaux comme ci-dessus. Comme $\mathcal{H}$ admet une base hilbertienne dense, on peut même prendre les espaces deux à deux orthogonaux.

Gardons ceci à l'esprit lorsque nous pensons à $\mathcal{M} = L^\infty(X, \mathcal{F}, \mu) \subset \mathcal{B}(L^2(X, \mathcal{F}, \mu))$. Dans ce cas les projecteurs sont exactement les opérateurs de multiplication T_f, où f est mesurables, et tels que $T_f^2 = T_f$. Or $T_f^2 = T_{f^2}$, ce qui impose $f^2 = f$ d'où $f(X) \subset \{0, 1\}$. Autrement dit, les projecteurs de $\mathcal{M}$ sont exactement les fonctions indicatrices de $\mathcal{F}$, avec $P_A := T_{\chi_A}$ pour $A \in \mathcal{F}$ associé au sous-espace fermé de L^2, $\{f : f_{|X\setminus A} = 0\}$. Dans ce cas on a précisément $P_{A\wedge B} = P_{A\cap B}, P_{A\vee B} = P_{A\cup B}$, et on voit, que pour ne pas rencontrer de problèmes de mesurabilité, il est sans doute plus sage de se restreindre au cas où $\mathcal{H}$ est séparable, ce qui est en particulier le cas si X est polonais.

Définition 1.25. *Soit $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ une algèbre de von Neumann. Deux projecteurs $P, Q \in P(\mathcal{M})$ sont dits **von Neumann-Murray équivalents**, et on note $P \sim Q$, si il existe une isométrie partielle $I \in \mathcal{M}$ telle que $I^*I = P$, $II^* = Q$. Moralement, ceci signifie que $\mathcal{M}$ reconnaît que les espaces $P\mathcal{H}$ et $Q\mathcal{H}$ sont géométriquement les mêmes.*

Si E et F sont des sous-espaces de dimension finie n dans $\mathcal{H}$, munissons les des bases hilbertiennes $(e_i)_{1\leqslant i\leqslant n}$, et $(f_i)_{1\leqslant i\leqslant n}$. Alors l'application $x = P(E^\perp) + \sum_{i=1\ldots n} \alpha_i e_i \mapsto \sum_{i=1\ldots n} \alpha_i f_i$ est une isométrie partielle de E sur F. On peut considérer que, géométriquement, les espaces E et F sont *les mêmes*, et on voit qu'ils sont précisément von Neumann-Murray équivalents. L'idée de von Neumann et Murray est d'utiliser cette identification pour comprendre la structure du treillis $P(\mathcal{M})$. En effet, on peut montrer que cette relation est une relation d'équivalence, et que l'ordre partiel sur $P(\mathcal{M})$ passe au quotient pour donner un ordre partiel sur $P(\mathcal{M})/\sim$. De surcroît, si $\mathcal{M}$ est un facteur, alors on peut montrer que cet ordre est total. Ceci conduit à classifier les facteurs en fonction de la nature de cet ordre. Il s'agit d'un vaste sujet, qui pourrait nous emmener très loin, et qui est développé dans [Sun87]. Pour l'heure, laissons de côté les algèbres de von Neumann,

mais pas notre conviction qu'elles constituent une structure intéressante. Regardons à présent les espaces de probabilité non-commutatifs.

2. Probabilités non-commutatives

On peut enfin aborder le sujet des probabilités non-commutatives. L'objectif de cette section est d'exposer brièvement les concepts fondamentaux de la théorie, pour ensuite aborder le sujet de l'indépendance libre, qui joue dans le cadre commutatif un rôle analogue à celui de l'indépendance classique, mais dont la combinatoire est beaucoup plus complexe. Enfin nous aborderons un résultat d'indépendance libre asymptotique entre des grandes matrices, qui justifie heuristiquement l'idée que les variables non-commutatives peuvent être approximées par de grandes matrices, idée qui est à la source de la définition de l'entropie libre, sujet de la prochaine section. L'exposé est nécessairement bref, tout le détail peut être trouvé dans le livre de Nica et Speicher [Spe06].

2.1. Espaces de probabilités non-commutatifs et distributions non-commutatives.

Définition 2.1. *Un **espace de probabilités non-commutatif** est un couple $(\mathcal{M}, \varphi)$, où $\mathcal{M}$ est une $\mathbb{C}$-algèbre unitaire et φ est une forme $\mathbb{C}$-linéaire vérifiant $\varphi(1_{\mathcal{M}}) = 1$. On dit que φ est un **état**.*

*Si $\varphi(ab) = \varphi(ba)\ \forall a, b \in \mathcal{M}$, on dit que l'état est **tracial**, et l'espace de probabilités non-commutatif $(\mathcal{M}, \varphi)$ est un **espace tracial**.*

*Si $\mathcal{M}$ est une $*$-algèbre et que φ est **positive**, au sens où elle prend des valeurs positives sur chaque élément positif :*

$$\varphi(a^*a) \geqslant 0\ \forall a$$

*alors $(\mathcal{M}, \varphi)$ est un **$*$-espace de probabilités**.*

*Si on a l'implication $\varphi(a^*a) = 0 \implies a = 0$, φ est dite **fidèle.***

Si de plus $\mathcal{M}$ est une C-algèbre et $(\mathcal{M}, \varphi)$ est un $*$-espace de probabilité, on parle alors de **C*-espace de probabilités**.*

*Si $\mathcal{M}$ est une algèbre de von Neumann, que $(\mathcal{M}, \varphi)$ est un $*$-espace de probabilités tracial, on parle de **W*-espace de probabilités**.*

Dans un $*$-espace de probabilités, on a toujours que $\varphi(a^*) = \overline{\varphi(a)}$. En effet, en faisant la décomposition $a = x + iy$ en parties réelle et imaginaire auto-adjointes, il suffit de se ramener à montrer que pour a auto-adjoint, $\varphi(a) \in \mathbb{R}$, mais ceci est impliqué par la positivité de φ, en décomposant a en différence de deux éléments positifs.

Il s'ensuit que $(a, b) \mapsto \varphi(b^*a)$ est une forme sesquilinéaire positive semi-définie, et vérifie par conséquent l'inégalité de Cauchy-Schwartz

$$|\varphi(b^*a)| \leqslant \sqrt{\varphi(a^*a)\varphi(b^*b)}$$

Il s'ensuit en particulier que si $(\mathcal{A}, \varphi)$ est un C*-espace de probabilités, alors φ est automatiquement continue. Regardons d'abord le cas positif, $p \in \mathcal{A}+$. Alors on a $\varphi(p) \geqslant 0$ par positivité, et comme $\|p\| = \rho(p)$, car p est normal, on peut définir $b = f(p) = (\|p\| - p)^{\frac{1}{2}}$, qui est auto-adjoint car $\overline{f} = f$ où $f(\lambda) = (\|p\| - \lambda)^{\frac{1}{2}}$. Donc on a $\varphi(b^*b) = \varphi(\|p\| - p) \geqslant 0$ par positivité de φ, ce qui implique $\varphi(p) \leqslant \|p\|$.

Si a est à présent un élément quelconque, en considérant l'élément positif a^*a, On a $|\varphi(a)| \leqslant \sqrt{\varphi(a^*a)}$ par Cauchy-Schwartz, puis

$$|\varphi(a)| \leqslant \sqrt{\|a^*a\|} = \|a\|$$

En utilisant le cas positif et l'identité (C*1).

La première chose à noter est que ce point de vue recouvre les espaces de probabilités usuels à travers leurs algèbres de variables aléatoires : $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ et $L^{-\infty}(\Omega, \mathcal{F}, \mathbb{P})$, munies l'intégration par rapport à $\mathbb{P}$, $\mathbb{E} : X \mapsto \int_\Omega X(\omega)\mathbb{P}(\mathrm{d}\omega)$, sont respectivement un W*-espace de probabilités et un $*$-espace de probabilité tracial, et $\mathbb{E}$ est fidèle puisque $\mathbb{E}[\overline{X}X] = 0 \implies |X|^2 = 0$ p.p..

Dans la mesure où toutes les indicatrices $\mathbb{1}_A, A \in \mathcal{F}$ sont des éléments de ces algèbres, aucune information n'est perdue en adoptant ce point de vue.

L'algèbre de von Neumann des matrices $M_n(\mathbb{C})$ munie de la forme $\mathrm{tr} : A \mapsto \frac{1}{n}\mathrm{Tr}(A)$, fait de $(M_n(\mathbb{C}), \tau)$ un W*-espace de probabilités avec état fidèle. La positivité et la fidélité de tr découlent du fait que $\mathrm{tr}(A^*A)$ est simplement le carré de sa norme de Frobenius, normalisé pour avoir $\mathrm{tr}(I_n) = 1$.

On peut également munir la $*$-algèbre $M_n(L^{\infty-}(\Omega, \mathcal{F}\mathbb{P}))$ des matrices carrées aléatoires de la forme

$$\tau : M = (M_{ij})_{1\leqslant i,j\leqslant n} \mapsto \int_\Omega \mathrm{tr}(M(\omega))\mathbb{P}(\mathrm{d}\omega)$$

Du point de vue algébrique, il s'agit d'un cas d'extension des scalaires :

$$M_n(L^{\infty-}(\Omega, \mathcal{F}\mathbb{P})) \cong L^{\infty-}(\Omega, \mathcal{F}\mathbb{P}) \otimes_{\mathbb{C}} M_n(\mathbb{C})$$

Où la matrice élémentaire $(\delta_{kl,ij}X)_{1\leqslant i,j\leqslant n}, X \in L^{\infty-}$ s'identifie au tenseur élémentaire $E_{kl} \otimes X$, et τ à $\mathrm{tr} \otimes \mathbb{P}$. C'est un $*$-espace de probabilités tracial.

Si $\mathcal{H}$ est un espace de Hilbert complexe, $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ une sous-$*$-algèbre, et $x_0 \in \mathcal{H}$ est tel que $\|x_0\| = 1$, alors $\varphi : \mathcal{A} \mapsto \mathbb{C}, A \mapsto \langle Ax_0, x_0\rangle$ fait de $(\mathcal{A}, \varphi)$ un $*$-espace de probabilités. Un des intérêts de cette description est que la continuité faible d'opérateur de $A \mapsto \langle Ax_0, x_0\rangle$ permet d'étendre φ à l'adhérence faible de $\mathcal{A}$, par prolongement. Si τ est tracial, on obtient en particulier un W*-espace de probabilités

$\mathbb{C}[G]$, muni de la forme $\tau : (\alpha_g)_{g\in G} \mapsto \alpha_e$ est un $*$-espace de probabilités tracial. Notons que dans $\mathcal{B}(L^2(G))$, on a toujours $\tau((\alpha_g)_{g\in G}) = \langle(\sum_{g\in G} \alpha_g\lambda_g)\delta_e, \delta_e\rangle$, et comme δ_e est de norme 1, l'inclusion de $\mathbb{C}[G]$ dans $\mathcal{B}(L^2(G))$ est de la forme donnée au paragraphe précédent. En surchargant la définition de τ, on obtient que $(L(G), \tau)$ est un W*-espace de probabilités. Il est souvent utile de voir les espaces de probabilités comme des sous-algèbres d'opérateurs, ce qui motive la définition suivante.

Définition 2.2. *Une **représentation** d'un $*$-espace de probabilités $(\mathcal{A}, \varphi)$ correspond à la donnée d'un espace de Hilbert complexe $\mathcal{H}$, d'un morphisme de $*$-algèbres unitaires $\Phi : \mathcal{A} \mapsto \mathcal{B}(\mathcal{H})$ et d'un vecteur $x_0 \in \mathcal{H}$ de norme 1 vérifiant $\varphi(a) = \langle\Phi(A)x_0, x_0\rangle$. Si ce morphisme est injectif, alors on parle de représentation fidèle.*

*Plus généralement, on parlera de **morphisme** $(\mathcal{A}, \varphi) \xrightarrow{f} (\mathcal{B}, \psi)$ d'espaces de probabilités pour désigner les morphismes $\mathcal{A} \mapsto \mathcal{B}$ vérifiant $\varphi = \psi \circ f$, et on fera varier la notion de morphisme en fonction de la structure algébrique de $\mathcal{A}$.*

L'injection $L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \mapsto \mathcal{B}(L^2(\Omega, \mathcal{F}\mathbb{P}))$ et le vecteur $\mathbb{1}_\Omega$ forment une représentation. Il est facile de vérifier que qu'il n'existe pas de vecteur x_0 dans tel que l'identité fournisse une représentation de $(M_n(\mathbb{C}), \mathrm{tr})$ dans $\mathbb{C}^n$ (autrement dit l'état tr n'est pas représenté par un vecteur), mais si $\Phi : M_n(\mathbb{C}) \mapsto \mathcal{B}(\mathbb{C}^{n^2}) \cong M_{n^2}(\mathbb{C})$ est l'injection diagonale selon une base $(e_{ij})_{1\leqslant i,j\leqslant n}$, on voit que le vecteur $x_0 = (\frac{\delta_{ij}}{\sqrt{n}})_{1\leqslant i,j\leqslant n})$ fournit une représentation fidèle. Par extension, on obtient une représentation de $M_n(L^\infty(\Omega, \mathcal{F}\mathbb{P}))$ sur $\mathcal{B}(L^2(\Omega, \mathcal{F}, \mathbb{P})^{n^2})$ grâce à $X_0 = (\frac{\mathbb{1}_\Omega\delta_{ij}}{\sqrt{n}})_{1\leqslant i,j\leqslant n})$.

Si $(\mathcal{A}, \tau)$ est un $*$-espace de probabilités tracial, on a immédiatement que

$$\forall a_1, a_2, \cdots, a_n \in \mathcal{A},\ \varphi(a_1 \cdots a_n) = \varphi(a_2 \cdots a_n a_1) = \varphi(a_{\sigma(1)} \cdots a_{\sigma(n)})$$

Pour n'importe quelle permutation cyclique σ de $\{1, \cdots, n\}$. En particulier, si $u \in \mathcal{A}$ est unitaire, alors on a $\varphi(u^*au) = \varphi(auu^*) = \varphi(a)$, c'est-à-dire que φ est invariant sous l'action par conjugaison du groupe multiplicatif des éléments unitaires, et même plus généralement sous l'action de son groupe d'éléments inversibles.

En particulier lorsque $\mathcal{A} = M_n(\mathbb{C})$ et $A \in \mathcal{A}$ est une matrice normale, le théorème spectral affirme que A est unitairement conjuguée à la matrice diagonale de ses valeurs propres $\Lambda = \text{diag}((\lambda_i)_{1 \leqslant i \leqslant n}), \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n : A = U^*\Lambda U$ pour une certaine matrice unitaire $U \in \mathcal{U}_n$. On obtient donc par exemple que

$$\forall k,\ \text{tr}(A^k) = \text{tr}(U^*\Lambda^k U) = \frac{1}{n}\sum_{i=1}^{n} \lambda_i^k$$

Les identités de Newton permettent alors de calculer $\chi_A(\lambda) = \det(A - \lambda I)$, le polynôme caractéristique de A, puis toutes ses valeurs propres.Autrement dit la collection $\{\text{tr}(M^k), k \geqslant 0\}$ caractérise complétement Λ, puis l'orbite de A sous l'action de $\mathcal{U}(n)$ par conjugaison.

La même observation s'étend aux conjugaisons simultanées par un même élément unitaire : si $(a_1, \cdots, a_n) \in \mathcal{A}^n$, $u \in \mathcal{A}$ unitaire, on a $\varphi(u^*a_1u \cdots u^*a_1u) = \varphi(u^*a_1(uu^*)\cdots(uu^*)a_nu) = \varphi(u^*a_1 \cdots a_nu) = \varphi(a_1 \cdots a_n)$.

Il existe aussi dans ce cas un résultat de théorie des invariants [Pro76]qui affirme que les *moments mixtes* de $(A_1, \ldots, A_n)$ où les $A_1, \cdots, A_n \in M_n(\mathbb{C})$ caractérisent complètement l'orbite de $(A_1, \ldots, A_n)$ sous l'action de conjugaison simultanée de $\mathcal{U}_n$. Plus explicitement,

$$\text{tr}(A_{i_1} \cdots A_{i_p}) = \text{tr}(B_{i_1} \cdots B_{i_p})\ \forall p \in \mathbb{N} \forall (i_1, \ldots, i_p) \in \{1, \ldots, n\}^p$$

$$\iff \exists U \in \mathcal{U}_n : U^*A_iU = B_i\ \forall i$$

Géométriquement, ceci signifie que les moments mixtes encodent toutes les valeurs propres des A_i, mais aussi les orientations relatives des sous-espaces propres associés dans $\mathbb{C}^n$.

Soit $\mathcal{A}$ un C*-espace de probabilités, et $a \in \mathcal{A}$ normal. On considère $\text{C}^*\langle 1, a\rangle \subset \mathcal{A}$, qui est une sous $*$-algèbre commutative, isomorphe à $\mathcal{C}(\sigma(a), \mathbb{C})$, via $f \mapsto f(a)$ par le théorème de Gelfand. Alors $f \mapsto \varphi(f(a))$ est une forme linéaire continue, et positive sur $\mathcal{C}(\sigma(a), \mathbb{C})$. Le théorème de Riesz affirme donc qu'il existe une unique mesure de Radon borélienne μ_a à support dans $\sigma(a)$ vérifiant

$$\varphi(f(a)) = \int_{\sigma(a)} f \text{d}\mu_a = \varphi(f(a))\ \forall f \in \mathcal{C}(\sigma(a), \mathbb{C})$$

Comme $\varphi(1) = 1$, on a aussi que $\mu_a(\sigma(a)) = 1$, si bien que μ_a est une mesure de probabilités. Pour $P \in \mathbb{C}\langle X, X^*\rangle$, notons $\tilde{\mu}_a : \mathbb{C}\langle X, X^*\rangle \mapsto \mathbb{C}, P \mapsto \varphi(P(a, a^*))$. Autrement dit, par linéarité, et puisque a et a^* commutent, l'information portée par $\tilde{\mu}_a$ est aussi celle portée par la collection des $\{\varphi((a^*)^m a^n), m \geqslant 0, n \geqslant 0\}$. Or, l'application $\mathbb{C}\langle X^*, X\rangle \xrightarrow{\iota} \mathcal{C}(\sigma(a, \mathbb{C}))$ qui à un polynôme associe sa fonction polynomiale est telle que, par calcul fonctionnel, $\tilde{\mu}_a = \mu_a \circ \iota$ (où on note $\mu_a : f \mapsto \int f \text{d}\mu_a = \varphi(f(a))$). Mais comme $\sigma(A)$ est compact, le théorème de Stone-Weierstrass implique que μ_a est déterminée par la valeur qu'elle prend sur les fonctions polynomiales, si bien que la connaissance de $\tilde{\mu}_a$ suffit à déterminer μ_a.
Dans les cas où a n'est pas normal, on perd la commutativité de $\text{C}^*\langle 1, a\rangle$, et l'interprétation analytique de $\tilde{\mu}_a$, mais celle-ci motive cependant la généralisation suivante.

Définition 2.3. *Soit $\mathcal{A}$ un $*$-espace de probabilités, et $a \in \mathcal{A}$ un élément quelconque. On appelle* ***distribution non-commutative*** *ou* ***$*$-distribution*** *la forme linéaire*

$$\mu_a : \mathbb{C}\langle X, X^*\rangle; P \mapsto P(a, a^*)$$

Si M est un monôme non-commutatif, c'est-à-dire de la forme

$$M = \prod_{i=1}^{k} X^{\epsilon_i},\ \epsilon_i \in \{1, *\}\ \forall i$$

*pour un certain $k \in \mathbb{N}$, le nombre $\mu_a(M)$ est l'$*$-**moment** de a associé à M, et la donnée de l'$*$-distribution de a correspond exactement à la donnée de ses $*$-moments.*
Si de plus, a est normal dans un C-espace de probabilités, la remarque ci-dessus montre qu'il existe une unique mesure de probabilité borélienne sur $\sigma(a)$, telle que $\mu_a(P)$ est donnée par l'intégrale de P sous cette mesure, que l'on désignera comme la **distribution analytique** de a. Elle est en particulier caractérisée par l'équation :*

$$\varphi(a^{*m}a^n) = \int_{\sigma(a)} \overline{z}^m z^n d\mu_a\ \forall m, n \in \mathbb{N}$$

Dans le cas où $\mathcal{A}$ est un C-espace de probabilités, les propriétés spectrales permettent de préciser la nature de μ_a en fonction de celle de a :*

Si a est auto-adjoint, $\mathrm{supp}(\mu_a) \subset \mathbb{R}$

Si a est positif, $\mathrm{supp}(\mu_a) \subset \mathbb{R}_+$

Si a est positif, $\mathrm{supp}(\mu_a) \subset \mathbb{S}^1$*, le cercle unité de $\mathbb{C}$*

Si a est un projecteur , $\mathrm{supp}(\mu_a) \subset \{0, 1\}$

Proposition 2.4. *Dans le cas où $(\mathcal{A}, \varphi)$ est un C*-espace de probabilités tel que φ soit fidèle, et $a \in \mathcal{A}$ est normal, on a exactement :*

$$\mathrm{supp}\,\mu_a = \sigma(a)$$

De plus, pour $f \in \mathcal{C}(\sigma(a))$, on a $\mu_{f(a)} = f_*\mu_a$, la mesure image de μ_a par f.

Démonstration. Pour $(\mathcal{A}, \varphi)$ un C*-espace de probabilités et a normal, on a toujours l'inclusion $\mathrm{supp}\,\mu_a \subset \sigma(a)$. Pour l'inclusion réciproque, supposons qu'il existe un $\lambda \in \sigma(a)$ tel que $\lambda \notin \mathrm{supp}\,\mu_a$. Comme $\mathbb{C} \setminus \mathrm{supp}\,\mu_a$ est un ouvert, prenons $r > 0$ tel que $B(\lambda, r) \subset \mathbb{C} \setminus \mathrm{supp}\,\mu_a$. Soit f la fonction en tipi valant 1 en λ et 0 en dehors de $B(, \lambda, r)$, $f(z) = (1 - \frac{|z-\lambda|}{r})^+ \leqslant \mathbb{1}_{B(\lambda,r)}$. Alors on a $\|f(a)\| = 1$ par calcul fonctionnel continu, et on a

$$\varphi(f(a)^* f(a)) = \varphi(f(a)^2) = \int_{\sigma(a)} f(a)^2 \mathrm{d}\mu_a \leqslant \mu_a(B(\lambda, r)) = 0$$

Donc φ n'est pas fidèle. On conclut à la contraposée.

Pour toute $g \in \mathcal{C}(\sigma(f(a)), \mathbb{C})$, et comme $f(a)$ est normal, on a

$$\int_{\sigma(a)} (g \circ f) \mathrm{d}\mu_a = \varphi(g(f(a)) = \int_{\sigma(f(a))} g \mathrm{d}\mu_{f(a)}$$

Or par changement de variable,

$$\int_{\sigma(a)} g \circ f \mathrm{d}\mu_a = \int_{f(\sigma(a))} g \mathrm{d} f_*\mu_a$$

Comme $\sigma(f(a)) = f(\sigma(a))$, on conclut. □

Exemple 9. Calculons quelques exemples de distributions analytiques dans le cas commutatif, c'est-à-dire normal, et voyons qu'il n'y a rien de très surprenant.

Vérifions que dans le cas d'une C*-algèbre classique de variables aléatoires, la distribution analytique d'une variable $Y \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ est sa distribution au sens classique, μ_Y. Il suffit de vérifier que pour un polynôme $P \in \mathbb{C}\langle X, X^*\rangle$, on a bien

$$\mathbb{E}[P(Y)] = \mathbb{E}[\tilde{P}(Y)] = \int_{\mathbb{C}} \tilde{P}(z, \overline{z}) \mathrm{d}\mu_Y(z) = \int_{\mathbb{C}} P(z, \overline{z}) \mathrm{d}\mu_Y(z)$$

, où on note $\tilde{P} = \mathbb{1}_{B(0,R)} P$, avec $\mathbb{P}(|Y| < R) = 1$, qui est mesurable et bornée.

Soit $\mathcal{A} = M_n(\mathbb{C})$ et $A \in \mathcal{A}$ une matrice normale. Alors le spectre de A est l'ensemble discret de ses valeurs propres, et la distribution analytique de A est l'unique mesure de probabilités sur $\sigma(A)$ vérifiant pour toute f continue :

$$\frac{1}{n} \operatorname{Tr}(f(A)) = \int_{\sigma(A)} f(A) \mathrm{d}\mu_A$$

En prenant $f = \mathbb{1}_{\{\lambda\}}$ où $\lambda \in \sigma(A)$, on obtient $\frac{m_\lambda}{n} = \mu_A(\lambda)$. On en déduit que

$$\mu_A = \sum_{\lambda \in \sigma(A)} \frac{m_\lambda}{n} \delta_\lambda$$

où m_λ est la multiplicité de la valeur propre λ pour A. C'est donc la mesure de comptage normalisée sur $\sigma(A)$, où les valeurs multiples sont comptées avec leur multiplicité.

Soit $\mathcal{A} = \mathbb{C}G$, G discret, muni de la trace τ_G, dont on peut montrer la fidélité. On considère l'élément $g \in \mathbb{C}G$, qui est normal, et unitaire. Sa distribution analytique est donc portée par $\mathbb{S}^1$, et la donnée de ses $*$-moments correspond donc à la donnée des $\varphi(g^k),\ k \in \mathbb{Z}$. De surcroît on a :

$$\tau_G(g^k) = \begin{cases} 1 & g^k = e \\ 0 \text{ sinon} & \end{cases}$$

En particulier on a $\tau_G(g^k) = \mathbb{1}_{\mathrm{ord}(g)|k}$. La distribution analytique de g est donc l'unique mesure de probabilité sur le cercle unité vérifiant :

$$\int_0^{2\pi} e^{i\theta k} \mu_g(\mathrm{d}\theta) = \mathbb{1}_{\mathrm{ord}(g)|k}$$

Si $\mathrm{ord}(g) = \infty$, ceci équivaut à $k = 0$, et on voit que la mesure $\mathrm{d}\mu_g = \frac{\mathrm{d}\theta}{2\pi}$ est celle qui convient, et si $\mathrm{ord}(g) = p \in \mathbb{N}$, alors il s'agit de la mesure $\frac{1}{p} \sum_{k=0}^{p-1} \delta_{e^{2ik\pi/p}}$. Ce sont dans les deux cas des mesures de Haar, respectivement sur le cercle unité et sur le groupe cyclique des n-ièmes racines de l'unité.

Exemple 10. Examinons maintenant un premier exemple simple utilisant un élément non-normal.

Soit $\mathcal{H} = \ell^2(\mathbb{N})$, et $S \in \mathcal{B}(\mathcal{H})$. On définit l'espace de probabilités non-commutatif $(\mathcal{H}, \varphi)$ où $\varphi : A \mapsto \langle A\delta_0, \delta_0\rangle$, où on note $(\delta_n)_{n\in\mathbb{N}}$ la base canonique de $\mathcal{H}$. Considérons l'opérateur de décalage T : défini par $T\delta_k = \delta_{k+1}$. Son adjoint est alors défini par $T^*\delta_k \mapsto \delta_{k-1}$ pour $k \geqslant 1$, et $T^*\delta_0 = 0$. On a clairement $T * T = 1$, mais T^* n'est pas inversible, donc T est une isométrie non-unitaire, en particulier non-normal.

En revanche, l'élément $T + T^*$ est normal, et même auto-adjoint, ce qui implique que son $*$-distribution est déterminée par la valeur des

$$\langle (T + T^*)^n \delta_0, \delta_0\rangle,\ n \in \mathbb{N}$$

Or, on a

$$(T + T^*)^n = \sum_{\epsilon \in \{1,*\}^n} T^{\epsilon_n} \cdots T^{\epsilon_1}$$

Et donc on en déduit que :

$$\varphi((T+T^*)^n) = |\{\epsilon \in \{1,*\}^n : \ T^{\epsilon_n} \cdots T^{\epsilon_1}\delta_0 = \delta_0\}|$$

Observons à présent que si il existe un $1 \leqslant k \leqslant n$ tel que $N_k(*) := |\{1 \leqslant j \leqslant k : \epsilon_j = *\}| > |\{1 \leqslant j \leqslant k : \epsilon_j = 1\}| := N_k(1)$, alors on a

$$T^{\epsilon_k} \cdots T^{\epsilon_1}\delta_0 = 0$$

et si $N_k(1) > N_k(*) \forall 1 \leqslant k \leqslant n$, alors

$$T^{\epsilon_k} \cdots T^{\epsilon_1}\delta_0 = \delta_{N_n(1)-N_n(*)} \neq \delta_0$$

Il s'ensuit que

$$\varphi(T+T^*)^n = D_n$$

Où D_n est le nombre de suites $\{\epsilon_1, \ldots, \epsilon_n\}$ telles que $N_n(1) = N_n(\epsilon), N_k(1) \geqslant N_k(\epsilon)\ \forall 1 \leqslant k \leqslant n$. Il s'agit en fait d'un problème de combinatoire classique, l'énumération des chemins de Dyck, dont la réponse est donnée par :

$$D_{2n+1} = 0, D_{2n} = C_n$$

où C_n est le n-ième nombre de Catalan, $C_n = \frac{1}{n+1}\binom{2n}{n} = \binom{2n}{n} - \binom{2n}{n+1}$. Or on peut vérifier que les moments de la densité $f(x) = \mathbb{1}_{[-2,2]}(x)\sqrt{4-x^2}$, en les exprimant en fonction d'intégrales de Wallis, sont donné par $C_{n/2}$ lorsque n est pair, les moments impairs étant manifestement nuls. En d'autres termes,

$$\int_{\mathbb{R}} x^n f(x)\mathrm{d}x = D_n$$

ce qui montre que $T+T^*$ a comme distribution analytique la mesure sur $[-2,2]$ de densité f. Un tel élément est appelé *semi-circulaire* standard.

Dans le cas non-normal, on a vu que la non-commutativité a priori de la sous-algèbre engendrée nous pousse à considérer sa distribution comme la donnée de ses $*$-moments. On procède de même pour définir les distributions jointes. Formellement :

Définition 2.5. *Soient $(a_1, \ldots, a_n) \in \mathcal{A}^n$ où $\mathcal{A}$ est un $*$-espace de probabilités. On définit alors encore l'$*$-**distribution jointe**, comme la donnée des moments mixtes des $(a_1, \ldots, a_n)$, c'est-à-dire la forme linéaire*

$$\mu_{a_1,\ldots,a_n} : \mathbb{C}\langle X_1, \ldots, X_n, X_1^*, \ldots, X_n^*\rangle \mapsto \mathbb{C}; P \mapsto \varphi(P(a_1, \ldots, a_n))$$

Remarquons que si φ est fidèle, la distribution d'un n-uplet $(a_1, \ldots, a_n)$ ne dépend pas de l'espace sur lequel on le représente, au sens où on a un isomorphisme de $*$-espaces de probabilités

$$\mathcal{A} \ni *\langle 1, a_1, \ldots, a_n\rangle \xrightarrow{\Phi} *\langle 1, b_1, \ldots, b_n\rangle \subset \mathcal{B}$$

où $(\mathcal{B}, \psi)$ est un $*$-espace de probabilité muni d'un état fidèle, dès qu'on a $\mu_{b_1,\ldots,b_n} = \mu_{a_1,\ldots,a_n}$. Dénotons cette forme μ pour alléger les notations. Pour $P, Q \in \mathbb{C}\langle a_1, \ldots, a_n, a_1^*, \ldots, a_n^*\rangle$, on a $P(a_1, \ldots, a_n) = Q(a_1, \ldots, a_n) \iff \varphi((P-Q)^*(a_1, \ldots, a_n)(P-Q)(a_1, \ldots, a_n)) = 0 \iff \mu((P-Q)^*(P-Q) \iff P(b_1, \ldots, b_n)$, par fidélité de ψ.

L'application $\Phi : *\langle 1, a_1, \ldots, a_n\rangle \mapsto *\langle 1, b_1, \ldots, b_n\rangle$ est donc bien définie par $P(a_1, \ldots, a_n) \mapsto P(b_1, \ldots, b_n)$, et bijective, c'est clairement un $*$-morphisme, et on a

$$\varphi(P(a_1, \ldots, a_n)) = \mu(P) = \psi(P(b_1), \ldots, b_n) = \psi \circ \Phi(P(a_1, \ldots, a_n))$$

Φ est donc un isomorphisme d'$*$-espaces de probabilités.

On peut également montrer[Spe06, p. 66], grâce à un analogue de la formule de Gelfand pour φ, que si $\mathcal{A}$ et $\mathcal{B}$ sont des C*-espaces de probabilité, que ce $*$-isomorphisme se prolonge à un

C*-isomorphisme (c'est-à-dire un $*$-isomorphisme isométrique) entre les C*-algèbres respectives engendrées par ces n-uplets.

Définissons enfin la notion de convergence en distribution non-commutative, de la seule manière qu'on puisse espérer, c'est-à-dire par convergence des moments

Définition 2.6. *Soit $(\mathcal{A}_k, \varphi_k)_{k\in\mathbb{N}}$ une suite d'espace de probabilités non-commutatifs, et pour chaque k, on fixe un n-uplet $(a_1^{(k)}, \dots, a_n^{(k)})$. Soit également $(\mathcal{A}, \varphi)$ un espace de probabilités non-commutatif, $(a_1, \dots, a_n) \in \mathcal{A}^n$. On dit alors que la suite $(a_1^{(k)}, \dots, a_n^{(k)})_k$* ***converge en distribution non-commutative*** *si pour tout $P \in \mathbb{C}\langle X_1, \dots, X_k, X_1^*, \dots, X_k^*\rangle$, on a*

$$\varphi_n(P(a_1^{(k)}, \dots, a_n^{(k)})) \xrightarrow{n\to\infty} \varphi(P(a_1, \dots, a_n))$$
$$\iff \mu_{(a_1^{(k)},\dots,a_n^{(k)})} \xrightarrow{n\to\infty} \mu_{(a_1,\dots,a_n)}$$

En tout point. On peut naturellement étendre cette définition à des familles $(a_\alpha^{(k)})_{k\in\mathbb{N},\alpha\in I}$ arbitraires en disant qu'une telle famille converge en distribution non-commutative vers $(a_\alpha)_{\alpha\in I}$ si pour tout $J \subset I$ fini, on a la convergence en distribution non-commutative $(a_\alpha^{(k)})_{\alpha\in J} \to (a_\alpha)_{\alpha\in J}$

2.2. **Indépendance libre.** Dans un espace de probabilités classique $(\Omega, \mathcal{F}, \mathbb{P})$, une famille $(\mathcal{F}_\alpha)_{\alpha\in I}$ de sous-σ-algèbres est dite indépendante si et seulement si :

$$\forall J \subset I \text{ fini},\ \forall (A_\alpha)_{\alpha\in J}, A_\alpha \in \mathcal{F}_\alpha\ \forall \alpha$$
$$\mathbb{P}(\bigcap_{\alpha\in J} A_\alpha) = \prod_{\alpha\in J} \mathbb{P}(A_\alpha)$$

On peut reformuler cette définition de façon équivalente : la famille $(\mathcal{F}_\alpha)_{\alpha\in I}$ est indépendante si et seulement si

$$\forall \alpha_1 \neq \alpha_2 \neq \dots \neq \alpha_k \in I,\ \forall (X_{\alpha_j})_{j=1\dots,k}, X_\alpha \in L^\infty(\Omega, \mathcal{F}_\alpha, \mathbb{P})\ \forall \alpha$$
$$\mathbb{E}[X_{\alpha_j}] = 0\ \forall 1 \leqslant j \leqslant k \implies \mathbb{E}[\prod_{1\leqslant j\leqslant k} X_{\alpha_j}] = 0$$

Le sens $\implies$ est évident. Pour l'autre sens, on considère, pour $(A_\alpha)_{\alpha\in J}$ donné, les variables centrées $(\mathbb{1}_\alpha - \mathbb{P}(A_\alpha))_{\alpha\in J}$. On a alors, en montrant par récurrence sur $|J|$, le cas $|J| = 1$ étant tautologique, que $\mathbb{P}(\bigcap_{\alpha\in J} A_\alpha) = \prod_{\alpha\in J} \mathbb{P}(A_\alpha)\ \forall A_\alpha \in \mathcal{F}_\alpha$.

$$\mathbb{E}[\prod_{\alpha\in J}(\mathbb{1}_{A_\alpha} - \mathbb{P}(A_\alpha))] = 0 = \mathbb{E}[\prod_{\alpha\in J} \mathbb{1}_{A_\alpha}] + \sum_{\varnothing\neq K\subset J} (-1)^{|K|}\mathbb{E}[\prod_{\gamma\in J\setminus K} \mathbb{1}_{A_\gamma}] \prod_{\beta\in K} \mathbb{P}(A_\beta)$$

Tous les espérances dans la somme sont connues par un cas de récurrence de rang inférieur à $|J|$, puisque K est non-vide, donc on obtient :

$$\mathbb{P}(\bigcap_{\alpha\in J} A_\alpha) = -\prod_{\alpha\in J} \mathbb{P}(A_\alpha) \sum_{\varnothing\neq K\subset J} (-1)^{|K|}$$

Et on se ramène à évaluer la somme

$$\sum_{\varnothing\neq K\subset J} (-1)^{|K|} = \sum_{k=1}^{|J|} \binom{|J|}{k} (-1)^k = \sum_{k=0}^{|J|} \binom{|J|}{k} (-1)^k - 1 = (1-1)^{|J|} - 1 = -1$$

Ce qui conclut. La notion d'indépendance libre est celle qu'on obtient quand on remplace la condition

$$\alpha_1 \neq \alpha_2 \neq \dots \neq \alpha_k$$

par la condition :

$$\alpha_1 \neq \alpha_2, \alpha_2 \neq \alpha_3, \ldots, \alpha_{k-1} \neq \alpha_k$$

Définition 2.7. *On se donne un espace de probabilités non-commutatif $(\mathcal{A}, \varphi)$, et une famille $(\mathcal{A}_\alpha)_{\alpha \in I}$ de sous-algèbres unitaires. On dit qu'elles forment une famille* ***librement indépendante****, si :*

$$\forall k \in \mathbb{N},\ \forall \alpha_1 \neq \alpha_2, \alpha_2 \neq \alpha_3, \ldots, \alpha_{k-1} \neq \alpha_k,\ (\alpha_1, \ldots, \alpha_k) \in I^k$$
$$\forall\ (a_j)_{j=1\ldots k},\ a_j \in \mathcal{A}_{\alpha_j}\ \forall j$$

On a :

$$\varphi(a_j) = 0 \forall j \implies \varphi(\prod_{j=1\ldots k} a_j) = 0$$

De la même façon qu'en probabilités classiques, on définit l'indépendance libre de variables aléatoires non-commutatives comme celle des sous-algèbres unitaires engendrées, et, si on est dans un $*$-espace de probabilités, on définit l'$*$-indépendance libre comme l'indépendance libre des $*$-algèbres engendrées.

Il est clair que deux variables aléatoires indépendantes ne sont en général pas librement indépendantes : par exemple, si $X_1 = (-1)^{B_1}, Y_1 = (-1)^{B_2}, X_2 = -X_1, Y_2 = -Y_1$ où $X \sim Y$ sont des variables de Bernoulli indépendantes de paramètre $\frac{1}{2}$, on a $\mathbb{E}[X_1] = \mathbb{E}[Y_1] = \mathbb{E}[X_2] = \mathbb{E}[Y_2] = 0$, mais $\mathbb{E}[X_1 Y_1 X_2 Y_2] = 1$. Pour voir ce que donnerait ce calcul dans le cas libre, posons a_1,b_1,a_2 et b_2 des éléments d'un espace de probabilités non-commutatif $(\mathcal{A}, \varphi)$, en provenance de deux sous-algèbres librement indépendantes. Considérons les variables $a_i - \varphi(a_i)$, $b_i - \varphi(b_i)$ pour $i = 1, 2$, telles que $\varphi(a_i - \varphi(a_i)) = \varphi(b_j - \varphi(b_j)) = 0$, et qui sont librement indépendantes, puisque contenues dans les sous-algèbres unitaires engendrées respectivement par a_i et b_j (on vérifie facilement que deux sous-algèbres respectives de deux algèbres librement indépendantes sont elles-même librement indépendantes). Calculons alors :

$$\begin{aligned}
&\varphi((a_1 - \varphi(a_1))(b_1 - \varphi(b_1))(a_2 - \varphi(a_2))(b_2 - \varphi(b_2))) \\
&= \varphi(a_1 b_1 a_2 b_2) - \varphi(b_2)\varphi(a_1 b_1 a_2) - \varphi(a_2)\varphi(a_1 b_1 b_2) + \varphi(a_2)\varphi(b_2)\varphi(a_1 b_1) \\
&-\varphi(b_1)\varphi(a_1 a_2 b_2) + \varphi(b_1)\varphi(b_2)\varphi(a_1 a_2) + \varphi(b_1)\varphi(a_2)\varphi(a_1 b_2) - \varphi(a_1)\varphi(b_1)\varphi(a_2)\varphi(b_2) \\
&-\varphi(a_1)\varphi(b_1 a_2 b_2) + \varphi(a_1)\varphi(b_2)\varphi(b_1 a_2) + \varphi(a_1)\varphi(a_2)\varphi(b_1 b_2) - \varphi(a_1)\varphi(a_2)\varphi(b_1)\varphi(b_2) \\
&+\varphi(a_1)\varphi(b_1)\varphi(a_2 b_2) - \varphi(a_1)\varphi(a_2)\varphi(b_1)\varphi(b_2) - \varphi(a_1)\varphi(a_2)\varphi(b_1)\varphi(b_2) + \varphi(a_1)\varphi(a_2)\varphi(b_1)\varphi(b_2) \\
&= 0
\end{aligned}$$

On peut continuer à simplifier. On peut vérifier, en utilisant la même technique de substitution par des variables centrées, que, $\varphi(cd) = \varphi(c)\varphi(d)$, pour c et d librement indépendantes, puis également que $\varphi(cdc') = \varphi(cc')\varphi(d)$, si c, c' sont dans la même sous-algèbre librement indépendante de d. En appliquant ceci au calcul ci-dessus, on obtient, après moult simplifications :

$$\phi(a_1 b_1 a_2 b_2) = \varphi(a_1 a_2)\varphi(b_1)\varphi(b_2) + \varphi(a_1)\varphi(a_2)\varphi(b_1 b_2) - \varphi(a)\varphi(a_2)\varphi(b_1)\varphi(b_2)$$

Le résultat est différent du $\varphi(a_1 a_2)\varphi(b_1 b_2)$ qu'on aurait obtenu dans le cas classique, mais similaire dans le sens où pour le calculer, on a *in fine* seulement besoin de connaître

$$\varphi(a_1), \varphi(a_2), \varphi(a_1 a_2), \varphi(b_1), \varphi(b_2), \varphi(b_1 b_2)$$

Autrement dit, comme dans le cas classique, ce moment mixte de (a, b) se calcule en fonction des moments (non-mixtes) de a, b. Ce fait se généralise dans la proposition suivante, dont la démonstration est assez analogue du cas classique, mais qui montre bien à quel point la combinatoire se complique dans le cadre non-commutatif.

Proposition 2.8. [Bia98] *Soit $r \in \mathbb{N}$. Pour une partition π de $\{1, \dots, n\}$, écrivons*

$$\varphi_\pi = \prod_{\substack{\{i_1,\dots,i_r\}\in\pi \\ i_1<\dots<i_r}} \varphi(a_{i_1} \dots a_{i_r})$$

Et notons $\pi_1 \leqslant \pi_2$ l'ordre sur les partitions défini par $\pi_1 \leqslant \pi_2$ si et seulement si toute classe de π_1 s'écrit comme une union (forcément disjointe) de classes de π_2.

Soient $(\mathcal{A}_\alpha)_{\alpha\in I}$ une famille de sous-algèbres librement indépendantes, donnons nous une famille finie $(a_1, \dots, a_n)$ telle que $a_k \in \mathcal{A}_{\alpha_k}$ $\forall k$. Partitionnons $\{1, \dots, n\} := \mathbf{n}$ via la relation d'équivalence $i \sim j \iff \alpha_i = \alpha_j$, et notons Π cette partition. Alors on a le résultat : pour chaque $\pi \leqslant \Pi$, il existe des coefficients $c(\pi, \Pi)$ ne dépendant que de (π, Π) tels que

$$\varphi(a_1 \dots a_n) = \sum_{\pi\leqslant\Pi} c(\pi, \Pi)\varphi_\pi$$

Démonstration. On raisonne par récurrence sur k. Si $k = 1$, la seule partition est la partition triviale 1, et on pose $c(1, 1) = 1$. Sinon, on considère les variables recentrées $a'_i = a_i - \varphi(a_i)$, et on écrit :

$$0 = \varphi(a'_1 \cdots a'_n) = \varphi(a_1 \cdots a_n) + \sum_{\varnothing\neq K\subset\mathbf{n}} (-1)^{|K|} \prod_{k\in K} \varphi(a_k)\varphi(\prod_{j\in\mathbf{n}\setminus K} a_j)$$

Où le produit à l'intérieur de φ est ordonné. Considérons, pour chaque K, la partition Π_K de $\mathbf{n} \setminus K$ par la relation d'équivalence $\sim$. Comme K est non-vide, l'hypothèse de récurrence nous dit qu'on peut écrire

$$\varphi(\prod_{j\in\mathbf{n}\setminus K} a_j) = \sum_{\pi\leqslant\Pi_K} c(\pi, \Pi_K)\varphi_\pi$$

En substituant, on obtient

$$\begin{aligned}\varphi(a_1 \cdots a_n) &= -\sum_{\varnothing\neq K\subset\mathbf{n}} (-1)^{|K|} \prod_{k\in K} \varphi(a_k) \sum_{\pi\leqslant\Pi_K} c(\pi, \Pi_K)\varphi_\pi \\ &= -\sum_{\varnothing\neq K\subset\mathbf{n}} (-1)^{|K|} \sum_{\pi\leqslant\Pi_K} c(\pi, \Pi)\varphi_\pi \prod_{k\in K} \varphi(a_k)\end{aligned}$$

Or, pour $\pi \leqslant \Pi_K$ une partition donnée, $\varphi_\pi \prod_{k\in K} \varphi(a_k) = \varphi_{\pi'}$, où $\pi' = \pi'(\pi, K)$ est la partition de $\mathbf{n}$ telle que la classe de tout élément de K soit un singleton, et telle que la classe de tout élément $x \in \mathbf{n} \setminus K$ soit $\mathrm{cl}_\pi(x) \setminus K$. Il est clair que $\pi' \leqslant \Pi$: si $x \sim_{\pi'} x'$, alors soit $x' \in K \vee x \in K$ ce qui implique $x = x'$, soit $x \sim_\pi x'$ et $x \sim x'$ puisque $\pi \leqslant \Pi$. On a donc finalement :

$$\varphi(a_1 \cdots a_n) = \sum_{\varnothing\neq K\subset\mathbf{n}} \sum_{\pi\leqslant\Pi_K} (-1)^{|K|+1} c(\pi, \Pi_K)\varphi_{\pi'}$$

qui est bien une somme sur les sous-partitions de Π.

Comme chaque φ_π avec $\pi \leqslant \Pi$ est l'espérance d'un produit d'éléments appartenant à la même classe d'équivalence modulo $\sim$, ce qui signifie que tous ces éléments appartiennent à une même sous-algèbre, il s'ensuit qu'étant donné des sous-algèbres librement indépendantes $(\mathcal{A}_\alpha)_{\alpha\in I}$ et une famille d'éléments $(a_1, \dots, a_n)$ appartenant à la réunion de ces sous-algèbres, il suffit en principe de connaître les restrictions $(\varphi_{|\mathcal{A}_\alpha})_{\alpha\in I}$ pour pouvoir calculer $\varphi(a_1 \cdots a_n)$. C'est un peu trompeur, puisqu'il faut aussi connaître les $c(\pi, \Pi)$, ce qui constitue le point de départ de la riche théorie combinatoire développée par Roland Speicher. □

On peut abstraire le cas de l'indépendance classique au cadre des espaces non-commutatifs, en disant que des sous-algèbres sont classiquements indépendantes si et seulement si tout moment $\varphi(a_1 \dots a_n)$ se factorise en φ_Π, ce qui revient à dire que, dans cette théorie de l'indépendance classique, on a $c(\Pi, \Pi) = 1,\ c(\pi, \Pi) = 0\ \forall \pi < \Pi$. On aurait pû également adopter la définition *centrée* qu'on a vu plus haut.

Soient $(\mathcal{A}, \varphi), (\mathcal{B}, \psi)$ deux espaces deux probabilités non-commutatifs. On peut munir la $\mathbb{C}$-algèbre $\mathcal{A} \otimes \mathcal{B}$ d'une structure d'espace de probabilités non-commutatif, grâce à l'état $\varphi \otimes \psi$ déterminé par $\varphi \otimes \psi(a \otimes b) = \varphi(a)\varphi(b)\ \forall a \in \mathcal{A},\ b \in \mathcal{B}$. Il est très facile de voir que les sous-algèbres unitaires $\mathcal{A} \otimes 1_{\mathcal{B}}\mathbb{C}$ et $1_{\mathcal{A}}\mathbb{C} \otimes \mathcal{B}$ sont classiquement indépendantes. Cela justifie la terminologie usuelle, qui qualifie l'indépendance classique d'**indépendance tensorielle**.

Il se trouve que la notion d'indépendance libre correspond également à une une construction algébrique, celle de produit libre, qui est plus simple à comprendre dans le cadre des groupes. Une *présentation* pour un groupe G est la donnée d'un ensemble de *générateurs* $\mathcal{G} = \{g_\alpha\}_{\alpha \in I}$, d'un ensemble de *relations* $\mathcal{R} \subset \mathbb{F}_\mathcal{G}$, qui sont des mots sur le groupe libre engendré par $\mathcal{G}$, et d'un isomorphisme de groupes $G \cong \mathbb{F}_\mathcal{G}/\langle\mathcal{R}\rangle^{\mathbb{F}_\mathcal{G}}$ où on quotiente par la clôture conjuguée de $\mathcal{R}$, le plus petit sous-groupe distingué contenant $\mathcal{R}$. On écrit alors $G = \langle\mathcal{G}|\mathcal{R}\rangle$. Concrètement, les relations donnent les équations $f(g_1, \dots, g_k) = e$ non-triviales dans G, c'est-à-dire non imposées par la seule structure de groupe. Ainsi, le groupe cyclique admet la présentation $C_n = \langle g|g^n\rangle$, le groupe libre $\mathbb{F}_n = \langle g_1, \dots g_n|\varnothing\rangle$ le groupe *abélien* libre $\mathbb{Z}^n = \langle g_1, \dots, g_n|g_i g_j g_i^{-1} g_j^{-1},\ i \neq j\rangle$, et ainsi de suite. Soit maintenant une famille $(G_\alpha)_{\alpha \in I}$ de groupes, avec $G_\alpha = \langle\mathcal{G}_\alpha|\mathcal{R}_\alpha\rangle$ d'élément neutre e_α. Alors on définit le produit libre :

$$P = *_{\alpha \in I} G_\alpha = \langle \bigcup_{\alpha \in I} \mathcal{G}_\alpha | \bigcup_{\alpha \in I} \mathcal{R}_\alpha \cup \{e_\alpha\}\rangle$$

En particulier, on voit que si on se donne une famille $(g_1, \dots, g_n)$ avec $g_i \neq e_{\alpha_i} \in G_{\alpha_i},\ 1 \leqslant i \leqslant n$ et $\alpha_1 \neq \alpha_2, \dots, \alpha_{n-1} \neq \alpha_n$, on a $g_1 \cdots g_n \neq e$ dans P, puisqu'il n'y a aucune relation permettant de simplifier deux éléments en provenance de groupes différents. Par analogie, on dit d'une famille de sous-groupes $(G_\alpha)_\alpha$ de G est libre si il n'y a aucune relation non-triviale entre eux, au sens ou tout produit d'éléments non-triviaux dont les termes adjacents sont en provenance de sous-groupes différents est lui-même non-trivial. Un premier lien entre la notion de liberté algébrique et celle d'indépendance lire est donnée par la proposition suivante :

Proposition 2.9. [Spe06, p. 78] *Soit G un groupe, $(G_\alpha)_{\alpha \in I}$ une famille de sous-groupes. Alors on a l'équivalence :*

$$(G_\alpha)_{\alpha \in I} \text{ est libre dans } G \iff (\mathbb{C}G_\alpha)_{\alpha \in I} \text{ est librement indépendante dans } \mathbb{C}G$$

Démonstration. On rappelle que pour $g \in \mathbb{C}G$ un vecteur de la base canonique, on a $\tau(g) = 0 \iff g \neq e$. Soient $(g_1, \dots, g_n)$une famille telle que $g_i \neq e\ \forall i,\ g_i \in G_{\alpha_i}$, et vérifiant la condition de non-adjacence habituelle $\alpha_1 \neq \alpha_2, \dots, \alpha_{n-1} \neq \alpha_n$. Alors par indépendance libre des $\mathbb{C}G_\alpha$, $\tau(g_1 \cdots g_n) = 0 \implies g_1 \cdots g_n \neq e$ dans G, ce qui prouve le sens $\Longleftarrow$.

Pour l'autre sens, on fixe $a_j = \sum_{g \in G_{\alpha_j}} c_{g,j} g \in \mathbb{C}G_{\alpha_j}$, $j = 1 \cdots n$ tels que $\alpha_1 \neq \alpha_2, \dots, \alpha_{n-1} \neq \alpha_n$, et $\tau(a_j) = 0$, ce qui signifie que $c_{e,j} = 0\ \forall j$. Alors on a

$$\tau(a_1 \cdots a_n) = \sum_{\substack{g_1 \cdots g_n = e \\ (g_1, \dots, g_n) \in G_{\alpha_1} \times \cdots \times G_{\alpha_n}}} c_{g_1,1} \cdots c_{g_n,n}$$

Or la liberté des $(G_\alpha)_{\alpha \in I}$ dans G nous dit exactement que toute famille $(g_1, \dots, g_n)$ de cette forme doit contenir au moins un élément neutre $g_j = e$, et donc le produit dans chaque terme de la somme

contient au moins un coefficient $c_{e,j} = 0$, ce qui montre que $\varphi(a_1 \cdots a_n) = 0$, et l'indépendance libre des $(\mathbb{C}G_\alpha)_{\alpha \in I}$. □

On peut également définir une notion de produit libre d'espaces de probabilités non-commutatifs. La construction est la suivante. Soit $(\mathcal{A}_\alpha, \varphi_\alpha)_{\alpha \in I}$ une famille d'espaces de probabilités non-commutatifs. On note $\mathcal{A}_\alpha^0 = \ker(\varphi_\alpha)$ le sous-espace des variables centrées. Pour tout n-uplets d'indices $t = (\alpha_1, \ldots, \alpha_n)$ vérifiant la condition $\alpha_1 \neq \alpha_2, \ldots, \alpha_{n-1} \neq \alpha_n$, on considère le $\mathbb{C}$-espace vectoriel

$$\mathcal{W}_t = \bigotimes_{\alpha \in t} \mathcal{A}_\alpha^0$$

qu'on comprend intuitivement comme l'espace engendré par les mots de type t, c'est à-dire des concaténations formelles d'éléments $a_1 \ldots a_n$ où chaque $a_j \in \mathcal{A}_{\alpha_i}^0$. Notons également T l'ensemble de tous les types. Alors on pose

$$\mathcal{A} = \mathbb{C} \oplus \left(\bigoplus_{t \in T} \mathcal{W}_t \right)$$

C'est clairement un $\mathbb{C}$-espace vectoriel, la seule difficulté est de définir la multiplication sur le terme de droite pour en faire une algèbre. En effet, pour $z, z' \in \mathbb{C}$ et l, l' des combinaisons linéaires de mots, on pose $(z + l) \cdot (z' + l') = zz' + zl' + lz' + ll'$, et il suffit de spécifier la valeur de ll'. En fait, il suffit même de spécifier la règle de multiplication sur les mots, puisque ceux-ci engendrent $\bigoplus_t \mathcal{W}_t$, c'est-à-dire les produits de la forme :

$$(a_1 \otimes \cdots \otimes a_n) \cdot (b_1 \otimes \cdots \otimes b_m)$$

où $(a_1, \ldots, a_n)$ est de type $(\alpha_1, \ldots, \alpha_n)$ et $(b_1, \ldots, b_m)$ est de type $(\beta_1, \ldots, \beta_m)$. Si $\alpha_n \neq \beta_1$, il n'y a pas de difficulté, on pose

$$(a_1 \otimes \cdots \otimes a_n) \cdot (b_1 \otimes \cdots \otimes b_m) = a_1 \otimes \cdots \otimes a_n \otimes b_1 \otimes \cdots \otimes b_m$$

Si $\alpha_n = \beta_1$, la simple concaténation ne fonctionne plus, puisqu'on n'a alors pas nécessairement $\varphi_{\alpha_n}(a_n b_1) = 0$, $a_n b_1 \in \mathcal{A}_{\alpha_n}^0$, la condition qu'on aurait pu espérer pour obtenir un mot valide $a_1 \otimes \cdots \otimes a_n b_1 \otimes \cdots b_m$. L'astuce est de définir la multiplication récursivement, en utilisant la variable centrée $a_n b_1 - \varphi_{\alpha_n}(a_n b_1)$.

On prescrit la règle pour la multiplication de deux mots a, a' de même type (α) : $aa' = \varphi_\alpha(aa') + (aa' - \varphi_\alpha(aa'))$. Sinon, comme le produit $(a_1 \otimes \cdots \otimes a_{n-1}) \cdot (b_2 \otimes \cdots \otimes b_m)$ est bien défini par récurrence, on peut définir

$$(a_1 \otimes \cdots \otimes a_n) \cdot (b_1 \otimes \cdots \otimes b_m)$$

comme l'expression

$$a_1 \otimes \cdots \otimes (a_n b_1 - \varphi_{\alpha_n}(a_n b_1)) \otimes \cdots \otimes b_n + \varphi_{\alpha_n}(a_n b_1)(a_1 \otimes \cdots \otimes a_{n-1}) \cdot (b_2 \otimes \cdots \otimes b_m)$$

À titre d'exemple, on peut essayer de calculer $c \otimes b \cdot b' \otimes c'$ avec $b, b' \in \mathcal{B} \neq \mathcal{C} \ni c, c'$, pour voir comment ce produit se calcule récursivement :

$$c \otimes b \cdot b' \otimes c' = c \otimes (bb' - \varphi_{\mathcal{B}}(bb')) \otimes c' + \varphi_{\mathcal{B}}(bb')c \otimes c$$

$$b \otimes c' = cc' - \varphi_{\mathcal{C}}(cc') + \varphi_{\mathcal{C}}(cc')$$

$$c \otimes b \cdot b' \otimes c' = \varphi_{\mathcal{B}}(bb')\varphi_{\mathcal{C}}(cc')c \otimes (bb' - \varphi_{\mathcal{B}}(bb')) \otimes c' + \varphi_{\mathcal{B}}(bb')(cc' - \varphi_{\mathcal{C}}(cc'))$$

Il est un peu fastidieux, mais direct, de vérifier, par récurrence sur la somme de leurs longueurs, l'associativité sur les mots, qui s'étend distributivement pour faire de $\mathcal{A}$ une $\mathbb{C}$-algèbre unitaire d'élément neutre $1 + 0$. De plus, les inclusions unitaires

$$\iota_\alpha : \mathcal{A}_\alpha \mapsto \mathcal{A}; a \mapsto \varphi_\alpha(a) + (a - \varphi_\alpha(a))$$

montrent qu'on peut voir chaque $\mathcal{A}_\alpha$ comme une sous-algèbre de $\mathcal{A}$. En définissant finalement φ comme la projection $\varphi(z+l) = z$ où z est la partie constante et l est la partie centrée, on voit que $\varphi(1) = 1$, et $\varphi \circ \iota_\alpha = \varphi_\alpha$, ce qui correspond exactement à la propriété universelle qu'on attend d'un *produit libre* dans le contexte des algèbres unitaires.

Définition 2.10. *L'espace de probabilités non-commutatif $(\mathcal{A}, \varphi)$ est appelé le* ***produit libre*** *des $(\mathcal{A}_\alpha, \varphi_\alpha)$. On voit immédiatement par la construction que les $(\iota_\alpha(\mathcal{A}_\alpha))_{\alpha\in\mathcal{A}}$ sont des sous-algèbres librement indépendantes de $\mathcal{A}$.*

Cette construction n'est pas qu'un amusement formel : elle permet entre autres de construire rigoureusement des familles de variables aléatoires librement indépendantes et identiquement distribuées. Il est en outre possible de montrer que si tous les $\mathcal{A}_\alpha$ sont des $*$-espaces de probabilités, alors le produit libre est un $*$-espace de probabilités, avec l'involution définie sur les mots par

$$(a_1 \otimes \cdots \otimes a_n)^* = a_n^* \otimes \cdots \otimes a_1^*$$

2.3. **Liberté asymptotique pour les grandes matrices.**

Définition 2.11. *Soit I un ensemble, et P une partition de I. On se donne une suite $(\mathcal{A}_k, \varphi_k)_{k\in\mathbb{N}}$ d'espaces de probabilités non-commutatifs, et une suite de familles $((a_\alpha^{(k)})_{\alpha\in I})_{k\in\mathbb{N}}$ de variables aléatoires, avec chaque $a_\alpha^{(k)} \in \mathcal{A}_k$. On dit que cette suite est* ***asymptotiquement librement indépendante****, ou plus simplement* ***asymptotiquement libre*** *par rapport à la partition P s'il existe un espace de probabilités non-commutatif $(\mathcal{A}, \varphi)$ et une famille de variables de $\mathcal{A}$ $(a_\alpha)_{\alpha\in I}$ telles que $(a_\alpha^{(k)})_{\alpha\in I}$ converge en distribution non-commutative vers $(a_\alpha)_{\alpha\in I}$, et telles que les familles $((a_\beta)_{\beta\in p})_{p\in P}$ soient libres dans $\mathcal{A}$.*

Une découverte importante de Voiculescu dans [Voi91] est un ensemble de résultats concernant le comportement asymptotique des distributions de certaines grandes matrices aléatoires quand leur dimension tend vers $+\infty$, qui s'inscrit maintenant dans une famille plus large de résultats abondants dans ce sens. Attelons nous ici à la tâche d'esquisser le cas le plus simple, celui de deux matrices gaussiennes hermitiennes et indépendantes. La preuve dans ce cas s'appuie sur des considérations combinatoires.

Proposition 2.12. *Soit $X = (X_1, \ldots, X_n)$ un vecteur gaussien centré. Alors pour tous $1 \leqslant i_1, \ldots i_k \leqslant n$, on a*

$$\mathbb{E}[X_{i_1} \cdots X_{i_k}] = \sum_{\pi\in\mathcal{P}_2(k)} \prod_{(r,s)\in\pi} \mathbb{E}[X_{i_r} X_{i_s}]$$

Où $\mathcal{P}_2$ désigne l'ensemble des accouplements de $\{1, \ldots, k\}$, c'est-à-dire les partitions en classes de 2 éléments. Ceci est la formule de Wick.

Démonstration. On esquisse rapidement la preuve, due à Nelson [Nel73]. Les deux côtés sont des formes linéaires en chaque X_i, et symétriques (sur le sous-espace des variables gaussiennes centrées, si on veut). On a la formule de polarisation pour les formes multilinéaires symétriques :

$$\Phi(x_1, \ldots, x_n) = \frac{1}{n!} \sum_{k=1}^{n} \sum_{1\leqslant i_1 \leqslant \cdots \leqslant i_k \leqslant n} (-1)^{n-k} \Phi(x_{i_1} + \cdots + x_{i_k}, \ldots, x_{i_1} + \cdots + x_{i_k})$$

Dans le cas où $k = 2l$ est pair, on se ramène par polarisation à vérifier que $\mathbb{E}[X^{2l}] = C_{2l}\mathbb{E}[X^2]^l$ où C_{2l} est le nombre d'accouplements de $\{1, \ldots, 2l\}$. Le côté gauche est un exercice d'intégration par parties, l'autre se montre facilement par récurrence.

Le cas impair suit en observant d'un côté l'imparité de la loi de $X_{i_1} \cdots X_{i_{2l+1}}$, et de l'autre l'absence d'accouplements d'un ensemble de cardinal impair.

□

On va s'intéresser au cas des matrices $A \in M_n(L^{\infty-}(\Omega, \mathcal{F}, \mathbb{P}))$, définies par les conditions suivantes.

A est *gaussienne* : chacun de ses coefficients est une variable gaussienne complexe centrée, au sens où ses parties réelles et imaginaires sont gaussiennes centrées.

A est *hermitienne* : $A_{ij} = \overline{A_{ji}}$ $\forall 1 \leqslant i \leqslant j \leqslant n$.

La matrice des covariances de A est donnée par

$$\mathbb{E}[A_{ij}A_{kl}] = \frac{1}{n}\delta_{il}\delta_{jk} \ \forall 1 \leqslant i, j, k, l \leqslant n$$

Cette condition est en particulier réalisée lorsque les A_{ii} sont indépendantes de même loi $\mathcal{N}(0, \frac{1}{n})$, et où les termes non-diagonaux sont indépendants au-dessus de la diagonale, avec, pour $i < j$, $A_{ij} \sim \frac{1}{\sqrt{2n}}(G + iF)$ avec G, F deux gaussiennes standard indépendantes, et $A_{ji} = \overline{A_{ij}}$. Pour ne pas alourdir les notations, on omet la dimension n, mais il faut considérer A comme un terme d'une suite $(A_n)_{n\in\mathbb{N}}$ de matrices aléatoires dont la taille tend vers l'infini.

On pose $\varphi(A) = \mathbb{E}[\mathrm{tr}(A)]$. On peut calculer les moments de A.

$$\begin{aligned}
\varphi(A^m) &= \frac{1}{n} \sum_{1\leqslant i_1 \leqslant \cdots \leqslant i_m \leqslant n} \mathbb{E}[A_{i_1 i_2} \cdots A_{i_m i_1}] \\
\text{(Par Wick avec } i_{m+1} := i_1 \text{ :)} \quad &= \frac{1}{n} \sum_{1\leqslant i_1 \leqslant \cdots \leqslant i_m \leqslant n} \sum_{\pi\in\mathcal{P}_2(m)} \prod_{(r,s)\in\pi} \mathbb{E}[A_{i_r i_{r+1}} A_{i_s i_{s+1}}] \\
&= \frac{1}{n} \sum_{1\leqslant i_1 \leqslant \cdots \leqslant i_m \leqslant n} \sum_{\pi\in\mathcal{P}_2(m)} \prod_{(r,s)\in\pi} \frac{\delta_{i_r i_{s+1}}\delta_{i_{r+1} i_s}}{n} \\
&= \frac{1}{n^{1+\frac{m}{2}}} \sum_{1\leqslant i_1 \leqslant \cdots \leqslant i_m \leqslant n} \sum_{\pi\in\mathcal{P}_2(m)} \prod_{(r,s)\in\pi} \delta_{i_r i_{s+1}}\delta_{i_{r+1} i_s} \\
&= \frac{1}{n^{1+\frac{m}{2}}} \sum_{\pi\in\mathcal{P}_2(m)} \sum_{1\leqslant i_1 \leqslant \cdots \leqslant i_m \leqslant n} \prod_{(r,s)\in\pi} \delta_{i_r i_{s+1}}\delta_{i_{r+1} i_s}
\end{aligned}$$

En identifiant chaque accouplement $\pi \in \mathcal{P}_2(m)$ à un produit de 2-cycles dans S_m, donné par $\pi(r) = s, \pi(s) = r \iff \{r, s\} \in \pi$, et en notant $\gamma = (1\ 2\ \ldots\ m) \in S_m$, on peut encore écrire

$$\begin{aligned}
\varphi(A^m) &= \frac{1}{n^{1+\frac{m}{2}}} \sum_{\pi\in\mathcal{P}_2(m)} \sum_{1\leqslant i_1 \leqslant \cdots \leqslant i_m \leqslant n} \prod_{r=1}^{m} \delta_{i_r i_{\pi(r)+1}} \\
&= \frac{1}{n^{1+\frac{m}{2}}} \sum_{\pi\in\mathcal{P}_2(m)} \sum_{1\leqslant i_1 \leqslant \cdots \leqslant i_m \leqslant n} \prod_{r=1}^{m} \delta_{i_r i_{\gamma\pi(r)}}
\end{aligned}$$

La condition

$$\prod_{r=1}^{m} \delta_{i_r i_{\gamma\pi(r)}}$$

impose exactement que la fonction $j \mapsto i_j$ est constante sur les cycles de $\gamma\pi$. Autrement dit, c'est une fonction de l'ensemble des cycles de $\gamma\pi$ dans $\{1, \dots, n\}$. La quantité

$$\sum_{1 \leqslant i_1 \leqslant \dots \leqslant i_m \leqslant n} \prod_{r=1}^{m} \delta_{i_r i_{\gamma\pi(r)}}$$

est donc le cardinal de l'ensemble des fonctions sur les cycles de $\gamma\pi$ à valeurs dans $\{1, \dots, n\}$, soit $n^{\#(\gamma\pi)}$, où on note $\#(\sigma)$ le nombre de cycles de $\sigma \in S_m$. Finalement, on obtient

$$\varphi(A^m) = \sum_{\pi \in \mathcal{P}_2(m)} n^{\#(\gamma\pi) - 1 - \frac{m}{2}}$$

Un argument similaire montre que pour deux matrices gaussiennes $A^{(1)}, A^{(2)}$ indépendantes, de la forme prescrite plus haut, les moments mixtes sont donnés, pour $(p_i)_{i=1\dots m} \in \{1, 2\}^m$, par

$$\varphi(A^{(p_1)} \cdots A^{(p_m)}) = \sum_{\pi \in \mathcal{P}_2^{(p)}(m)} N^{\#(\gamma\pi) - 1 - \frac{m}{2}}$$

Où cette fois la somme est sur les accouplements $\pi \leqslant \Pi$ où Π est la partition associée à la relation $p_i \sim p_j \iff p_i = p_j$ (autrement dit ceux qui accouplent seulement des indices correspondant à la même matrice).

Il est clair que les moments d'ordre impair dans le cas d'une matrice, comme dans le cas mixte, sont nuls. En outre, par des arguments combinatoires, on peut montrer que pour $m = 2k$, $\#(\gamma\pi) \leqslant k + 1$, et que cette borne est atteinte exactement quand π a la propriété particulière d'être un accouplement *non-croisé*, ce qui signifie qu'il n'existe pas d'indices $1 \leqslant r < s < t < u \leqslant m$ avec $\{r, t\} \in \pi$ et $\{s, u\} \in \pi$. Asymptotiquement lorque $n \to \infty$, le moment $\varphi(A^m)$ compte donc le nombre D_m d'accouplements non-croisés de $\{1, \dots, m\}$, et le moment mixte $\varphi(A^{(p_1)} \cdots A^{(p_m)})$ compte le nombre $D_m^{(p)}$ d'accouplements non-croisés de $\mathcal{P}_2^{(p)}(m)$. Il se trouve que D_m est donné par le nombre de Catalan $C_{m/2}$, et que la théorie combinatoire de Speicher pour l'indépendance libre prescrit[Spe06, Ch. 11] que les moments mixtes d'une paire d'éléments semi-circulaires standards sont exactement donnés par les $D_m^{(p)}$. On en déduit :

Théorème 2.13. *$A \xrightarrow{n\to\infty}$ a en distribution non-commutative, où a est un élément semi-circulaire standard tel que décrit dans l'exemple 10.*

$(A^{(1)}, A^{(2)}) \xrightarrow{n\to\infty} (a, b)$ en distribution non-commutative où (a, b) est une paire librement indépendante d'éléments semi-circulaires standard. En particulier $A^{(1)}, A^{(2)}$ sont asymptotiquement libres.

Il existe de nombreuses variantes de ces résultats, en prenant par exemple une des matrices déterministes, où des lois de Haar sur $\mathcal{U}_n$. Les lois limites peuvent varier, mais ce qui transparaît, de manière très vague, est que sous de nombreuses conditions d'invariance unitaire des distributions jointes en jeu, les grandes matrices se comportent approximativement comme des variables librement indépendantes. Ceci fournit le terreau intuitif pour le développement d'une notion d'entropie pour les variables non-commutatives, via un processus d'approximation par des matrices.

3. L'entropie libre

La notion d'entropie libre telle que nous la discutons ici a été introduite par Dan Voiculescu[Voi94]. L'objectif était de pouvoir définir des invariants pour attaquer le problème de la classification des algèbres $L(\mathbb{F}_n)$, mais la formulation du concept est vraiment probabiliste, et non sans lien avec l'entropie de Shannon et la thermodynamique. L'objectif de cette section est de présenter l'entropie de Shannon de manière analogue à la définition de l'entropie libre par Voiculescu, puis d'introduire celle-ci et de discuter de quelques unes de ses propriétés.

3.1. **L'entropie de Shannon.** Commencons par examiner l'approche microscopique de l'entropie dans le cas classique. L'approche est essentiellement donnée par la preuve du théorème de Sanov dans [Zei98], et suit la démarche de [Aus17]. Dans cette section, on se place dans le cadre d'un ensemble fini A, et on pose $\mathcal{M}_1(A)$ l'ensemble des mesures de probabilités sur A. C'est un espace compact pour la topologie faible qui coïncide dans ce cas avec la topologie induite par $\mathbb{R}^A$ (on voit $\mathcal{M}_1(A)$ comme un simplexe). Étant donné $p \in \mathcal{M}_1(A)$, une question naturelle est celle de savoir si $p = \sum_{a\in A} p_a\delta_a$ est une distribution *typique*. Une manière d'aborder la question est de se demander, parmis les $|A|^n$ suites de n éléments à valeurs dans A, quelle proportion ont une distribution empirique *proche* de p. Notons donc, pour $u \in A^n$, $p(u) = \sum_{a\in A} \frac{u^{-1}(\{a\})}{n}\delta_a$ sa distribution empirique.

On peut d'abord essayer de compter le nombre exact de u telles que $p(u) = p$. Ainsi on définit les ensembles

$$\Gamma(p; n) = \{u \in A^n : p(u) = p\}$$

Si p est irrationnelle, au sens où il existe un $p_a \notin \mathbb{Q}$, on a clairement $\Gamma(p; n) = \varnothing\ \forall n$, et de même si il existe un p_a tel que $p_a \neq \frac{k}{n}$, $\forall k \in \mathbb{N}$, on a encore $\Gamma(p; n) = \varnothing$. On dit donc que n est un dénominateur admissible pour p si chaque $p_a = \frac{k_a}{n}$ pour un certain $k_a \in \mathbb{N}$, et en particulier $np_a \in \mathbb{N}\ \forall a$. Dans ce cas, le cardinal de $\Gamma(p; n)$ est donné par le coefficient multinomial :

$$|\Gamma(p; n)| = \binom{n}{(p_a n)_{a\in A}} = \frac{n!}{\prod_{a\in A}(p_a n)!}$$

En particulier, on peut estimer le logarithme de cette quantité grâce à la formule de Stirling pour obtenir

$$\frac{1}{n}\log|\Gamma(p; n)| \to H(p) := -\sum_{a\in A} p_a \log p_a$$

Le long d'une suite de dénominateurs admissibles tendant vers $+\infty$. En fait, on a les estimations :

Proposition 3.1. *Pour n un dénominateur admissible pour p,*

$$e^{nH(p)-o(n)} \leqslant |\Gamma(p; n)| \leqslant e^{nH(p)}$$

Démonstration. En notant $p^{\otimes n}$ la n-ième puissance de p sur A^n, on a

$$\forall\ u \in \Gamma(p; n), p^{\otimes n}(u) = \prod_{i=1}^{n} p_{u_i} = \prod_{a\in A} p_a^{|u^{-1}(a)|} = \prod_{a\in A} p_a^{np_a} = e^{-nH(p)}$$

Comme $p^{\otimes n}(\Gamma(p; n)) \leqslant 1$, on obtient en particulier que

$$p^{\otimes n}(\Gamma(p; n)) = \sum_{u\in\Gamma(p;n)} p^{\otimes n}(u) = |\Gamma(p; n)|e^{-nH(p)} \implies |\Gamma(p; n)| \leqslant e^{nH(p)}$$

La borne supérieure est prouvée.

Pour la borne inférieure, on note que $\Gamma(p;n)$ est l'état le plus probable sous $p^{\otimes n}$. Soit $q \in \mathcal{M}_1(A)$. Pour $v \in \Gamma(q;n)$, on a $p^{\otimes n}(v) = \prod_{a\in A} p_a^{nq_a}$, ce qui donne :

$$\frac{p^{\otimes n}(\Gamma(p;n))}{p^{\otimes n}(\Gamma(q;n))} = \frac{n!\prod_a p_a^{np_a}\prod_a (q_a n)!}{n!\prod_a p_a^{nq_a}\prod_a (p_a n)!}$$

$$= \prod_{a\in A} \frac{(q_a n)!}{(p_a n)!} p_a^{np_a - nq_a}$$

En utilisant l'inégalité suivante avec $p_a n$ et $q_a n$:

$$\frac{m!}{n!} = \begin{cases} (n+1)(n+2)\cdots(m) \geqslant n^{m-n} & m \geqslant n \\ \frac{1}{m+1}\frac{1}{m+2}\cdots\frac{1}{n} \geqslant n^{m-n} & m < n \end{cases}$$

on obtient

$$\frac{p^{\otimes n}(\Gamma(p;n))}{p^{\otimes n}(\Gamma(q;n))} \geqslant \prod_{a\in A} (p_a n)^{nq_a - np_a} p_a^{np_a - nq_a} = \prod_{a\in A} n^{nq_a - np_a} = n^{n\sum(q_a - p_a)} = n^0 = 1$$

Comme pour chaque $u \in A^n$, on a $u \in \Gamma(p(u);n)$, il s'ensuit que

$$1 = p^{\otimes}(A^n) = p^{\otimes}\left(\bigcup_{q\in\mathcal{M}_1} \Gamma(q;n)\right)$$

et que cette union est disjointe et finie, puisque l'ensemble des mesures q admettant n comme dénominateur admissible est contenu dans $\{\frac{0}{n}, \frac{1}{n}, \ldots, \frac{n-1}{n}, \frac{n}{n}\}^{|A|}$, de cardinal $(n+1)^{|A|}$. On a alors :

$$1 = \sum_{q\in\mathcal{M}_1, q \text{ admettant } n} p^{\otimes n}(\Gamma(q;n)) \leqslant (n+1)^{|A|} p^{\otimes n}(\Gamma(p;n)) = (n+1)^{|A|}|\Gamma(p;n)|e^{-nH(p)}$$

$$\implies |\Gamma(p;n)| \geqslant (n+1)^{-|A|} e^{nH(p)} = e^{nH(p) - o(n)}$$

□

La deuxième tentative consiste à relâcher notre définition pour se permettre approcher p jusqu'à à un certain degré d'approximation ϵ, et en particulier pouvoir traiter les mesures irationnelles. On pose donc

$$\Gamma(p;n,\epsilon) = \{u \in A^n : \|p(u) - p\| < \epsilon\}$$

Où $\|p - q\| = \sum_{a\in A} |p_a - q_a|$ est la distance en variation totale de p à q, qui est aussi la distance L^1 sur $\mathbb{R}^{|A|}$.
Alors on a :

Proposition 3.2.

$$e^{nH(p)+o(n)} \leqslant |\Gamma(p;n,\epsilon)| \leqslant e^{nH(p)+nh(\epsilon)+o(n)}$$

Où $h(\epsilon) \xrightarrow{\epsilon\to 0} 0$. *Par conséquent :*

$$\lim_{\epsilon\to 0}\lim_{n\to\infty} \frac{1}{n}\log|\Gamma(p;n,\epsilon)| = H(p)$$

Démonstration. En effet, $p \mapsto H(p)$ est continue, donc uniformément continue sur le compact $\mathcal{M}_1(A)$. En notant

$$h(\epsilon) = \sup_{\|p-q\|<\epsilon} |H(p) - H(q)| \xrightarrow{\epsilon\to 0} 0$$

On a :

$$\Gamma(p;n,\epsilon) = \bigcup_{\|p-q\|<\epsilon} \Gamma(q;n)$$

Seuls les q qui ont un dénominateur admissible n contribuent à cette union, qui est disjointe. De plus, pour un tel q, on a

$$|\Gamma(q;n)| \leqslant e^{nH(q)} \leqslant e^{n(H(p)+h(\epsilon))} \implies |\Gamma(p;n,\epsilon)| \leqslant |\{q : \|q-p\| < \epsilon, n \text{ admissible pour } q\}| e^{n(H(p)+h(\epsilon))}$$

Puis en bornant grossièrement le cardinal de cet ensemble, on a :

$$|\Gamma(p;n,\epsilon)| \leqslant (n+1)^{|A|} e^{nH(n)+nh(\epsilon)} = e^{nH(n)+nh(\epsilon)+o(n)}$$

Pour la borne inférieure, on approxime p de la façon suivante : fixons une énumération de $A = \{a_1, \cdots, a_{|A|}\}$, et définissons $q^n \in \mathcal{M}_1(A)$ par

$$q^n_{a_i} = \frac{\lfloor np_{a_i} \rfloor}{n} \ \forall 1 \leqslant i < |A|, \ q^n_{a_{|A|}} = 1 - (q^n_{a_1} + \cdots + q^n_{a_{|A|-1}})$$

On a

$$\begin{aligned} \|q^n - p\| &= \sum_{i=1}^{|A|-1} \left| \frac{\lfloor np_{a_i} \rfloor}{n} - p_{a_i} \right| + |q^n_{a_{|A|}} - p_{a_{|A|}}| \\ &\leqslant 2 \sum_{i=1}^{|A|-1} \left| \frac{\lfloor np_{a_i} \rfloor}{n} - p_{a_i} \right| \leqslant \frac{2|A|}{n} \end{aligned}$$

Pour $\epsilon > 0$ fixé, prenons n tel que $\|q^n - p\| < \epsilon$. On a alors, comme $\Gamma(p;n,\epsilon) = \bigcup_{\|p-q\|<\epsilon} \Gamma(q;n)$,

$$|\Gamma(p;n,\epsilon)| \geqslant |\Gamma(q^n;n)| \geqslant e^{nH(q^n)-o(n)} \geqslant e^{n(H(p)-h(\|q^n-p\|)-o(n))}$$

Il suffit à présent d'observer que $h(\|q^n - p\|) \to 0 \implies \frac{n}{n} h(\|q^n - p\|) \to 0 \implies nh(\|q^n - p\|) = o(n)$ pour obtenir

$$|\Gamma(p;n,\epsilon)| \geqslant e^{nH(p)-o(n)}$$

□

La quantité $H(p) = -\sum_{a \in A} p_a \log p_a$ est appelée **entropie de Shannon** de p. L'approche de Shannon pour la définition de H n'était cependant pas combinatoire, mais axiomatique : il a postulé certaines propriétés que devrait avoir une mesure $H(p)$ de l'information apportée par l'observation d'une variable aléatoire de loi p, puis montré que $p \mapsto H(p)$ était la seule fonction satisfaisant ses postulats. En un sens, le chemin emprunté ci-dessus est plus proche de l'idée d'entropie thermodynamique : si la distribution p correspond au *macro-état* du système observé, les fonctions A^n correspondent à des *micro-états* discrets de taille n, dont la distribution empirique se conforme plus ou moins bien à l'observation p. L'ensemble $\Gamma(p;n,\epsilon)$ s'interprète donc comme l'ensemble des micro-états qui concordent avec l'observation à un degré de précision en deçà d'un seuil de tolérance fixé. Si on interprète la quantité $\frac{|\Gamma(p;n,\epsilon)|}{|A|^n}$ géométriquement comme la proportion des micro-états qui se conforment à l'observation, on s'attend à ce que cette quantité décroisse géométriquement avec n, la *finesse de la discrétisation*, si bien qu'il est naturel de regarder à l'échelle $\frac{1}{n} \log \cdot$ cette quantité pour obtenir à la limite, et à une constante additive près, l'entropie $H(p)$.

3.2. **Micro-états matriciaux et entropie libre.** L'idée de Voiculescu s'inspire de la même idée thermodynamique, excepté que les objets à approcher par des micro-états ne sont plus des distributions classiques, mais des distributions non-commutatives.

On se place à partir de maintenant dans un W*-espace de probabilités fini tracial $(\mathcal{M}, \tau)$, c'est-à-dire que $\mathcal{M}$ est une algèbre de von Neumann finie, et on fixe n éléments auto-adjoints $X_1, \cdots, X_n \in \mathcal{M}$, et on note $\|X_i\| = \sqrt{\tau(X^2)}$.

L'ensemble $M_k(\mathbb{C})^{\text{sa}}$ des matrices auto-adjointes de taille k est un sous-espace de $M_k(\mathbb{C})$, admettant la $\mathbb{R}$-base

$$\{E_{ii}, 1 \leqslant i \leqslant k\} \cup \{R_{ij} := \frac{1}{\sqrt{2}}(E_{ij} + E_{ji}), \leqslant j < i \leqslant k\} \cup \{I_{ij} := \frac{i}{\sqrt{2}}(E_{ij} - E_{ji}), 1 \leqslant j < i \leqslant k\}$$

On peut donc le voir comme un espace euclidien de dimension $k + 2\frac{(k-1)k}{2} = k^2$, et on voit que la norme euclidienne associée n'est autre que la norme de Frobenius. On pose de plus λ_{k^2} la mesure de Lebesgue sur $M_k(\mathbb{C})^{\text{sa}}$. On note également, pour $A \in M_k(\mathbb{C})^{\text{sa}}$, $\|A\| = \sqrt{\text{tr}(A^2)} = \frac{1}{\sqrt{n}}\|A\|_{\text{F}}$, où $\|\cdot\|_{\text{F}}$ est la norme de Frobenius.

Définition 3.3. *On définit, pour $R, \epsilon > 0$, et $m, k \in \mathbb{N}$ l'ensemble :*

$$\Gamma_R(X_1, \cdots, X_n; m, k, \epsilon)$$

comme l'ensemble des n-uplets $(A_1, \cdots, A_n)$ $M_k(\mathbb{C})^{sa}$ tels que :

$$\|A_i\| \leqslant R \ \forall 1 \leqslant i \leqslant n$$

$$|\tau(X_{i_1} \cdots X_{i_p}) - \text{tr}(A_{i_1} \cdots A_{i_p})| < \epsilon \ \forall 1 \leqslant p \leqslant m \ \forall i_1, \ldots, i_p \in \{1, \ldots, n\}$$

On définit de plus successivement les quantités :

$$\chi_R(X_1, \ldots, X_n; m, k, \epsilon) = \log \lambda_{nk^2}(\Gamma_R(X_1, \cdots, X_n; m, k, \epsilon))$$

$$\chi_R(X_1, \ldots, X_n; m, \epsilon) = \limsup_{k \to \infty} \frac{1}{k^2} \chi_R(X_1, \ldots, X_n; m, k, \epsilon) + \frac{n}{2} \log k$$

$$(\triangle) \qquad \chi_R(X_1, \ldots, X_n) = \lim_{\epsilon \to 0, m \to \infty} \chi_R(X_1, \ldots, X_n; m, \epsilon)$$

$$\chi(X_1, \ldots, X_n) = \sup_{R > 0} \chi_R(X_1, \ldots, X_n)$$

C'est cette dernière quantité qu'on nomme ***entropie libre*** *de $(X_1, \ldots, X_n)$.*

Un élément de $\Gamma_R(X_1, \ldots, X_n; m, k, \epsilon)$ est un *micro-état approchant* matriciel. Le paramètre k s'interprète comme le degré de finesse de la discrétisation, m, ϵ correspondent au degré d'approximation tolérée sur les moments, et R est une borne fixée pour pouvoir travailler dans un domaine de mesure finie. Comme dans le cas classique, l'entropie libre de $(X_1, \ldots, X_n)$ est alors définie par une succession de limites, en faisant d'abord tendre k vers l'infini, puis le degré de tolérance vers 0.

La condition sur les normes nous disent que pour $1 \leqslant i \leqslant n$, $\text{tr}(A_i^2) = \frac{1}{k}\text{Tr}(A_i^2) \leqslant R^2$. Ceci implique que $A_i \in B_{k^2}(0, \sqrt{k}R)$ pour la norme euclidienne sur $\mathbb{R}^{k^2}$, et donc que $\Gamma_R(X_1, \ldots, X_n; m, k, \epsilon)$ est inclus dans le produit de ces n boules P. On a donc

$$\log|P| \approx n \log|B_{k^2}(0, \sqrt{k}R)| = n \log V_{k^2} + nk^2 \log(\sqrt{k}R) \approx n \log V_{k^2} + \frac{nk^2}{2} \log k$$

Au premier ordre d'approximation (on omet un terme en $O(k^2)$), et où V_{k^2} est le volume de la boule unité de $\mathbb{R}^{k^2}$. Or par la formule de Stirling, on a

$$\log V_{k^2} \approx \frac{k^2}{2} \log \pi - (\frac{k^2}{2} + 1) \log(\frac{k^2}{2}) \approx -k^2 \log k$$

encore au premier ordre, ce qui nous donne, à une constante additive près

$$\frac{1}{k^2}\log|P| \approx -\frac{n}{2}\log k$$

Cette interprétation géométrique, exactement comme dans le cas classique, justifie l'échelle asymptotique choisie lorsqu'on a pris la limite du pas de discrétisation $k \to \infty$.

Pour voir que la limite $(\triangle)$ est bien définie, on observe que pour $m \leqslant m'$, on a

$$\Gamma_R(X_1,\dots,X_n;m',k,\epsilon) \subset \Gamma_R(X_1,\dots,X_n;m,k,\epsilon)$$

, et que pour $0 < \epsilon < \epsilon'$, on a

$$\Gamma_R(X_1,\dots,X_n;m',k,\epsilon) \subset \Gamma_R(X_1,\dots,X_n;m,k,\epsilon')$$

si bien que l'expression dont on prend la limite $(\triangle)$ tend de façon décroissante vers sa borne inférieure.

Notons que χ ne dépend,comme dans le cas classique, que de la $*$-distribution jointe de $(X_1,\dots,X_n)$, il faut donc voir l'entropie libre comme une propriété de $(\tau, X_1,\dots,X_n)$, en particulier elle ne dépend pas de la représentation choisie. Notons également que, par tracialité de tr, Les ensembles $\Gamma_R(X_1,\dots,X_n;m,k,\epsilon)$ sont invariants par conjugaison unitaire : si $U \in M_k(\mathbb{C})$ est unitaire, $(A_1,\dots,A_n) \in \Gamma_R(X_1,\dots,X_n;m,k,\epsilon)$, alors $(U^*A_1U,\dots,U^*A_n,U) \in \Gamma_R(X_1,\dots,X_n;m,k,\epsilon)$, puisque $\|U^*AU\| = \|A\|$, et que toutes les conditions sur les moments se simplifient par invariance de tr sous l'action de $\mathcal{U}_n$.

Proposition 3.4. *On a l'inégalité suivante, où* $C^2 = \tau(X_1^2 + \cdots + X_n^2)$

$$\chi(X_1,\dots,X_n) \leqslant \frac{n}{2}\log(2\pi e\frac{C^2}{n})$$

En particulier, l'entropie libre est soit finie, soit égale à $-\infty$.

Démonstration. Il suffit de montrer

$$\chi_R(X_1,\dots,X_n;m,k,\epsilon) \leqslant \frac{nk^2}{2}(\log(2\pi e\frac{C^2+n\epsilon}{n}) - \log k)$$

En effet, comme $\limsup_{k\to\infty} \frac{nk^2}{2k^2}(\log(2\pi e\frac{C^2+n\epsilon}{n}) - \log k) + \frac{n}{2}\log k = \frac{n}{2}\log(2\pi e\frac{C^2+n\epsilon}{n})$, cette inégalité impliquera

$$\chi_R(X_1,\dots,X_n;m,\epsilon) \leqslant \frac{n}{2}\log(2\pi e\frac{C^2+n\epsilon}{n})$$

puis la conclusion, puisque cette borne est uniforme en $m \geqslant 2$ et en R, et continue en 0 par rapport à ϵ.

Le résultat s'appuie sur une inégalité que Voiculescu appelle l'inégalité de Shannon, prouvée dans le lemme ci-dessous.

Si $\Gamma := \Gamma_R(X_1,\cdots,X_n;m,k,\epsilon)$, est vide, il n'y a rien à prouver. Sinon, on pose $\lambda := \lambda_{nk^2}, |\Gamma| := \lambda(\Gamma)$, et avec $f(x) = \frac{1}{|\Gamma|}\mathbb{1}_\Gamma(x)$ la densité uniforme sur Γ, l'inégalité de Shannon implique :

$$-\frac{1}{|\Gamma|}\int_\Gamma \log\frac{1}{|\Gamma|}\mathrm{d}\lambda = \log|\Gamma| \leqslant \frac{nk^2}{2}\log(\frac{2\pi e a^2}{nk^2})$$

où

$$a^2 = \frac{1}{|\Gamma|}\int_\Gamma \|(A_1,\dots,A_n)\|^2\lambda(\mathrm{d}(A_1,\dots,A_n))$$

(La norme ci-dessus est la norme euclidienne). Or, $(A_1, \dots, A_n) \in \Gamma,\ m \geqslant 2$ implique $|\frac{1}{k}\|A_i\|^2 - \tau(X_i^2)| < \epsilon\ \forall 1 \leqslant i \leqslant n$. Il s'ensuit que

$$\frac{1}{k}\|(A_1, \cdots, A_n)\|^2 \leqslant \sum_{i=1}^{n}(\tau(X_i^2) + \epsilon) = C^2 + n\epsilon$$

On conclut, puisque on a alors

$$a^2 \leqslant k(C^2 + n\epsilon)$$

puis

$$\log|\Gamma| \leqslant \frac{nk^2}{2}\log(\frac{2\pi ek(C^2+n\epsilon)}{nk^2}) = \frac{nk^2}{2}(\log(\frac{2\pi e(C^2+n\epsilon)}{n}) - \log k)$$

□

Lemme 3.5. *Montrons l'inégalité de Shannon. Soit f une densité de probabilités sur $\mathbb{R}^p$, on définit l'entropie différentielle de f :*

$$H(f) = -\int_{\mathbb{R}^p} f(x_1, \dots, x_p)\log(f(x_1, \dots, x_p))dx_1 \dots dx_p := -\int_{\mathbb{R}^p} f\log f \mathbb{1}_{\{f>0\}} d\lambda_p$$

Notons $a^2 = \int_{\mathbb{R}^p}(x_1^2 + \cdots + x_p^2)f(x_1, \dots, x_p)dx_1 \dots dx_p$. Alors on a l'inégalité de Shannon :

$$H(f) \leqslant \frac{1}{2}p\log(\frac{2\pi ea^2}{p})$$

Démonstration. On commence par rappeler quelques propriétés de l'entropie. Soient f, g deux densités de probabilités sur $\mathbb{R}^p$ telles que f soit absolument continue par rapport à g (au sens de leur mesures correspondantes respectives).

On définit l'entropie relative :

$$\mathrm{D_{KL}}(f\|g) = \int_{\mathbb{R}^p} f(x)\log(\frac{f(x)}{g(x)})\mathbb{1}_{\{f>0\}}\mathrm{d}x = \int_{\mathbb{R}^p} f(x)\log(\frac{f(x)}{g(x)})\mathrm{d}x$$

En convenant que $0 \cdot \log(0) = 0 = 0 \cdot \log(0/0)$.Remarquons que

$$\begin{aligned}\mathrm{D_{KL}}(f\|g) &= -\int_{\mathbb{R}^p} f(x)\log(\frac{g(x)}{f(x)})\mathrm{d}x \\ \geqslant -\log\left(\int_{\mathbb{R}^p} f(x)\frac{g(x)}{f(x)}\mathrm{d}x\right) &= -\log\left(\int_{\mathbb{R}^p} g(x)\mathrm{d}x\right) \\ &= 0\end{aligned}$$

Où l'inégalité est celle de Jensen. Ceci implique que $\mathrm{D_{KL}}(f\|g)$ est bien définie et positive.

Intéressons nous au cas particulier où f est une densité sur $\mathbb{R}^p \times \mathbb{R}^q$ et g est donnée comme le produit des densités marginales de f sur $\mathbb{R}^p$ et $\mathbb{R}^q$,

$$g(x, y) = \left(\int_{\mathbb{R}^q} f(x, y)\mathrm{d}y\right)\left(\int_{\mathbb{R}^p} f(x, y)\mathrm{d}x\right)$$

Notons respectivement f_p et f_q ces marginales. En terme de lois, si ν est la mesure associée à f, g est la densité associée à $\pi^p_*(\nu) \otimes \pi^q_*(\nu)$, où π^p, π^q sont les projections sur $\mathbb{R}^p$ et $\mathbb{R}^q$, et calculons :

$$\begin{aligned}
\mathrm{D_{KL}}(f\|g) &= \int_{\mathbb{R}^p\times\mathbb{R}^q} f(x,y)\log(\frac{f(x,y)}{f_p(x)f_q(y)})\mathrm{d}x\mathrm{d}y \\
&= \int_{\mathbb{R}^p\times\mathbb{R}^q} f(x,y)\log(f(x,y)) - f(x,y)\log(f_p(x)) - f(x,y)\log(f_q(y))\mathrm{d}x\mathrm{d}y \\
&= -H(f) - \int_{\mathbb{R}^p}\log(f_p(x))\int_{\mathbb{R}^q} f(x,y)\mathrm{d}y\mathrm{d}x - \int_{\mathbb{R}^q}\log(f_q(y))\int_{\mathbb{R}^p} f(x,y)\mathrm{d}x\mathrm{d}y \\
&= -H(f) - \int_{\mathbb{R}^p} f_p(x)\log(f_p(x))\mathrm{d}x - \int_{\mathbb{R}^q} f_q(y)\log(f_q(y))\mathrm{d}y \\
&= -H(f) + H(f_p) + H(f_q) \geqslant 0
\end{aligned}$$

De proche en proche, on en déduit que pour f une densité sur $\mathbb{R}^p$ de densités marginales $f_1, \cdots, f_p$ sur $\mathbb{R}$, on a $H(f) \leqslant H(f_1) + \cdots + H(f_n)$.

De plus, pour g une densité sur $\mathbb{R}$ telle que $\int x^2 g(x)\mathrm{d}x = \sigma^2$, on a $H(g) \leqslant \frac{1}{2}\log(2\pi e\sigma^2)$, et cette borne est atteinte pour g une densité gaussienne de variance σ^2.

En effet, considérons la quantité $\mathrm{D_{KL}}(g\|f)$, où g est une densité de variance σ^2 et f est la densité d'une gaussienne centrée de variance σ^2 (en particulier $\mathrm{d}x \ll f(x)\mathrm{d}x$, donc $g(x)\mathrm{d}x \ll f(x)\mathrm{d}x$). On a :

$$\begin{aligned}
\mathrm{D_{KL}}(g\|f) &= \int_{\mathbb{R}} g(x)\log(\frac{g(x)}{f(x)})\mathrm{d}x \\
&= -H(g) - \int_{\mathbb{R}} g(x)\log(\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}})\mathrm{d}x \\
&= -H(g) - \log(\frac{1}{\sqrt{2\pi\sigma^2}}) + \int_{\mathbb{R}} g(x)\frac{x^2}{2\sigma^2}\mathrm{d}x \\
&= -H(g) - \log(\frac{1}{\sqrt{2\pi\sigma^2}}) + \frac{\sigma^2}{2\sigma^2} \\
&= -H(g) + \frac{1}{2}(\log(2\pi\sigma^2) + 1) \\
&= -H(g) + \frac{1}{2}(\log(2\pi e\sigma^2)) \geqslant 0
\end{aligned}$$

Nous pouvons à présent conclure. On pose $\sigma_i^2 = \int x^2 f_i(x^2)\mathrm{d}x$, si bien que $a^2 = \sigma_1^2 + \cdots + \sigma_p^2$.

$$\begin{aligned}
H(f) &\leqslant \sum_{i=1}^{p} H(f_i) \\
&\leqslant \sum_{i=1}^{p} \frac{1}{2}\log(2\pi e\sigma_i^2) \\
&= \frac{1}{2}\log\left((2\pi e)^p \prod_{i=1}^{p} \sigma_i^2\right) \\
&= \frac{p}{2}\log\left(2\pi e\left(\prod_{i=1}^{p}\sigma_i^2\right)^{\frac{1}{p}}\right) \\
&\leqslant \frac{p}{2}\log\left(\frac{2\pi e\sum_{i=1}^{n}\sigma_i^2}{p}\right) \text{ par l'inégalité arithmético-géométrique} \\
&= \frac{p}{2}\log\left(\frac{2\pi e a^2}{p}\right)
\end{aligned}$$

□

On a l'analogue de la propriété $H(f) \leqslant H(f_p) + H(f_q)$.

Proposition 3.6. *Si $1 \leqslant p \leqslant n$, alors*

$$\chi(X_1, \ldots, X_n) \leqslant \chi(X_1, \ldots, X_p) + \chi(X_{p+1}, \ldots, X_n)$$

Démonstration. On a manifestement l'inclusion

$$\Gamma_R(X_1, \ldots, X_n; m, k, \epsilon) \subset \Gamma_R(X_1, \ldots, X_p; m, k, \epsilon) \times \Gamma_R(X_{p+1}, \ldots, X_n; m, k, \epsilon)$$

La condition sur les normes est évidemment vérifiée, et les conditions sur les moments restent vérifiées en restriction aux choix indices dans $\{1, \ldots, p\}$ et $\{p+1, \ldots, n\}$, respectivement. Ceci implique immédiatement :

$$\log\lambda(\Gamma_R(X_1, \ldots, X_n; m, k, \epsilon)) \leqslant \log\lambda(\Gamma_R(X_1, \ldots, X_p; m, k, \epsilon)) + \log\lambda(\Gamma_R(X_{p+1}, \ldots, X_n; m, k, \epsilon))$$

Par linéarité des opérations limitantes, on a donc successivement

$$\begin{aligned}
\chi_R(X_1, \ldots, X_n; m, k, \epsilon) &\leqslant \chi_R(X_1, \ldots, X_p; m, k, \epsilon) + \chi_R(X_{p+1}, \ldots, X_n; m, k, \epsilon) \\
\chi_R(X_1, \ldots, X_n; m, \epsilon) &\leqslant \chi_R(X_1, \ldots, X_p; m, \epsilon) + \chi_R(X_{p+1}, \ldots, X_n; m, \epsilon) \\
\chi_R(X_1, \ldots, X_n) &\leqslant \chi_R(X_1, \ldots, X_p) + \chi_R(X_{p+1}, \ldots, X_n)
\end{aligned}$$

□

La proposition suivante montre que la fonction $R \mapsto \chi_R(X_1, \ldots, X_n; m, k, \epsilon)$ est stationnaire dès que $R > \max\{\|X_1\|, \ldots, \|X_n\|\}\} := \rho$, si bien que $\chi_R(X_1, \ldots, X_n) = \chi(X_1, \ldots, X_n)$, pour n'importe quel R choisi suffisament grand.

Proposition 3.7. *Si $\rho < R_1 < R_2$, alors*

$$\chi_{R_1}(X_1, \ldots, X_n) = \chi_{R_2}(X_1, \ldots, X_n)$$

Démonstration. On donne une idée de la preuve de Voiculescu.

Fixons $\rho < R_0 < R_1 < R_2$, où $\rho = \max(\|X_1\|, \ldots, \|X_n\|)$. On a automatiquement $\chi_{R_2}(X_1, \ldots, X_n) \geqslant \chi_{R_1}(X_1, \ldots, X_n)$, il suffit donc de montrer l'inégalité réciproque.

Définissons g la fonction affine par morceaux sur $[-R_2, R_2]$ définie par :

$$g : x \mapsto \begin{cases} x & x \in [0, R_0] \\ R_0 + \frac{R_1 - R_0}{R_2 - R_0}(x - R_0) & x \in [R_0, R_2] \\ -g(-x) & x \in [-R_2, 0] \end{cases}$$

Autrement dit g interpole linéairement par morceaux entre les points $(-R_2, -R_1), (-R_0, -R_0), (R_0, R_0)$ et (R_2, R_1). Notons G l'application qui applique g à $(A_1, \ldots, A_n)$ coordonnée par coordonnée. Fixons m et ϵ, $\delta > 0$. Alors on peut montrer par un argument spectral que $\exists \epsilon_1 < \epsilon$, $m_1 \geqslant m$ tels que

$$G(\Gamma_{R_2}(X_1, \ldots, X_n; m_1, k, \epsilon_1)) \subset \Gamma_{R_1}(X_1, \ldots, X_n; m, k, \epsilon) \ \forall k \geqslant 0$$

Avec $\|g(A_i) - A_i\|, \mathrm{tr}(\mu_{A_i}(B(0, R_2) \backslash B(0, R_0))) < \delta$, pour $(A_1, \ldots, A_n) \in \Gamma_{R_2}(X_1, \ldots, X_n; m_1, k, \epsilon_1)$. ($\mu_A$ est la mesure spectrale de A, donc $\mathrm{tr}(\mu_A(B))$ est simplement $\frac{1}{k}|\sigma(A) \cap B|$). On a donc

$$\log \int_{\mathbb{R}^{nk^2}} \mathbb{1}_{G(\Gamma_{R_2}(X_1, \ldots, X_n; m_1, k, \epsilon_1))} \mathrm{d}\lambda_{nk^2} \leqslant \chi_{R_1}(X_1, \ldots, X_n; m, k, \epsilon)) \ \forall k$$

Il s'agit donc de borner

$$\int_{\mathbb{R}^{nk^2}} \mathbb{1}_{G(\Gamma_{R_2}(X_1, \ldots, X_n; m_1, k, \epsilon_1))} = \int_{\Gamma_{R_2}(X_1, \ldots, X_n; m_1, k, \epsilon_1)} |J_G(A_1, \ldots, A_n)| \mathrm{d}\lambda(A_1, \ldots, A_n)$$

Où J_G est le jacobien du changement de variable $(A_1, \ldots, A_n) \leftrightarrow (g(A_1), \ldots, g(A_n))$. On cherche donc à borner ce dernier. C'est un changement de variables diagonal, donc celui-ci est donné par le produit des Jacobiens de $A_i \leftrightarrow g(A_i)$.

L'idée est alors d'exprimer ce dernier via la composition :

$$M_k(\mathbb{C})^{\mathrm{sa}} \xrightarrow{\phi} \mathcal{U}_k/T \times \mathbb{R}^k \xrightarrow{g} \mathcal{U}_k/T \times \mathbb{R}^k \xrightarrow{\phi^{-1}} M_k(\mathbb{C})^{\mathrm{sa}}$$

Qui correspond à la procédure de diagonaliser A, appliquer g à la diagonale, puis de dédiagonaliser A pour revenir au système de coordonnées initiales. ϕ est donnée par $A \mapsto ([U], (\lambda_1, \ldots, \lambda_n))$ où U est une matrice unitaire telle que $U^*AU = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Cependant U n'est pas unique, comme on peut le voir en multipliant U par n'importe quelle matrice unitaire diagonale à gauche, ce qui explique qu'on prenne la classe $[U]$ de U dans $\mathcal{U}_k/T$, où T est le sous-groupe des matrices unitaires diagonales. La formule intégrale de Weyl[And10, p. 190] implique que la contribution de ϕ au Jacobien total au point A est donnée par

$$c \prod_{1 \leqslant i < j \leqslant k} |\lambda_j - \lambda_i|^2$$

Où c est une certaine constante de normalisation. La contribution de g au point $([U], (\lambda_1, \ldots, \lambda_k))$ est ensuite donnée, puisque g n'agit que sur le spectre de A, par

$$\prod_{i=1 \ldots k} g'(\lambda_i) = \left(\frac{R_1 - R_0}{R_2 - R_0}\right)^{|\sigma(A) \backslash B(0, R_0)|}$$

Pour chaque A tel que $R_0 \notin \sigma(A)$, en particulier pour presque tous A. Enfin, la contribution de ϕ^{-1} au point $([U], (g(\lambda_1), \dots, g(\lambda_k))$ est donnée par

$$\frac{1}{c} \prod_{1 \leqslant i < j \leqslant n} |g(\lambda_j) - g(\lambda_i)|^{-2}$$

Finalement, on obtient que le Jacobien de la transformation $A \leftrightarrow g(A)$ est donné par

$$J_A = \prod_{1 \leqslant i < j \leqslant k} \frac{|\lambda_j - \lambda_i|^2}{|g(\lambda_j) - g(\lambda_i)|^2} \left(\frac{R_1 - R_0}{R_2 - R_0} \right)^{|\sigma(A) \setminus B(0, R_0)|}$$

Mais si $|\lambda_i|, |\lambda_j| < R_0$, on a $\frac{|\lambda_j - \lambda_i|^2}{|g(\lambda_j) - g(\lambda_i)|^2} = 1$, et si $|\lambda_i|, |\lambda_j| > R_0$, cette fraction est donnée par $\left(\frac{R_1 - R_0}{R_2 - R_0} \right)^{-2}$. En exploitant la condition

$$\operatorname{tr}(\mu_{A_i}(B(0, R_2) \setminus B(0, R_0))) < \delta$$
$$s := |\sigma(A) \setminus B(0, R_0)| < k\delta$$

On peut arriver à la borne

$$J_A \geqslant \left(\frac{R_1 - R_0}{R_2 - R_0} \right)^{k + k^2(2\delta - \delta^2)}$$

puis

$$\chi_{R_2}(X_1, \dots X_n; m_1, k, \epsilon_1) + n(k + k^2(2\delta - \delta^2)) \log \left(\frac{R_1 - R_0}{R_2 - R_0} \right) \leqslant \log \int_{\mathbb{R}^{nk^2}} \mathbb{1}_{G(\Gamma_{R_2}(X_1, \dots, X_n; m_1, k, \epsilon_1))} \mathrm{d}\lambda_{nk^2}$$
$$\leqslant \chi_{R_1}(X_1, \dots, X_n; m, k, \epsilon)$$

En faisant tendre $k \to \infty$, on obtient

$$\chi_{R_2}(X_1, \dots, X_n; m_1, \epsilon_1) + n(\delta - \delta^2) \log \left(\frac{R_1 - R_0}{R_2 - R_0} \right) \leqslant \chi_{R_1}(X_1, \dots, X_n; m, k, \epsilon)$$

Puis on conclut en faisant tendre $\delta \to 0$ □

Finissons notre discussion de ces quelques propriétés de l'entropie libre par un résultat de semi-continuité.

Proposition 3.8. *Soit $(X_1^{(p)}, \dots, X_n^{(p)})_{p \in \mathbb{N}}$ une suite convergente en distribution non-commutative vers $(X_1, \dots, X_n)$. On suppose que toutes ces variables vivent dans (M, τ). Alors on a*

$$\limsup_{p \to \infty} \chi_R(X_1^{(p)}, \dots, X_1^{(p)}) \leqslant \chi_R(X_1, \dots, X_n)$$

Si $\sup_{p \in \mathbb{N}, 1 \leqslant i \leqslant n} \left\| X_i^{(p)} \right\| < \infty$, on a

$$\limsup_{p \to \infty} \chi(X_1^{(p)}, \dots, X_1^{(p)}) \leqslant \chi(X_1, \dots, X_n)$$

Démonstration. Soient $m \in \mathbb{N}, \epsilon > 0$. Comme pour chaque $P \in \mathbb{C}\langle Z_1, \dots, Z_n \rangle$

$$\lim_{p \to \infty} \tau(P(X_1^{(p)}, \dots, X_n^{(p)})) = \tau(P(X_1, \dots, X_n))$$

Il s'ensuit, comme il n'existe qu'un nombre fini de moments d'ordre au plus m, et en écrivant

$$|\operatorname{tr}(A_{i_1} \dots A_{i_r}) - \tau(X_{i_1} \dots X_{i_n})| \leqslant |\operatorname{tr}(A_{i_1} \dots A_{i_r}) - \tau(X_{i_1}^{(p)} \dots X_{i_n}^{(p)})| + |\tau(X_{i_1}^{(p)} \dots X_{i_n}^{(p)}) - \tau(X_{i_1} \dots X_{i_n})|$$

qu'il existe p_0 tel que $\forall p \geqslant p_0$

$$\Gamma_R(X_1^{(p)}, \dots, X_n^{(p)}; m, k, \epsilon) \subset \Gamma_R(X_1, \dots, X_n; m, k, 2\epsilon)$$

On en déduit

$$\chi_R(X_1^{(p)}, \dots, X_n^{(p)}; m, k, \epsilon) \leqslant \chi_R(X_1, \dots, X_n; m, k, 2\epsilon)$$
$$\chi_R(X_1^{(p)}, \dots, X_n^{(p)}; m, \epsilon) \leqslant \chi_R(X_1, \dots, X_n; m, 2\epsilon)$$

En prenant la limite à gauche, qui est décroissante en m et en ϵ, on obtient

$$\chi_R(X_1^{(p)}), \dots, X_n^{(p)}) \leqslant \chi_R(X_1, \dots, X_n; m, 2\epsilon)$$

Puis, comme la borne est uniforme en p,

$$\limsup_{p\to\infty} \chi_R(X_1^{(p)}), \dots, X_n^{(p)}) \leqslant \chi_R(X_1, \dots, X_n; m, 2\epsilon)$$

On peut maintenant faire tendre m et ϵ vers 0 pour obtenir le résultat. Si on a

$$\sup_{p\in\mathbb{N}, 1\leqslant i\leqslant n} \left\| X_i^{(p)} \right\| < R < \infty$$

au vu de la proposition précédente, on a, quitte a prendre $R > \max(\|X_1\|, \dots, \|X_n\|)$,

$$\limsup_{p\to\infty} \chi(X_1^{(p)}, \dots, X_n^{(p)}) = \limsup_{p\to\infty} \chi_R(X_1^{(p)}, \dots, X_n^{(p)}) \leqslant \chi_R(X_1 \dots, X_n) = \chi(X_1, \dots, X_n)$$

□

Références

[And10] Alice Guionnet Ofer Zeitouni & Greg W. Anderson. *An Introduction to Random Matrices*. Cambridge University Press, 2010.

[Arv76] William Arveson. *An Invitation to C*-Algebras*. Springer-Verlag New York, 1976.

[Aus17] Tim Austin. Entropy and ergodic theory. 2017. `https://www.math.ucla.edu/~tim/entropycourse.html`.

[Bia98] Philippe Biane. Free probability for probabilists. 1998. `https://arxiv.org/abs/math/9809193v1`.

[Con90] John B. Conway. *A Course in Functional Analysis*. Springer-Verlag New York, 1990.

[Die60] Jean Dieudonné. *Foundations of Modern Analysis*. Academic Press, 1960.

[Nel73] Edward Nelson. Probability theory and euclidean field theory. In *Constructive Quantum Field Theory*, pages 94–124. Springer-Verlag, 1973.

[Pet13] Jesse Peterson. Notes on von neumann algebras. 2013. `https://math.vanderbilt.edu/peters10/teaching/spring2013/vonNeumannAlgebras.pdf`.

[Pop10] Claire Anantharaman & Sorin Popa. *An introduction to* Π_1 *factors*. 2010.

[Pro76] Claudio Procesi. The invariant theory of n×n matrices. *Advances in Mathematics*, 19 :306–381, 1976.

[Spe06] Alexandru Nica & Roland Speicher. *Lectures on the Combinatorics of Free Probability Theory*. Cambridge University Press, 2006.

[Sun87] Vaikalathur Shankar Sunder. *An Invitation To von Neumann Algebras*. Springer-Verlag New York, 1987.

[Voi91] Dan Voiculescu. Limit laws for random matrices and free products. *Inventiones mathematicae*, 104 :201–220, 1991.

[Voi94] Dan Voiculescu. The analogues of entropy and fischer information in free probability theory ii. *Inventiones mathematicae*, 118 :411–440, 1994.

[Zei98] Amir Dembo & Ofer Zeitouni. *Large Deviations Techniques and Applications*. Springer-Verlag New York, 1998.