Entropie Quantique et Théorèmes de Loewner

RAPPORT D'EA MAT581

Thomas Louis
X2020
Tuteur : Mathieu Lewin

Table des matières

I Introduction et motivations

L'outil central de la physique statistique est **l'entropie**. On la définit de la manière suivante : supposons qu'on ne connait pas un système discret de manière deterministe. On connait seulement la probabilité que le système soit dans un certain état, donnée par une mesure de probabilité ρ. Alors :

$$S_c = -\sum_{i=1}^{n} \rho(x_i) \log(\rho(x_i))$$

Ou l'indice "c" indique qu'on parle de l'entropie **classique**, et $\{x_1, ..., x_n\}$ désigne l'ensemble des états possibles.

Dans le cadre de la mécanique quantique, on considerera plutôt un "mélange statistique" d'états $|\psi_i\rangle$. Cela signifie qu'on sait que le système est dans l'état $|\psi_i\rangle$ avec probabilité p_i. On représente cette situation par une matrice densité ρ définie par

$$\rho = \sum p_i \, |\psi_i\rangle \langle\psi_i|$$

qui prend en compte à la fois l'aspect probabiliste quantique (dans les $|\psi_i\rangle$), mais aussi la description probabiliste classique (à travers les p_i).

Remarquons que ρ est une matrice hermitienne vérifiant $\mathrm{tr}(\rho) = 1$. De plus,

$$\langle\psi| \, \rho \, |\psi\rangle = \sum_{i=1}^{n} p_i |\langle\psi|\psi_i\rangle|^2 \geq 0$$

Ce qui montre que ρ est définie positive.

Cette notion nous permet de généraliser la notion d'entropie au cas quantique : on note

$$S(\rho) = -\operatorname{Tr}(\rho \log \rho) = -\sum_{i=1}^{n} \lambda_i \log(\lambda_i)$$

l'entropie quantique, équivalent quantique de S_c. L'idée est que beaucoup des propriétés vraies pour l'entropie classique restent vraies dans le cas quantique, mais les preuves deviennent plus élaborées. On détaille dans ce rapport la preuve d'une de ces propriétés : la **concavité**.

On étudiera également le cas de **l'entropie quantique relative**, qui mesure la dissimilarité entre deux matrices densité A et B :

$$S(A|B) = \mathrm{Tr}[A(\log A - \log B)]$$

et on montrera la **convexité jointe** de cette entropie relative, de laquelle on déduira la **sous-additivité forte**.

Ces preuves constituent l'occasion de s'intéresser aux notions d'opérateur monotonie / convexité, qui permettent d'étendre les notions de fonctions croissantes et convexes sur les matrices hermitiennes. On prouvera en guise de conclusion, un résultat important, le **théorème de Loewner**, qui caractérise l'ensemble de ces fonctions.

II Entropie quantique d'une variable

II.1 Matrices Hermitiennes

Définition 1. *Soit $H_n \subset \mathcal{M}_n(\mathbb{C})$ l'ensemble des matrices hermitiennes, c'est à dire vérifiant $A^* = A$.*

*On dit que $A \in H_n$ est **positive** si $\forall v \in \mathbb{C}^n$,*

$$\langle v, Av \rangle \geq 0$$

*et qu'elle est **définie positive** si cette inégalité est stricte*

On notera H_n^+ l'ensemble des matrices **positives** et H_n^{++} l'ensemble des matrices **définies positives**.

On rappelle dans la suite sans démonstration les propriétés classiques des matrices positives, qui sont des conséquences du théorème spectral.

Théorème II.1. *Soit $A \in H_n$ Les propositions suivantes sont équivalentes*

1. *$A \in H_n^+$ (resp. H_n^{++})*
2. *Toutes les valeurs propres de A sont positives (resp. strictement positive)*
3. *Il existe $N \in \mathcal{M}_n(\mathbb{C})$ (resp. $GL_n(\mathbb{C})$) telle que $A = N^*N$*

Proposition II.1. *Soit $A \in H_n^+$ (resp. H_n^{++}). Alors il existe une unique $A^{\frac{1}{2}} \in H_n^+$ (resp. H_n^{++}) telle que $\left(A^{\frac{1}{2}}\right)^2 = A$*

Proposition II.2. *Soit $A, B \in H_n^+$. Alors $ABA \in H_n^+$.*

Preuve. Soit $\{\lambda_1, ..., \lambda_n\}$ les valeurs propres de A et $\{u_1, ..., u_n\}$ une base orthogonale de vecteurs propres de A. Alors :

$$\langle u_i, ABAu_i \rangle = \langle Au_i, BAu_i \rangle = \lambda^2 \langle u_i, Bu_i \rangle \geq 0$$

$\square$

L'ensemble des matrices qu'on considère, comme vu dans l'introduction est

$$S_n = \{\rho \in H_n^{++}, \, \mathrm{Tr}(\rho) = 1\}$$

II.2 Fonctions opérateur concaves / opérateurs monotones

Dans l'introduction, on a utilisé la notation $\log \rho$ sans vraiment expliquer à quoi cela correspondait. En fait, dans le cadre des matrices hermitiennes, on dispose d'une manière naturelle de définir des fonctions de matrices à partir de fonctions réelles.

Définition 2. *Soit $A \in H_n$ et $f : \mathbb{R} \to \mathbb{R}$. Soit $\{\lambda_1, ..., \lambda_n\}$ les valeurs propres de A et $\{u_1, ..., u_n\}$ une base orthogonale de vecteurs propres de A. Soit P_i la projection orthogonale sur Vect(u_i). On note alors :*

$$f(A) = \sum_{i=1}^{n} f(\lambda_i) P_i$$

Exemple II.2.1. Soit $f : x \mapsto \sum_{k=1}^{r} c_k x^k$ une fonction polynomiale. Alors, soit $v = \sum_{i=1}^{n} v_i u_i \in \mathbb{C}^n$:

$$f(A)v = \sum_{i=1}^{n} f(\lambda_i) P_i v = \sum_{k=1}^{r} c_k \sum_{i=1}^{n} \lambda_i^k v_i u_i = \sum_{k=1}^{r} c_k A^k v$$

Donc la définition choisie est cohérente.

On peut alors généraliser les notions connnues sur les fonctions réelles (monotonie, convexité) aux fonctions matricielles.

Définition 3. *Soit $A, B \in H_n$. On définit un ordre partiel sur H_n par :*

$$A \geq B \Leftrightarrow A - B \text{ est positive}$$

Définition 4. *Soit $f : I \to \mathbb{R}$. On dit que*

1. *f est **opérateur n-croissante** (ou n-croissante) si $\forall A \leq B \in H_n$ tels que $\sigma(A), \sigma(B) \subset I$*

   $$A \leq B \Rightarrow f(A) \leq f(B)$$

   *On dit également que f est **opérateur croissante** si elle est n-croissante pour tout $n \geq 1$.*

2. *f est **opérateur n-convexe** (ou n-convexe) si, $\forall A \leq B \in H_n$ tels que*

$\sigma(A), \sigma(B) \subset I$

$$A \leq B \Rightarrow \forall t \in [0,1],\ f(tA + (1-t)B) \leq tf(A) + (1-t)f(B)$$

*On dit également que f est **opérateur convexe** si elle est n-convexe pour tout $n \geq 1$.*

3. *f est **(n-)opérateur concave** si $-f$ est (n-)opérateur convexe.*

Remarque 1. Comme en analyse réelle, si f est continue, il est suffisant de prouver que

$$f\left(\frac{A+B}{2}\right) \leq \frac{f(A)+f(B)}{2}$$

pour montrer que f est opérateur convexe. C'est ce qu'on fera dans la suite.

La notion d'opérateur monotonie / convexité est en fait beaucoup plus restrictive qu'elle ne parait. Les fonctions croissantes / convexes ne sont ainsi pas toujours opérateurs monontones / opérateurs convexes.

Exemple II.2.2. (tiré de [Bha96]). $f : t \mapsto t^2$ n'est pas opérateur-monotone sur H_n^+. En effet, prenons

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ et } B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

. Alors $B - A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \geq 0$. Mais

$$B^2 - A^2 = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} - \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}$$

qui n'est pas une matrice positive.

On commence par prouver une version plus forte de la convexité opérateur qui sera utile dans la suite. Les preuves proviennent [Car10].

On admettra le résultat suivant :

Théorème II.2. *(Inégalité de Sherman-Davis) Soit f opérateur convexe et $P \in \mathcal{M}_n(\mathbb{C})$ une projection orthogonale. Alors pour tout $A \in H_n$*

$$Pf(PAP)P \leq Pf(A)P$$

On peut alors montrer la propriété voulue :

Théorème II.3. *(Inégalité de Jensen opérateur) Soit f opérateur convexe. Alors pour tout $X_1, ..., X_K \in M_n(\mathbb{C})$ et tels que $\sum_{k=1}^{K} X_k X_k^* = 1$, et pour tout $A_1, ..., A_K \in H_n^+$ alors*

$$f\left(\sum_{k=1}^{K} X_k A_k X_k^*\right) \leq \sum_{k=1}^{K} X_k f(A_k) X_k^* \tag{1}$$

Preuve. Soit U une matrice unitaire définie par blocs :

$$U = \left[\begin{array}{c|c|c|c} V_1 & U_{1,1} & \dots & U_{1,n-1} \\ \hline \vdots & & \vdots & \vdots \\ \hline \vdots & & \vdots & \vdots \\ \hline V_K & U_{K,1} & \dots & U_{K,n-1} \end{array}\right]$$

où les blocs $U_{i,j}$ sont choisies de manière à ce que U soit bien unitaire. Par exemple, comme d'après 1, les V_i forment une famille orthonormée de $\mathbb{C}^{kn}$, on peut la compléter en une base orthonormale.

Soit également la matrice définie par blocs :

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & A_K \end{bmatrix}$$

P, matrice $Kn \times Kn$ (comme A) définie par blocs :

$$P = \begin{bmatrix} I_n & 0 & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix}$$

Alors on voit que U^*AU a $\sum_{k=1}^{K} X_k A_k X_k^*$ comme bloc en haut à gauche, et

$$f(UAU^*) = U^* f(A) U$$

a $\sum_{k=1}^{K} X_k f(A_k) X_k^*$ comme bloc en haut en gauche. Or par $II.2$:

$$Pf(PU^*ABP)P \leq Pf(U^*BU)P$$

$\square$

II.3 Operateur concavité de l'entropie

Dans la suite, on prouve que l'entropie quantique est opérateur concave. La preuve est tirée de [Car10].

Lemme II.4. *La fonction $t \mapsto t^{-1}$ est opérateur convexe sur H_n^{++}*

Preuve. Soit $A, B \in H_n^{++}$. On note $C = A^{\frac{1}{2}} B A^{\frac{1}{2}}$ Alors :

$$\frac{A^{-1} + B^{-1}}{2} - \left(\frac{A+B}{2}\right)^{-1} = A^{-\frac{1}{2}} \left(\frac{I + C^{-1}}{2} - \left(\frac{I+C}{2}\right)^{-1}\right) A^{-\frac{1}{2}}$$

On s'est ramené au cas plus simple de deux matrices qui commutent. Comme $C^{-1} \in H_n$, d'après le théorème spectral, on peut trouver une base orthogonale $\{u_1, ..., u_n\}$ de vecteurs propres de C^{-1}. Alors

$$\left\langle u_i, \left(\frac{I + C^{-1}}{2} - \left(\frac{I+C}{2}\right)^{-1}\right) u_i \right\rangle = \frac{1 + \lambda_i^{-1}}{2} - \left(\frac{1 + \lambda_i}{2}\right)^{-1}$$

Par l'inégalité des moyennes, le terme de droite est positif. Donc :

$$\frac{I + C^{-1}}{2} - \left(\frac{I+C}{2}\right)^{-1} \geq 0$$

On conclut avec la proposition II.2. $\square$

Remarque 2. On a utilisé dans la preuve "l'inégalité des moyennes". Il s'agit de dire que si $a_1, ..., a_n \geq 0$ et $n \geq 2$:

$$\frac{n}{a_1^{-1} + ... + a_n^{-1}} \leq (a_1 ... a_n)^{1/n} \leq \frac{1}{n}(a_1 + ... + a_n)$$

La partie de droite de l'inégalité est simplement l'inégalité arithmetico-géométrique.

La partie de gauche s'obtient en remarquant :

$$\frac{n}{a_1^{-1} + ... + a_n^{-1}} = \frac{a_1...a_n}{\frac{1}{n}\sum_{i=1}^n \prod_{k\neq i} a_k}$$

Le dénominateur correspond a la moyenne des $b_i = \prod_{k\neq i} a_k$ qu'on notera $m(b)$. Soit $g(b)$ la moyenne géométrique des b_i. Alors

$$g(b) = \left((a_1...a_n)^{n-1}\right)^{1/n} = (a_1...a_n)^{1-\frac{1}{n}}$$

Donc

$$\frac{n}{a_1^{-1} + ... + a_n^{-1}} = (a_1...a_n)^{1/n} \cdot \frac{g(b)}{h(b)} \leq (a_1...a_n)^{1/n}$$

Théorème II.5. *Pour tout $p \in [0,1]$ la fonction $t \mapsto t^p$ est opérateur concave. Pour tout $p \in [1,2]$ la fonction $t \mapsto t^p$ est opérateur convexe.*

Pour prouver le théorème ci-dessus, on aura besoin du lemme suivant. Il s'agit simplement d'un calcul d'intégral qu'on résout avec la formule des résidus.

Lemme II.6. *Soit $A \in H_n^{++}$.*

1. *Soit $0 < p < 1$. Alors :*

$$A^p = \frac{\pi}{\sin(\pi p)} \int_0^{+\infty} t^p \left(\frac{1}{t} - \frac{1}{t+A}\right) \mathrm{d}t$$

2. *Soit $1 < p < 2$. Alors :*

$$A^p = \frac{\pi}{\sin(\pi(p-1))} \int_0^{+\infty} t^p \left(\frac{A}{t} + \frac{t}{t+A} - 1\right) \mathrm{d}t$$

Preuve. (du théorème II.5)
Soit $0 < p < 1$. D'après le lemme II.4. $A \mapsto -(A+t)^{-1}$ est opérateur concave. La formule intégrale montre donc que $t \mapsto t^p$ est limite simple de sommes de fonctions opérateurs concaves. Comme la concavité passe à la limite, $t \mapsto t^p$ est opérateur concave.

Soit maintenant $1 < p < 2$. De même, $A \mapsto \frac{A}{t} + \frac{t}{t+A}$ est une somme de fonction opérateur convexe, est donc opérateur convexe. Par le même argument, $t \mapsto t^p$ est opérateur convexe.

Le cas $p = 1$ correspond à l'identité, qui est évidemment à la fois opérateur concave et opérateur convexe. $\square$

Théorème II.7. *$t \mapsto -t \log t$ est opérateur concave*

Preuve. On a :

$$A \log A = \lim_{p \to 1} \frac{A^p - A}{p - 1}$$

Mais $A \mapsto \frac{A^p - A}{p-1}$ est opérateur convexe pour tout $p \in [0, 2] \backslash \{1\}$ (cela est dû au changement de signe de $p - 1$ pour $p < 1$). Donc $A \mapsto A \log A$ est opérateur convexe, c'est à dire $A \mapsto -A \log A$ est opérateur concave. $\square$

Corollaire 1. *L'entropie quantique*

$$A \mapsto -\operatorname{Tr}(A \log A)$$

est concave.

Remarque 3. On pourra retrouver ce résultat de manière directe avec le théorème de Loewner, prouvé en section IV.

III Entropie relative

Ce qui suit est inspiré du cours ([Lew15]).

III.1 Premières propriétés et définition

Dans le cas d'une fonction réelle h à deux variables, la définition de la fonction d'opérateur correspondante est plus compliquée, car dans le cas général, deux matrices hermitiennes A et B ne sont pas co-diagonalisables, donc il n'est pas possible de définir $h(A, B)$ sur des espaces propres communs à A et B. On peut, cependant, le faire pour h sous une forme particulière, par exemple

$$h(x, y) = \sum_{k=1}^{r} c_k f_k(x) g_k(y)$$

C'est le cas de l'entropie relative :

Définition 5. *Soit $A, B \in S_n$. On note :*

$$S(A|B) = \mathrm{Tr}[A(\log A - \log B)]$$

L'entropie relative s'écrit également sous la forme suivante, qui est souvent utile.

Proposition III.1. *Soit $f : t \mapsto t \log t$ et $A, B \in S_n$. Alors*

$$S(A|B) = \mathrm{Tr}\left[f(A) - f(B) - f'(B)(A - B)\right]$$

Preuve. On a $f'(t) = \log t + 1$. Donc :

$$\begin{aligned} f(A) - f(B) - f'(B)(A - B) &= A \log A - B \log B - (\log B + 1)(A - B) \\ &= A(\log A - \log B) + A - B \end{aligned}$$

En prenant la trace, et comme $\mathrm{Tr}(A) = \mathrm{Tr}(B) = 1$:

$$\mathrm{Tr}\left[f(A) - f(B) - f'(B)(A - B)\right] = \mathrm{Tr}\left[A(\log A - \log B)\right]$$

□

On commence par un développement qui montre qu'on peut voir l'entropie relative comme une **distance**.

Proposition III.2. *Soit h définie sur $I \times J$ une fonction de la forme :*

$$h(x,y) = \sum_{k=1}^{r} c_k f_k(x) g_k(y)$$

telle que $h(x,y) \geq 0$ pour tout $x, y \in I \times J$. Alors pour tout $A, B \in H_n^+$,

$$\mathrm{Tr}[h(A,B)] \geq 0$$

Preuve. Soit $\{\lambda_1, ..., \lambda_n\}$ les valeurs propres de A et $\{u_1, ..., u_n\}$ une base orthogonale de vecteurs propres de A, respectivement $\{\mu_1, ..., \mu_n\}$ et $\{v_1, ..., v_n\}$ pour B. Soit P_i, Q_i respectivement les projecteurs sur $Vect(u_i), Vect(v_i)$ Alors :

$$\begin{aligned} h(A,B) &= \sum_{k=1}^{r} c_k f_k(A) g_k(B) \\ &= \sum_{k=1}^{r} c_k \sum_{i=1}^{n} f_k(\lambda_i) P_i \sum_{j=1}^{n} g_k(\mu_j) Q_j \end{aligned}$$

Donc :

$$\mathrm{Tr}(h(A,B)) = \sum_{i,j=1}^{n} \mathrm{Tr}(P_i Q_j) \sum_{k=1}^{r} c_k f_k(\lambda_i) g_k(\mu_j)$$

Or il est facile de voir que $\mathrm{Tr}(P_i Q_j) = |\langle u_i, v_j \rangle|^2$. Donc

$$\mathrm{Tr}(h(A,B)) = \sum_{i,j=1}^{n} |\langle u_i, v_j \rangle|^2 \underbrace{\sum_{k=1}^{r} c_k f_k(\lambda_i) g_k(\mu_j)}_{\geq 0} \geq 0$$

□

Corollaire 2. *Soit $A, B \in S_n$. Alors*

$$S(A|B) \geq \mathrm{Tr}\big((A-B)^2\big)$$

En particulier $A = B \Leftrightarrow S(A|B) = 0$

Preuve. Soit $g(x,y) = x(\log x - \log y) - (x-y)^2$ définie pour tout $x, y \in]0,1]^2$. Alors

$$\frac{\partial g}{\partial x} = \log x - \log y + 1 - 2(x-y)$$
$$\frac{\partial g}{\partial y} = -\frac{x}{y} + 2(x-y)$$

Mais $-\frac{x}{y} + 2(x-y) = 0 \Leftrightarrow x - 2xy + 2y^2 + = 0 \Leftrightarrow x + x^2 + (y-x)^2 = 0$, ce qui est impossible. Donc g n'a pas de point critique dans $]0,1[^2$. De plus, les valeurs (limites) de g au bord de ce rectangle sont toutes positives. Donc g est toujours positive. Donc, d'après la proposition III.2, on a l'inégalité recherchée. □

III.2 Convexité jointe de l'entropie relative

On veut maintenant montrer que l'entropie relative est conjointement convexe, c'est à dire que pour toutes matrices $A_1, A_2, B_1, B_2 \in S_n$, et pour tout $\lambda \in [0,1]$:

$$S(\lambda A_1 + (1-\lambda)A_2 | \lambda B_1 + (1-\lambda)B_2) \leq \lambda S(A_1|B_1) + (1-\lambda)S(A_2|B_2)$$

On va en fait prouver un résultat plus général, qui traite notamment des traces partielles :

Définition 6. *Soit $V = V_1 \otimes V_2$ un produit tensoriel et $A \in \mathcal{L}(V)$. Soit $e_1, ..., e_p$ une base de V_1 et $f_1, ..., f_q$ une base de V_2. Alors $(e_i \otimes f_j)_{i,j}$ est une base de V et A s'écrit*

$$A = [A_{ijkl}] = [\langle e_i f_j, A e_k f_l \rangle] \textit{ pour } 1 \leq i,k \leq p \textit{ et } 1 \leq j,l \leq q$$

Alors la ***trace partielle*** *de A sur V_2 est définie par*

$$\mathrm{Tr}_2(A) = \left[\sum_{j=1}^{q} A_{ijkj}\right] \textit{ pour } 1 \leq i,k \leq p$$

On peut construire cet opérateur sans avoir recours au représentation matricielles en utilisant la caractérisation :

$$\mathrm{Tr}((B \otimes I)A) = \mathrm{Tr}(B\,\mathrm{Tr}_2(A))$$

On retrouve la première construction en prenant les B_i projecteurs sur $Vect(e_i)\times V_2$

Théorème III.1. *Soit $(A,B)\in S_n$. Alors*

1. $(A,B)\mapsto S(A|B)$ *est convexe*
2. *Pour toute matrices $(X_1,...,X_n)\mathcal{M}_n\in(\mathbb{C})$ telles que*

$$\sum_{k=1}^{K} X_k^* X_k = 1$$

Alors

$$\begin{aligned} S\left(\sum_{k=1}^{K} X_k A X_k^* \middle| \sum_{k=1}^{K} X_k B X_k^*\right) &\leq \sum_{k=1}^{K} S\left(X_k A X_k^* | X_k B X_k^*\right) \\ &\leq S(A|B) \end{aligned}$$

3. *Si $V=V_1\otimes V_2$, alors*

$$S(\mathrm{Tr}_2(A)|\,\mathrm{Tr}_2(B))\leq S(A|B)$$

Preuve. En fait, ces trois propriétés sont équivalentes. On commence par :

<u>Preuve de 1. $\Rightarrow$ 2.</u>.

Supposons S convexe. Remarquons que S est homogène de degré 1, c'est à dire

$$S(tA|tB)=tS(A|B)\quad\forall t>0$$

Donc on écrit :

$$\begin{aligned} S\left(\sum_{k=1}^{K} X_k A X_k^* \middle| \sum_{k=1}^{K} X_k B X_k^*\right) &= KS\left(\frac{1}{K}\sum_{k=1}^{K} X_k A X_k^* \middle| \frac{1}{K}\sum_{k=1}^{K} X_k B X_k^*\right) \\ &\leq \sum_{k=1}^{K} S\left(X_k A X_k^* | X_k B X_k^*\right) \end{aligned}$$

par convexité de S. Pour prouver la deuxième inégalité, on peut voir que :

$$\sum_{k=1}^{K} S\left(X_k A X_k^* | X_k B X_k^*\right) = S\left(\begin{bmatrix} X_1 A X_1^* & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & X_K A X_K^* \end{bmatrix} \middle| \begin{bmatrix} X_1 B X_1^* & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & X_K B X_K^* \end{bmatrix}\right)$$

Soit v l 'application définie sur $\mathbb{C}^n$ par

$$v(x) = \begin{bmatrix} X_1 x \\ \vdots \\ X_K x \end{bmatrix}$$

Alors v est une isométrie : en effet,

$$\begin{aligned} \langle v(x), v(y) \rangle &= \sum_{k=1}^{K} x^* X_k^* X_k y \\ &= x^* \left(\sum_{k=1}^{K} X_k^* X_k \right) y = x^* y = \langle x, y \rangle \end{aligned}$$

Donc

$$S(A|B) = S(vAv^* | vBv^*)$$

Or vAv^* et vBv^* ont les mêmes diagonales (respectives) que

$$\begin{bmatrix} X_1 A X_1^* & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & X_K A X_K^* \end{bmatrix} \text{ et } \begin{bmatrix} X_1 B X_1^* & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & X_K B X_K^* \end{bmatrix}$$

Il suffit donc de montrer qu'en gardant uniquement les coefficients diagonaux des matrices, l'entropie jointe décroit. Cela revient à montrer le résultat dans le cas où les X_k sont des projecteurs P_k tels que $\sum_{k=1}^{K} P_k = 1$. Par récurrence, il suffit de montrer que si

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \text{ et } B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

sont des matrices définies par blocs, alors :

$$S\left(\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \middle| \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix}\right) \leq S(A|B)$$

Or,

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} = \frac{A}{2} + \frac{1}{2}\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} A \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$$

Et de même pour B. On conclut par convexité et par la proposition.

<u>Preuve de 2. $\Rightarrow$ 3.</u>

On a, d'après le lemme de Schur :

$$\int_{SU(d)} UXU^* \mathrm{d}U = \frac{\mathrm{Tr}(X)}{d}$$

En particulier :

$$\int_{SU(V_2)} (I \otimes U)A(I \otimes U^*)\mathrm{d}U = \frac{\mathrm{Tr}_2(A) \otimes I}{d_2}$$

Or, $\int_{SU(V_2)}(I \otimes U)A(I \otimes U^*)$ est limite d'opérateur de la forme

$$\frac{1}{K}\sum_{k}^{K}(I \otimes U_k)A(I \otimes U_k^*)$$

et de même pour B. Or, d'après 2. :

$$S\left(\frac{1}{K}\sum_{k}^{K}(I \otimes U_k)A(I \otimes U_k^*) \middle| \frac{1}{K}\sum_{k}^{K}(I \otimes U_k)A(I \otimes U_k^*)\right) \leq S(A|B)$$

Donc, par passage à la limite :

$$S\left(\frac{\mathrm{Tr}_2(A) \otimes I}{d_2} \middle| \frac{\mathrm{Tr}_2(B) \otimes I}{d_2}\right) \leq S(A|B)$$

Or :

$$\begin{aligned} S\left(\operatorname{Tr}_2(A)\otimes I|\operatorname{Tr}_2(B)\otimes I\right) &= \operatorname{Tr}[\operatorname{Tr}_2(A)\otimes I((\log \operatorname{Tr}_2(A)\otimes I) - (\log \operatorname{Tr}_2(B)\otimes I))] \\ &= \operatorname{Tr}[(\operatorname{Tr}_2(A)(\log \operatorname{Tr}_2 A - \log \operatorname{Tr}_2 B))\otimes I] \\ &= d_2 S(\operatorname{Tr}_2(A)|\operatorname{Tr}_2(B)) \end{aligned}$$

Ce qui donne le résultat par homogénéité de S.

Preuve de 1.
Définissons les actions L (action à gauche) et R (action à droite) sur S_n par :

$$\begin{aligned} L(M) =& AM \\ R(M) =& MB \end{aligned}$$

Alors on a :

$$S(A|B) = \operatorname{Tr}(A(\log A - \log B)) = \langle I, L(\log L - \log R)(I)\rangle$$

Sous cette formulation, on retrouve le même problème de départ, mais avec les opérateurs L et R qui commutent. Cependant, on a plus la trace, donc il faut montrer la convexité opérateur. On veut donc montrer que si L_1, L_2, R_1, R_2 sont des opérateurs tels que $[L_i, R_j] = 0$, $t \in [0,1]$, alors, en notant $L = tL_1 + (1-t)L_2$ et $R = tR_1 + (1-t)R_2$

$$L(\log L - \log R) \leq (1-t)L_1(\log L_1 - \log R_1) + tL_2(\log L_2 - \log R_2)$$

Par les relations de commutations et avec $f(t) = t\log t$)

$$\begin{aligned} & L(\log L - \log R) \\ =& R^{1/2} f(R^{-1/2} L R^{-1/2}) R^{1/2} \\ =& R^{1/2} f(tR^{-1/2} L_1 R^{-1/2} + (1-t) R^{-1/2} L_2 R^{-1/2}) R^{1/2}) \end{aligned}$$

Soit $X = \sqrt{t} R^{-1/2} R_1^{1/2}$ et $Y = \sqrt{1-t} R^{-1/2} R_2^{1/2}$. Alors on a bien $X^* X + Y^* Y = 1$ et on conclut grace à la proposition $II.3$. $\square$

Remarque 4. L'application

$$A \mapsto \sum_{k=1}^{K} X_k A X_k^*$$

est un exemple de "**canal quantique**" (quantum channel en anglais). Remarquons qu'une propriété de cette application est qu'elle préserve la trace :

$$\begin{aligned}
\mathrm{Tr}\left[\sum_{k=1}^{n} X_k A X_k^*\right] &= \sum_{k=1}^{n} \mathrm{Tr}[X_k A X_k^*] \\
&= \sum_{k=1}^{n} \mathrm{Tr}[X_k^* X_k A] \\
&= \mathrm{Tr}\left[\sum_{k=1}^{n} (X_k^* X_k) A\right] \\
&= \mathrm{Tr}[A]
\end{aligned}$$

Soit ρ_{123} une matrice densité définie sur $V = V_1 \otimes V_2 \otimes V_3$. On note $\rho_{12} = \mathrm{Tr}_3(\rho_{123})$, et de même pour les autres indices.

Théorème III.2. *Sous-additivité forte de l'entropie*

$$S(\rho_{123}) + S(\rho_2) \leq S(\rho_{12}) + S(\rho_{23})$$

Preuve. Appliquons le théorème $III.1$ à $E_1 = V_1 \otimes V_2$, $E_2 = V_3$ et aux opérateurs ρ_{123} et $\rho_1 \otimes \rho_{23}$:

$$S(\rho_{12} \,|\, \rho_1 \otimes \rho_3) \leq S(\rho_{123} \,|\, \rho_1 \otimes \rho_{23})$$

Or :

$$\begin{aligned}
&S(\rho_{123} \,|\, \rho_1 \otimes \rho_{23}) - S(\rho_{12} \,|\, \rho_1 \otimes \rho_3) \\
&= \mathrm{Tr}\left[\rho_{123}(\log \rho_{123} - \log \rho_1 \otimes \rho_{23}) - \rho_{12}(\log \rho_{12} - \log \rho_1 \otimes \rho_3)\right] \\
&= S(\rho_{12}) + S(\rho_{23}) - S(\rho_2) - S(\rho_{123}) \geq 0
\end{aligned}$$

□

IV Théoreme de Loewner :

IV.1 Enoncé et premières implications

On démontre maintenant un résultat très général, qui donne une caractérisation des fonctions opérateur monotones. La preuve qu'on présente ici est tirée de [Sim19].

Théorème IV.1. *Soit $f \in \mathcal{C}^1([-1,1],\mathbb{R})$. Alors les conditions suivantes sont équivalentes :*

1. *f est opérateur monotone*
2. *il existe une mesure de probabilité μ sur $[-1,1]$ telle que*

$$\forall t \in [-1,1],\ \ f(t) = f(0) + f'(0) \int_{-1}^{1} \frac{t}{1-\lambda t} \mathrm{d}\mu(\lambda)$$

3. *f est la restriction à $]-1,1[$ d'une fonction analytique sur $(\mathbb{C}\backslash\mathbb{R}) \cup]-1,1[$ telle que*

$$\operatorname{Im} z > 0 \rightarrow \operatorname{Im} f(z) > 0$$

Remarque 5. Il n'est pas compliqué de voir que (ii) $\Rightarrow$ (iii) en écrivant :

$$\operatorname{Im}\left(\frac{z}{1+\lambda z}\right) = \operatorname{Im}\left(\frac{z(1+\lambda\overline{z})}{|1+\lambda z|^2}\right) = \frac{\operatorname{Im} z}{|1+\lambda z|^2}$$

L'implication réciproque, (iii) $\Rightarrow$ (ii), se fait à l'aide d'un théorème de représentation (théorème de représentation de Herglotz), qu'on ne détaillera pas ici.

Remarque 6. L'implication (ii) $\Rightarrow$ (i) provient du fait que pour tout $\lambda \in [-1,1]$ la fonction

$$x \mapsto \frac{x}{1+\lambda x}$$

est opérateur monotone.

Dans la suite, on s'occupera de prouver l'implication difficile du théorème, c'est à dire le sens (i) $\Leftrightarrow$ (ii) et (iii). Plus précisément, on prouvera (i) $\Rightarrow$ (ii), et (iii) se déduira de la remarque ci-dessus.

IV.2 Matrices de Loewner

On commence par une condition nécessaire et suffisante de monotonie qui nous sera utile dans la suite.

Définition 7. *Soit* $f \in \mathcal{C}^1([-1,1])$*,* $-1 < x_1, ..., x_k < 1$*. La* ***matrice de Loewner*** *associée à* f *et à ces points est*

$$[L_k(x_1, ..., x_k)]_{ij} = \begin{cases} \frac{f(x_i)-f(x_j)}{x_i-x_j} & \text{pour } i \neq j \\ f(x_i) & \text{pour } i = j \end{cases}$$

L'objectif de cette section est de prouver le résultat suivant :

Théorème IV.2. *Soit* $f \in \mathcal{C}^1([-1,1])$*. Alors* f *est n-monotone si et seulement si pour tout* $-1 < x_1, ..., x_n < 1$

$$L_n(x_1...x_n; f) \geq 0$$

Pour prouver ce théorème, on aura besoin des définitions suivantes :

Définition 8. *Soit* $A, B \in \mathcal{M}_n(\mathbb{C})$*. On définit*

$$[A \odot B]_{ij} = A_{ij}B_{ij}$$

Lemme IV.3. *Si* $A, B \geq 0$*, alors* $(A \odot B) \geq 0$

Preuve. Soit φ un vecteur, et on définit

$$\left[Q^{(\varphi)}\right] = \overline{\varphi_i}\varphi_j$$

Or, A, B sont des sommes de Q^φ où les φ sont orthogonaux. Donc il suffit de prouver le résultat pour $A = Q^\varphi$ et $B = Q^\psi$. Mais $Q^\varphi \odot Q^\psi = Q^{\varphi \odot \psi} \geq 0$ □

Théorème IV.4. *Formule de Daleckii-Krein Soit* $-1 < x_1, ..., x_m < 1$ *et* $f \in \mathcal{C}([-1,1])$*. Soit*

$$A = \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_m \end{pmatrix}$$

et $C \in H_n$. Alors

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} f(A+\lambda C)\bigg|_{\lambda=0} = L_m(x_1, ..., x_m; f) \odot C$$

Preuve. Par la continuité des valeurs propres (voir IV.7), pour λ assez petit, il existe une base orthonormale $(\varphi_k(\lambda))_k$ et des valeurs propres $(x_k(\lambda))_k$ telles que

$$\begin{aligned}
&(A+\lambda C)\varphi_k(\lambda) = x_k(\lambda)\varphi_k(\lambda) \\
&\varphi_k(0) = e_k \\
&x_k(0) = x_k
\end{aligned}$$

et les x_k sont continus en λ. On écrit

$$\langle \varphi_k(\lambda), f(A+\lambda C)\varphi_l(0)\rangle = \langle f(A+\lambda C)\varphi_k(\lambda), \varphi_l(0)\rangle$$

Donc

$$\langle \varphi_k(\lambda), [f(A+\lambda C) - f(A)]\varphi_l(0)\rangle = [f(x_k(\lambda)) - f(x_l(0))]\langle \varphi_k(\lambda), \varphi_l(0)\rangle$$

En particulier, avec f la fonction identité :

$$\langle \varphi_k(\lambda), \lambda C\varphi_l(0)\rangle = [x_k(\lambda) - x_l(0)]\langle \varphi_k(\lambda), \varphi_l(0)\rangle$$

En combinant les deux équations, on obtient :

$$\left\langle \varphi_k(\lambda), \frac{f(A+\lambda C) - f(A)}{\lambda}\varphi_l(0)\right\rangle = \frac{f(x_k(\lambda)) - f(x_l(0))}{x_k(\lambda) - x_l(0)}\langle \varphi_k(\lambda), C\varphi_l(0)\rangle$$

$\lambda \to 0$ donne :

$$\left[\frac{\mathrm{d}}{\mathrm{d}\lambda} f(A+\lambda C)\bigg|_{\lambda=0}\right]_{kl} = L_m(x_1, ..., x_m; f)_{kl} C_{kl}$$

Ce qui est exactement le résultat demandé. □

On a maintenant tous les outils pour prouver le théorème IV.2

Preuve. (de IV.2) Supposons <u>f m-monotone</u> et soit $-1 < x_1 < ... < x_m < 1$. On reprend A comme dans IV.4. Soit $C \geq 0$. Alors $A + \lambda C \geq 0$ pour tout

$\lambda \geq 0$. Donc $f(A+\lambda C) \geq f(A)$.

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} f(A+\lambda C) \geq 0.$$

Donc, d'après la formule de Daleckii-Krein (IV.4)

$$L_m(x_1, ..., x_m; f) \odot C \geq 0$$

et ce pour toute matrice $C \geq 0$. On prend en particulier $C = \varphi\varphi^T$, et on a alors :

$$\mathrm{Tr}[L_m(x_1, ..., x_m; f) \odot C] = \langle \varphi, L_m(x_1, ..., x_m; f)\varphi \rangle \geq 0$$

Donc $L_m(x_1, ..., x_m; f) \geq 0$

Inversement, on suppose $\underline{L_m(x_1, ..., x_m; f) \geq 0}$. D'après la formule de Daleckii-Krein (IV.4), et le lemme IV.3, on a pour tout $C \geq 0$, et en supposant A diagonale à valeurs propres distinctes

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda}(A+\lambda C)\bigg|_{\lambda=0} \geq 0$$

Et comme la positivité ne dépend pas de la base choisie, ce fait est vrai pour tout A à valeurs propres distinctes. Remarquons ensuite que, pour $\mu \in [0,1]$ fixé,

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda}(A+\lambda C)\bigg|_{\lambda=\mu} = \frac{\mathrm{d}f}{\mathrm{d}\lambda}(A+(\mu+\lambda)C)\bigg|_{\lambda=0}$$

Donc, si $A+\mu C$ a des valeurs propres distinctes, on a

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda}(A+\lambda C)\bigg|_{\lambda=\mu} \geq 0$$

Prenons donc maintenant $X \leq Y$. On veut montrer que $f(X) \leq f(Y)$. Remarquons qu'en supposant que $X + \mu(Y-X)$ a des valeurs propres distinctes pour tout $\mu \in [0,1]$, alors on peut appliquer la formule précédente avec $A = X$ et $B = Y - X$ pour affirmer que

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda}\left(A+\lambda(B-A)\right)\bigg|_{\lambda=\mu} \geq 0 \qquad (2)$$

pour tout $\mu \in [0,1]$ ce qui impose que $f(A) \leq f(B)$.

Si maintenant on suppose seulement que X a des valeurs propres distinctes, en notant $Z(\mu) = X + \mu(Y - X)$, alors les valeurs propres $(x_j(\mu))_j$ de $Z(\mu)$ sont analytiques sur $[0,1]$ (voir corollaire 3). Ainsi, les fonctions $x_j - x_i$ sont analytiques, et par le principe des zéros isolés, admettent un nombre fini de zéros. Finalement, il existe un nombre fini de réels $0 < \mu_1 < ... < \mu_l \leq 1$ tels que $Z(\mu)$ a des valeurs propres dégénérées en $\mu_1, ..., \mu_l$. Or, par 2, pour ε assez petit :

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda}\left(A + \lambda(B - A)\right)\Bigg|_{\lambda=\mu_j-\varepsilon} \geq 0$$

et comme f est C_1, en passant à la limite $\varepsilon \to 0$, on revient au cas précedent.

Démntrons maintenant le cas général, c'est à dire si A admet des valeurs propres dégénérées. Alors il est facile de voir que A est limite de $(A_k)_k$ où les A_k ont des valeurs propres distinctes (il suffit de perturber les valeurs propres de A). Ainsi, on a :

$$f(A_k) \leq f(A_k + (B - A))$$

pour tout $k \geq 0$, en par passage à la limite on obtient le résultat. □

IV.3 Preuve de Loewner

On présente maintenant la peuve du théorème $IV.1$.

Soit f opérateur monotone. On remarque que quitte à ajouter une fonction bien choisie, i.e définir :

$$f_{\varepsilon,\delta}(x) = f(x(1-\delta)) + \varepsilon g(x)$$

on peut supposer que les matrices de Leowner de f sont toutes **strictement** positives. On fera donc cette supposition dans toute la suite.

Définition 9. *Soit* $-1 \leq x_1 < x_2 < ... < x_n \leq 1$. *On définit un produit scalaire par :*

$$\langle \alpha, \beta \rangle_L = \langle \alpha, L(x_1, ..., x_n; f)\beta \rangle$$

L'idée générale est qu'on arrivera à représenter f à travers ce produit scalaire.

Proposition IV.1. *Il existe un unique vecteur $\varphi \in \mathbb{C}^n$ tel que, pour tout $1 \leq j \leq n$:*

$$\langle \delta_j, \varphi \rangle_L = f(x_j) \tag{3}$$

et un unique $A \in H_n$ tel que

$$\langle \delta_j, A\delta_k \rangle_L = \begin{cases} x_j f'(x_j) + f(x_j) & si\ j = k \\ \frac{x_k f(x_j) - x_k f(x_k)}{x_j - x_k} & si\ j \neq k \end{cases} \tag{4}$$

De plus, on a :

$$(A - x_j)\delta_j = \varphi \tag{5}$$

Preuve. Comme $L > 0$, 3 définit bien de manière unique un vecteur φ et 4 les éléments de matrices de A. D'autre part, on a :

$$\langle \delta_j, A\delta_k \rangle_L = x_k \langle \delta_j, \delta_k \rangle_L + \langle \delta_j, \varphi \rangle_L$$

Donc :

$$\langle \delta, j(A - x_k)\delta_k \rangle = \langle \delta_j, \varphi \rangle_L$$

Donc, par identification, on a 5 □

Dans la suite, on notera P_j la projection sur $Vect(A - x_j)$. Alors $A - x_j$: $\mathrm{Im}(P_j) \to \mathrm{Im}(P_j)$ est une bijection, car on a

$$\mathbb{C}^n = \ker(A - x_j) \oplus \mathrm{Im}(A - x_j)$$

On peut donc définir sans problème

$$(A - x_j)^{-1} P_j$$

sur $\mathbb{C}^n$.

Lemme IV.5. *On a :*

$$(1 - P_j)\varphi = 0 \quad et \quad P_j\delta_j = (A - x_j)^{-1} P_j \varphi$$

et

$$\|(A - x_j)P_j\varphi\|_L^2 \leq \|\delta_j\|_L^2 \tag{6}$$

Preuve. Comme $\varphi \in \text{Im}(P_j)$ par 5, on a

$$(1 - P_j)\varphi = 0$$

De même, 5 donne directement

$$P_j\delta_j = (A - x_j)^{-1}P_j\varphi$$

En appliquant $\|\cdot\|_L$ à cette dernière égalité, en utilisant qu'une projection est de norme inferieure à 1, on a

$$\|(A - x_j)P_j\varphi\|_L^2 \leq \|\delta_j\|_L^2$$

□

Théorème IV.6. *Il existe une mesure μ sur $\mathbb{R}$ dont le support est inclus dans n points, telle que $\mu(\{x_j\}) = 0$ et*

$$f(x_j) = \int_{\mathbb{R}} (x - x_j)^{-1}\,\mathrm{d}\mu(x) \tag{7}$$

$$f'(x_j) \geq \int_{\mathbb{R}} (x - x_j)^{-2}\,\mathrm{d}\mu(x) \tag{8}$$

Preuve. Soit maintenant $\{\psi_\alpha\}$ une base orthonormée de vecteurs propres de A, de valeurs propres respectives λ_α. Soit μ la mesure discrète définie par :

$$\mu(\{\lambda_\alpha\}) = |\langle\psi_\alpha, \varphi\rangle|^2$$

Alors, si $x_j = \lambda_\beta$ est une valeur propre, on a $x_j \in Ker(A - x_j) = \text{Im}(A - x_j)^\perp = \text{Im}(P_j)$. Or $\varphi \in \text{Im}(P_j)$ donc :

$$\mu(\{x_j\}) = \langle\psi_\beta, \varphi\rangle_L = 0$$

et dans le cas où x_j n'est pas valeur propre, on a évidemment toujours $\mu(\{x_j)\} = 0$.

Finalement, on a :

$$\begin{aligned} f(x_j) &= \langle \delta_j, \varphi \rangle_L \\ &= \langle P_j \delta_j, \varphi \rangle_L \\ &= (A - x_j)^{-1} P_j \varphi, \varphi \rangle_L \end{aligned}$$

Et comme $f(x_j) \in \mathbb{R}$, on a :

$$f(x_j) = \langle \varphi, (A - x_j)^{-1} P_j \varphi \rangle_L$$

Or, en notant $\varphi = \sum_\alpha \langle \psi_\alpha, \varphi \rangle \psi_\alpha$. Alors par ce qui précède :

$$f(x_j) = \sum_\alpha \langle \psi_\alpha, \varphi \rangle (\lambda_\alpha - x_j)^{-1} \langle \varphi, \psi_\alpha \rangle = \sum_\alpha |\langle \psi_\alpha, \varphi \rangle|^2 (\lambda_\alpha - x_j)^{-1}$$

où on convient que $|\langle \psi_\alpha, \varphi \rangle|^2 (\lambda_\alpha - x_j)^{-1} = 0$ si $\langle \psi_\alpha, \varphi \rangle = 0$. Alors l'expression ci dessus est exactement :

$$f(x_j) = \int_{\mathbb{R}} (x - x_j)^{-1} \, \mathrm{d}\mu(x)$$

D'autre part, on a :

$$\|\delta_j\|_L^2 = \langle \delta_j, L(x_1, ..., x_n; f) \delta_j \rangle = [L(x_1, ..., x_n; f)]_j j = f'(x_j)$$

Donc, 6 devient

$$\|(A - x_j) P_j \varphi\|_L^2 \leq f'(x_j)$$

et, de manière similaire,

$$\|(A - x_j) P_j \varphi\|_L^2 = \int_{\mathbb{R}} (x - x_j)^{-2} \, \mathrm{d}\mu(x)$$

ce qui donne :

$$f'(x_j) \geq \int_{\mathbb{R}} (x - x_j)^{-2} \, \mathrm{d}\mu(x)$$

□

La suite de l'argument consiste en l'application de ce qui précède pour $x_1, ..., x_n$ des dyadiques.

Définition 10. *Soit* $x_k = \frac{k}{2^m}$ *pour* $k \in \{-2^m, ..., 2^m\}$. *Soit alors* μ_m *la mesure* *donnée par le théorème IV.6 associée aux* $(x_k)_k$

Proposition IV.2. *On a*

$$\mu_m\left(\left[\frac{j}{2^m}-\frac{1}{2^{m+1}},\frac{j}{2^m}+\frac{1}{2m+1}\right]\right)\leq 2^{-2m-2}f'\left(\frac{j}{2^m}\right)$$

donc, en sommant sur j ;

$$\mu_m([-1,1])\leq 2^{-m-1}(1+2^{-m-1})\|f'\|_\infty$$

Preuve. On a, en notant $I_j=\left[\frac{j}{2^m}-\frac{1}{2^{m+1}},\frac{j}{2^m}+\frac{1}{2m+1}\right]$

$$\begin{aligned}
f'(\frac{j}{2^m}) &\geq \int_{\mathbb{R}}(x-j2^{-m})^{-2}\,\mathrm{d}\mu(x) && (9)\\
&\geq \int_{I_j}(x-j2^{-m})^{-2}\,\mathrm{d}\mu(x) && (10)\\
&\geq \left(2^{m+1}\right)^2\int_{I_j}\mathrm{d}\mu(x) && (11)\\
&= 2^{2m+2}\mu_m(I_j) && (12)
\end{aligned}$$

On a ce qu'il faut. □

Définition 11. *On définit* $\tilde{\mu}_m=\mu_m|_{\mathbb{R}\setminus[-1,1]}$, *et*

$$f_m(x)=\int_{\mathbb{R}\setminus[-1,1]}\frac{\mathrm{d}\tilde{\mu}_m}{y-x}$$

Proposition IV.3. *1. Pour tout* $j\in\{-2^m,...,2^m\}$ *:*

$$\left|f_m\left(\frac{j}{2^m}\right)-f\left(\frac{j}{2^m}\right)\right|\leq c2^{-(m+1)/2}$$

2. $f'_m(0)\leq f'(0)$

Preuve. 1. On a, par inégalité triangulaire :

$$\left|f_m\left(\frac{j}{2^m}\right)-f\left(\frac{j}{2^m}\right)\right|\leq\int_{-1}^{1}\frac{\mathrm{d}\mu_m(y)}{y-j2^{-m}}$$

puis, par Cauchy-Schwarz :

$$\int_{-1}^{1} \frac{\mathrm{d}\mu_m(y)}{|y - j2^{-m}|} \leq \left(\int_{-1}^{1} \frac{\mathrm{d}\mu_m(y)}{|y - j2^{-m}|^2} \right)^{\frac{1}{2}} \mu_m([-1,1])^{\frac{1}{2}}$$

D'après la propriété précédente et 8

$$\begin{aligned} \left(\int_{-1}^{1} \frac{\mathrm{d}\mu_m(y)}{|y - j2^{-m}|^2} \right)^{\frac{1}{2}} \mu_m([-1,1])^{\frac{1}{2}} &\leq \|f'\|_\infty^{\frac{1}{2}} \left(2^{-m-1}(1 + 2^{-m-1}) \|f'\|_\infty \right)^{\frac{1}{2}} \\ &\leq c2^{-(m+1)/2} \end{aligned}$$

2. On dérive sous le signe intégrale :

$$f'_m(x) = \int_{\mathbb{R}\backslash[-1,1]} \frac{\mathrm{d}\tilde{\mu}_m}{(y-x)^2}$$

Donc, en utilisant 8

$$f'_m(0) = \int_{\mathbb{R}\backslash[-1,1]} \frac{\mathrm{d}\tilde{\mu}_m}{(y-x)^2} \leq \int_{\mathbb{R}} \frac{\mathrm{d}\mu_m}{(y-x)^2} \leq f(0)$$

□

En fait, l'idée de la construction des (f_m) était qu'on obtient maintenant des fonctions qui vérifient (iii) dans le théorème IV.1. En effet, en définissant :

$$f_m(z) = \int_{\mathbb{R}\backslash[-1,1]} \frac{\mathrm{d}\tilde{\mu}_m}{y - z} \quad \forall z \in (\mathbb{C}\backslash\mathbb{R}) \cup [-1,1]$$

on obtient une fonction analytique. De plus, si $\operatorname{Im} z > 0$, alors :

$$\operatorname{Im}\left(\frac{1}{y-z}\right) = \operatorname{Im}\left(\frac{y - \overline{z}}{|y-z|^2}\right) = \frac{\operatorname{Im} z}{|y-z|^2} > 0$$

Ainsi, comme on a admis (ii) $\Leftrightarrow$ (iii), on a l'existence d'une mesure ν_m sur $[-1,1]$ telle que pour tout $x \in]-1,1[$

$$f_m(x) = f_m(0) + \int_{-1}^{1} \frac{x}{1 + \lambda x} \mathrm{d}\nu_m(\lambda)$$

Il suffit maintenant de montrer qu'on a bien une convergence de chaque terme quand $m \to +\infty$. D'une part, par pour tout x dyadique, d'après IV.3,

$$\lim_{m \to +\infty} f_m(x) = f(x)$$

D'autre part :

$$f'_m(x) = \int_{-1}^{1} \frac{(1+\lambda x) - \lambda x}{(1+\lambda x)^2} \mathrm{d}\nu_m(\lambda) = \int_{-1}^{1} \frac{1}{(1+\lambda x)^2} \mathrm{d}\nu_m(\lambda)$$

Donc d'après IV.3

$$f'_m(0) = \nu_m([-1,1]) \leq f'(0)$$

Donc les mesures sont uniformément bornées, et de plus, par ce qui précède, la limite

$$\lim_{m \to +\infty} \int_{-1}^{1} \frac{x}{1+\lambda x} \mathrm{d}\nu_m(\lambda)$$

Or, on peut voir à l'aide du théorème de Weierstrass que la famille

$$\left(\lambda \mapsto \frac{x}{1+\lambda x}\right)_{x \text{ dyadique}}$$

est totale dans $\mathcal{C}([-1,1])$. Ainsi, ν_m converge faiblement vers une mesure ν telle qu'on ait, pour tout x dyadique :

$$f(x) = f(0) + \int_{-1}^{1} \frac{x}{1+\lambda x} \mathrm{d}\nu(\lambda)$$

Or, comme les dyadiques sont denses, cette formule est vraie pour tout $x \in [-1,1]$.

Annexe

Théorème IV.7. *Continuité des valeurs propres*
Soit A une matrice de $\mathcal{M}_n(\mathbb{C})$ diagonalisable et $\lambda \in \mathbb{C}$ une valeur propre simple de A. Alors il existe un voisinage ouvert U de $\mathcal{M}_n(\mathbb{C})$ et une fonction $\mu : U \to \mathbb{C}$ analytique telle que

$$\mu(A) = \lambda$$
$$\forall B \in U,\ \mu(B) \in \sigma(B)$$

Corollaire 3. *Soit $(A(\lambda))_{0\leq\lambda\leq 1}$ une famille de H_n telle que pour tout $\lambda \in [0,1]$, $A(\lambda)$ a des valeurs propres distinctes et $\lambda \mapsto A(\lambda)$ est analytique. Alors il existe des fonctions analytiques réelles $(\lambda \mapsto x_j(\lambda))$ et des fonctions analytiques vectorielles $(\lambda \mapsto \varphi_j(\lambda))$ sur $[0,1]$ telles que :*

$$A(\lambda)\varphi_j(\lambda) = x_j(\lambda)$$

et pour tout λ, $(\varphi_j(\lambda))_j$ est une base orthonormée de $\mathbb{C}^n$.

Références

[Bha96] R. Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics. Springer New York, 1996. ISBN : 9780387948461. URL : `https://books.google.fr/books?id=F4hRy1F1M6QC`.

[Car10] Eric Carlen. "Trace inequalities and quantum entropy : An introductory course". In : 529 (jan. 2010). DOI : `10.1090/conm/529/10428`.

[Lew15] Mathieu Lewin. "Mesure de Gibbs non linéaires". Cours à l'IHES. 2015. URL : `https://www.youtube.com/watch?v=O53f6QZXsKo&t=3607s`.

[Sim19] B. Simon. *Loewner's Theorem on Monotone Matrix Functions*. Grundlehren der mathematischen Wissenschaften. Springer International Publishing, 2019. ISBN : 9783030224226. URL : `https://books.google.fr/books?id=hSmsDwAAQBAJ`.