Mesure, information et entropie

Le but de ce cours est de faire un bilan de l'information gagnée ou perdue lors d'une mesure. Auparavant, nous rappelons quelques résultats relatifs à l'entropie statistique d'un état quantique

1 Entropie statistique associée à un état quantique

a. Information manquante associée à une loi de probabilité

– Loi de probabilité $\{p_m\}$, avec $m = 1, 2 \dots N$, associée à N évènements aléatoires

$$\sum_{m=1}^{N} p_m = 1 \tag{10.1}$$

– L'évènement attendu étant aléatoire, il nous manque une certaine information. On peut essayer de chiffrer cette information manquante. En postulant certaines propriétés générales (additivité, continuité ...), Shannon a montré que l'information manquante associée à la loi de probabilité $\{p_m\}$ était donnée par

$$S(\{p_m\}) = -k \sum_m p_m \operatorname{Log} p_m \tag{10.2}$$

où k est une constante arbitraire. Voir références 1, 2, 3 (§ 3.1) et 4 (complément I.G).

b. Entropie statistique de von Neumann

– En mécanique quantique, toute notre connaissance sur l'état d'un système est rassemblée dans l'opérateur densité D (hermitique, semipositif, de trace 1). Diagonalisons D sous la forme

$$D = \sum_m p_m |\varphi_m\rangle\langle\varphi_m| \tag{10.3}$$

où

$$\langle\varphi_m|\varphi_p\rangle = \delta_{mp} \tag{10.4.a}$$

$$p_m \text{ réel}, \geqslant 0 \qquad \sum_m p_m = 1 \tag{10.4.b}$$

plusieurs p_m pouvant être égaux (valeurs propres dégénérées).

– Von Neumann a introduit l'entropie statistique $S(D)$ associée à D (référence 5)

$$\begin{aligned} S(D) &= -k \sum_m p_m \operatorname{Log} p_m \\ &= -k \operatorname{Tr} D \operatorname{Log} D \end{aligned} \tag{10.5}$$

L'entropie statistique (10.5) coïncide avec l'information manquante de Shannon (10.2) associée à la loi de probabilité $\{p_m\}$.

On peut également considérer que $S(D)$ représente le désordre existant dans l'état D, ou encore l'information manquante sur le système dans l'état D. On pourrait essayer d'acquérir cette information manquante au moyen de mesures. Il ne faut pas oublier cependant qu'une mesure perturbe en général l'état du système. Par exemple, si l'on mesure A, de valeurs propres a_α, l'état du système après une mesure non lue de A passe de D à $D' = \sum_\alpha \Pi_\alpha D \Pi_\alpha$ où Π_α est le projecteur sur le sous espace propre a_α. $S(D)$ caractérise donc l'information manquante qu'on pourrait acquérir en mesurant une observable A qui ne perturbe pas D, c'est à dire pour laquelle $D' = D$, c'est à dire encore une observable diagonale dans la base $\{|\varphi_m\rangle\}$ (voir référence 3, § 3.2)

C - Quelques propriétés importantes de l'entropie statistique.

<u>Lemme</u>

$$\begin{cases} S(D) = -k \,\mathrm{Tr}\, D \,\mathrm{Log}\, D \leq -k \,\mathrm{Tr}\, D \,\mathrm{Log}\, D' \\ \text{égalité réalisée si et seulement si } D' = D \end{cases} \tag{10.6}$$

- Démonstration

$$D = \sum_m p_m |\varphi_m\rangle\langle\varphi_m| \qquad D' = \sum_n p'_n |\varphi'_n\rangle\langle\varphi'_n| \tag{10.7}$$

$$k\,\mathrm{Tr}(-D\,\mathrm{Log}\,D + D\,\mathrm{Log}\,D') = -k\sum_m p_m \,\mathrm{Log}\, p_m + k\,\mathrm{Tr}\left[\left(\sum_m p_m |\varphi_m\rangle\langle\varphi_m|\right)\left(\sum_n \mathrm{Log}\, p'_n |\varphi'_n\rangle\langle\varphi'_n|\right)\right]$$

$$= -k\sum_m p_m \,\mathrm{Log}\, p_m + \sum_m\sum_n p_m |\langle\varphi_m|\varphi'_n\rangle|^2 \,\mathrm{Log}\, p'_n$$

$$= k\sum_m\sum_n p_m |\langle\varphi_m|\varphi'_n\rangle|^2 \,\mathrm{Log}\, \frac{p'_n}{p_m} \tag{10.8}$$

Or,
$$\mathrm{Log}\, y \leq y - 1 \tag{10.9}$$
l'égalité étant réalisée si et seulement si $y = 1$. On en déduit, en posant $y = p'_n/p_m$

$$p_m \,\mathrm{Log}\,\frac{p'_n}{p_m} \leq p_m\left(\frac{p'_n}{p_m} - 1\right) = p'_n - p_m \tag{10.10}$$

ce qui donne

$$k\,\mathrm{Tr}(-D\,\mathrm{Log}\,D + D\,\mathrm{Log}\,D') \leq \sum_m\sum_n (p'_n - p_m)|\langle\varphi_m|\varphi'_n\rangle|^2 = \sum_n p'_n - \sum_m p_m = 0 \tag{10.11}$$

et démontre la 1ère ligne de (10.6).

L'égalité dans (10.11) est réalisée si, pour tout couple m, n, on a, ou bien $\langle\varphi'_n|\varphi_m\rangle = 0$, ou bien $p'_n = p_m$, c'est à dire si

$$\langle\varphi'_n|\varphi_m\rangle(p'_n - p_m) = 0 \qquad \forall\ n, m \tag{10.12a}$$

qu'on peut réécrire sous la forme

$$\langle\varphi'_n|D' - D|\varphi_m\rangle = 0 \qquad \forall\ n, m \tag{10.12b}$$

Comme les $\{|\varphi_m\rangle\}$ et les $\{|\varphi'_n\rangle\}$ forment des bases, cette égalité n'est réalisée que si $D' = D$, ce qui démontre la 2me ligne de (10.6).

<u>Maximum et minimum de $S(D)$</u>

- Supposons l'espace des états de dimension finie W, et prenons pour D' l'opérateur de trace 1, proportionnel à l'opérateur unité : $D' = \frac{1}{W}\sum_m |\varphi_m\rangle\langle\varphi_m|$. On a alors, d'après (10.6)

$$S(D) \leq -k\,\mathrm{Log}\, W \,\mathrm{Tr}\, D = k\,\mathrm{Log}\, W \tag{10.13}$$

On a donc
$$S_{Max} = k\,\mathrm{Log}\, W \tag{10.14}$$

Pour faire le raccord avec la thermodynamique, nous prendrons désormais $k = k_B$ (constante de Boltzmann) et Log désignera le logarithme népérien.

- Le minimum de $S(D)$ est realisé pour un état pur (tous les p_m nuls sauf un qui vaut 1)
$$S_{min} = 0 \tag{10.15}$$

Additivité

– Le système étudié $\mathcal{S}$ est formé par la réunion de 2 systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$

Si $\quad D(1,2) = D(1) \otimes D(2) \qquad S[D(1,2)] = S[D(1)] + S[D(2)]$ (10.16)

Pour démontrer (10.16), il suffit d'utiliser une base produit tensoriel d'une base $\{|\varphi_m(1)\rangle\}$ qui diagonalise $D(1)$ par une base $\{|\chi_n(2)\rangle\}$ qui diagonalise $D(2)$

$$S[D(1,2)] = -k_B \sum_m \sum_n P_m^{(1)} P_n^{(2)} \operatorname{Log}(P_m^{(1)} P_n^{(2)}) = -k_B \sum_m \sum_n P_m^{(1)} P_n^{(2)} [\operatorname{Log} P_m^{(1)} + \operatorname{Log} P_n^{(2)}]$$

$$= -k_B \sum_m P_m^{(1)} \operatorname{Log} P_m^{(1)} - k_B \sum_n P_n^{(2)} \operatorname{Log} P_n^{(2)} = S[D(1)] + S[D(2)] \qquad (10.17)$$

- Le résultat (10.17) exprime que le désordre dans le tout est la somme des désordres dans les parties, si les parties sont statistiquement indépendantes.

Corrélations

– En général, $D(1,2)$ ne se factorise pas. A partir de $D(1,2)$, on peut calculer

$$D(1) = \operatorname{Tr}_2 D(1,2) \qquad D(2) = \operatorname{Tr}_1 D(1,2) \qquad (10.18)$$

On a alors

$$\begin{cases} S[D(1)] + S[D(2)] \geq S[D(1,2)] \\ \text{égalité réalisée si et seulement si } D(1,2) = D(1) \otimes D(2) \end{cases} \qquad (10.19)$$

– Démonstration

$$S[D(1)] = -k_B \operatorname{Tr} D(1) \operatorname{Log} D(1) = -k_B \operatorname{Tr} D(1,2) \operatorname{Log} D(1) \otimes \mathbb{1}(2) \qquad (10.20.a)$$

$$S[D(2)] = -k_B \operatorname{Tr} D(2) \operatorname{Log} D(2) = -k_B \operatorname{Tr} D(1,2) \operatorname{Log} \mathbb{1}(1) \otimes D(2) \qquad (10.20b)$$

Or, en utilisant la base factorisée qui diagonalise $D(1)$ et $D(2)$, on démontre que

$$\operatorname{Log}[D(1) \otimes \mathbb{1}(2)] + \operatorname{Log}[\mathbb{1}(1) \otimes D(2)] = \operatorname{Log}[D(1) \otimes D(2)] \qquad (10.21)$$

En ajoutant (10.20.a) et (10.20.b) et en utilisant (10.21), on obtient alors, compte tenu de (10.6)

$$\begin{aligned} S[D(1)] + S[D(2)] &= -k_B \operatorname{Tr} D(1,2) \operatorname{Log}[D(1) \otimes D(2)] \\ &\geq -k_B \operatorname{Tr} D(1,2) \operatorname{Log} D(1,2) = S[D(1,2)] \end{aligned} \qquad (10.22)$$

ce qui démontre (10.19).

– L'inégalité (10.19) exprime physiquement que $D(1)$ et $D(2)$ contiennent ensemble moins d'information que $D(1,2)$, qui contient en plus l'information relative aux corrélations entre 1 et 2.

Concavité

– Soient D_a et D_b 2 opérateurs densités possibles d'un même système et soit λ un réel tel que $0 \leq \lambda \leq 1$. On a, pour tout λ, et quels que soient D_a et D_b

$$\begin{cases} S[\lambda D_a + (1-\lambda) D_b] \geq \lambda S[D_a) + (1-\lambda) S(D_b) \\ \text{égalité réalisée si et seulement si } D_a = D_b \end{cases} \qquad (10.23)$$

– Démonstration. Posons

$$D' = \lambda D_a + (1-\lambda) D_b \qquad (10.24)$$

et appliquons (10.6) avec $D = D_a$ puis $D = D_b$

$$-k_B\lambda \operatorname{Tr} D_a \operatorname{Log} D' \geqslant \lambda\, S(D_a) \tag{10.25.a}$$

$$-k_B(1-\lambda) \operatorname{Tr} D_b \operatorname{Log} D' \geqslant (1-\lambda)\, S(D_b) \tag{10.25.b}$$

Il suffit alors d'ajouter (10.25.a) et 10.25.b) et d'utiliser (10.24) pour obtenir (10.23)

- Physiquement, (10.23) exprime que la réunion en un seul mélange statistique de 2 états d'un même système augmente le désordre.

- Généralisation

$$\begin{cases} S\left(\sum_j \mu_j D_j\right) \geqslant \sum_j \mu_j\, S(D_j) \\ \mu_j \text{ réel}, \quad 0 \leqslant \mu_j \leqslant 1, \quad \sum_j \mu_j = 1 \end{cases} \tag{10.26}$$

Invariance de S(D) dans une transformation unitaire de D

- Si $D(t) = U(t)\, D(0)\, U^+(t)$, $S[D(t)] = S[D(0)]$ (10.27)

- Démonstration

$$S[D(t)] = -k_B \operatorname{Tr}\left(D(t) \operatorname{Log} D(t)\right) = -k_B \operatorname{Tr}\left(U(t)\, D(0) \operatorname{Log} D(0)\, U^+(t)\right)$$
$$= -k_B \operatorname{Tr}\left(\underbrace{U^+(t)\, U(t)}_{=1}\, D(0) \operatorname{Log} D(0)\right) = S[D(0)] \tag{10.28}$$

- L'entropie d'un système ne change donc pas au cours d'une évolution unitaire de ce système.

- Que se passe-t-il si l'hamiltonien qui régit l'évolution unitaire de D est mal connu ?

Il faut alors moyenner sur plusieurs évolutions possibles de probabilités respectives p_j $\left(\sum_j p_j = 1\right)$

$$D(t) = \sum_j p_j\, U_j(t)\, D(0)\, U_j^+(t) = \sum_j p_j\, D_j(t) \tag{10.29}$$

Comme $D_j(t) = U_j(t)\, D(0)\, U_j^+(t)$, on a d'après (10.27)

$$S(D_j(t)) = S(D(0)) \tag{10.30}$$

L'inégalité (10.26) (concavité) entraîne alors, compte tenu de (10.29) et (10.30)

$$S(D(t)) = S\left(\sum_j p_j\, D_j(t)\right) \geqslant \sum_j p_j\, S(D(0)) \tag{10.31}$$

c'est à dire encore, comme $\sum_j p_j = 1$

$$S(D(t)) \geqslant S(D(0)) \tag{10.32}$$

Si un système isolé évolue sous l'effet d'un hamiltonien mal connu, l'entropie de ce système augmente au cours du temps

2) Diverses entropies statistiques pouvant être introduites à propos d'une mesure idéale (voir référence 6)

a) Opérateurs densité du système avant et après la mesure.

- Mesure de
$$A = \sum_\alpha a_\alpha \Pi_\alpha \tag{10.33}$$
(où Π_α est le projecteur sur le sous-espace propre a_α) sur un système

$\mathcal{S}$ dans l'état décrit par D. Probabilité de trouver a_α :

$$\mathcal{P}(a_\alpha) = \mathrm{Tr}\, \Pi_\alpha D \qquad (10.34)$$

- Etat avant la mesure : D. On peut écrire

$$D = \sum_\alpha \sum_\beta \Pi_\alpha D \Pi_\beta \qquad (10.35)$$

- Le système $\mathcal{S}$ et l'appareil de mesure $\mathcal{M}$ interagissent, puis se séparent. On ne s'intéresse plus à $\mathcal{M}$.

Etat réduit de $\mathcal{S}$ après une mesure non lue de A

$$D \longrightarrow D' = \sum_\alpha \Pi_\alpha D \Pi_\alpha \qquad (10.36)$$

Les blocs non diagonaux de D, $\Pi_\alpha D \Pi_\beta$ avec $\alpha \neq \beta$, sont annulés

- On observe le résultat obtenu. C'est a_α. Etat après la mesure

$$D \to D_\alpha = \frac{\Pi_\alpha D \Pi_\alpha}{\mathrm{Tr}\, \Pi_\alpha D} = \frac{\Pi_\alpha D \Pi_\alpha}{\mathcal{P}(a_\alpha)} \qquad (10.37)$$

Grâce à (10.37), on peut réécrire (10.36) sous la forme

$$D' = \sum_\alpha \mathcal{P}(a_\alpha)\, D_\alpha \qquad (10.38)$$

b) <u>Entropies du système après la mesure</u>

- Après une mesure non lue de A

$$S(D') = S\left(\sum_\alpha \Pi_\alpha D \Pi_\alpha\right) \qquad (10.39)$$

- Après une mesure ayant donné le résultat a_α

$$S(D_\alpha) = S\left(\frac{\Pi_\alpha D \Pi_\alpha}{\mathcal{P}(a_\alpha)}\right) \qquad (10.40)$$

- Entropie moyenne après une mesure lue : Moyenne des $S(D_\alpha)$ pondérées par les probabilités $\mathcal{P}(a_\alpha)$ de trouver a_α

$$\langle S(D_\alpha) \rangle = \sum_\alpha \mathcal{P}(a_\alpha)\, S(D_\alpha) \qquad (10.41)$$

c) <u>Entropie statistique associée à la distribution de probabilité des résultats de mesure.</u>

On considère la loi de probabilité $\{\mathcal{P}(a_\alpha)\}$ associée aux divers résultats possibles a_α d'une mesure de A et on introduit l'entropie de Shannon S_A associée à une telle loi (voir (10.2))

$$S_A = -k_B \sum_\alpha \mathcal{P}(a_\alpha) \,\mathrm{Log}\, \mathcal{P}(a_\alpha) \qquad (10.42)$$

d) <u>Relation entre $S(D')$, $\langle S(D_\alpha)\rangle$ et S_A. Additivité de l'information</u>

- A partir de (10.38), on obtient, dans une base qui diagonalise les D_α

$$\begin{aligned} S(D') &= -k_B \mathrm{Tr} \sum_\alpha \mathcal{P}(a_\alpha) D_\alpha \,\mathrm{Log}\big(\mathcal{P}(a_\alpha) D_\alpha\big) \\ &= -k_B \sum_\alpha \mathcal{P}(a_\alpha) \,\mathrm{Log}\, \mathcal{P}(a_\alpha) \underbrace{\mathrm{Tr}\, D_\alpha}_{=1} - k_B \sum_\alpha \mathcal{P}(a_\alpha)\, \mathrm{Tr}\, D_\alpha \,\mathrm{Log}\, D_\alpha \end{aligned} \qquad (10.43)$$

c'est à dire encore, compte tenu de (10.42) et (10.41)

$$S(D') = S_A + \langle S(D_\alpha) \rangle \qquad (10.44)$$

- Le résultat (10.44) montre que l'information totale manquante $S(D')$ sur l'état D' après une mesure non lue de A est la somme de 2 termes :
  (i) l'information manquante S_A sur le résultat a_α de la mesure
  (ii) l'information manquante moyenne qui subsiste après la lecture d'un résultat de mesure a_α et qui est liée au fait que, si a_α est dégénérée, D_α n'est pas en général un état pur.

- Après la mesure, le fait de trier les systèmes suivant les résultats obtenus a_α diminue l'entropie de $S(D')$ à $\langle S(D_\alpha)\rangle$. L'information S_A gagnée par l'observateur lui permet de diminuer l'entropie des systèmes. L'information est transformée en "negentropie".

③ Bilan d'informations au cours d'une mesure idéale

a - Perte d'information après une mesure idéale non lue

- Montrons tout d'abord que

$$S(D') \geqslant S(D) \tag{10.45}$$

- Démonstration.
  D'après (10.36), et compte tenu de l'invariance d'une trace dans une permutation circulaire

$$S(D') = -k_B \,\mathrm{Tr}\Big(\sum_\alpha \Pi_\alpha D \Pi_\alpha \,\mathrm{Log}\, D'\Big) = -k_B\,\mathrm{Tr}\Big(\sum_\alpha D\,\Pi_\alpha \,\mathrm{Log}\, D'\,\Pi_\alpha\Big) \tag{10.46}$$

Comme D' n'a pas de blocs non diagonaux

$$\sum_\alpha \Pi_\alpha \,\mathrm{Log}\, D'\, \Pi_\alpha = \mathrm{Log}\, D' \tag{10.47}$$

On transforme alors (10.46) en

$$S(D') = -k_B \,\mathrm{Tr}\, D \,\mathrm{Log}\, D' \tag{10.48}$$

L'inégalité (10.45) résulte alors simplement de (10.6).

- Le résultat (10.45) signifie physiquement que, en laissant interagir $\mathcal{S}$ et $\mathcal{M}$ puis en les séparant, on perd l'information associée aux corrélations quantiques non séparables apparues entre $\mathcal{S}$ et $\mathcal{M}$ à la suite de leur interaction.
  L'entropie de $\mathcal{S}+\mathcal{M}$ reste constante au cours du temps. C'est le fait de prendre la trace sur $\mathcal{M}$ qui augmente l'entropie de $\mathcal{S}$.

- On peut dire encore que $S(D')-S(D)$ représente la perte d'information associée à la disparition des blocs non-diagonaux de D. L'information contenue dans D sur les observables non compatibles avec A est détruite lors de l'interaction $\mathcal{S}$-$\mathcal{M}$.

b - L'état D' est un état d'entropie maximale

- Considérons tous les opérateurs densités $\tilde{D}$ conduisant aux mêmes valeurs moyennes que D pour toutes les observables C compatibles avec A

$$\mathrm{Tr}\, C\tilde{D} = \mathrm{Tr}\, C D \qquad \forall\ C \quad \text{avec } [C,A]=0 \tag{10.49}$$

– Dans la base où A est diagonal, les observables C sont représentées par des matrices diagonales par blocs

$$\Pi_\alpha C \Pi_\beta = 0 \qquad si \quad \alpha \neq \beta \tag{10.50}$$

Par contre, $\Pi_\alpha C \Pi_\alpha$ est a priori quelconque. Pour que l'égalité (10.49) soit valable quel que soit C, il faut donc que $\tilde{D}$ et D aient les même blocs diagonaux. Les opérateurs densités $\tilde{D}$ introduits plus haut sont donc tels que

$$\Pi_\alpha \tilde{D} \Pi_\alpha = \Pi_\alpha D \Pi_\alpha \qquad \forall \alpha \tag{10.51}$$

d'où l'on déduit que

$$D' = \sum_\alpha \Pi_\alpha D \Pi_\alpha = \sum_\alpha \Pi_\alpha \tilde{D} \Pi_\alpha \tag{10.52}$$

– La démonstration donnée dans le § 3a plus haut s'applique donc à $\tilde{D}$ et conduit à

$$S(D') \geqslant S(\tilde{D}) \tag{10.53}$$

L'état D' peut donc être également interprété comme l'état d'entropie maximale parmi tous les états $\tilde{D}$ conduisant aux mêmes valeurs moyennes que D pour toutes les observables C compatibles avec A.

Si l'on ne s'intéresse qu'aux observables C compatibles avec A, l'opérateur densité D' est donc l'opérateur densité le "moins biaisé" conduisant aux mêmes valeurs moyennes que D

C- Information moyenne gagnée par la lecture

– La lecture du résultat de mesure renseigne, non pas sur l'état D (état avant la mesure), mais sur l'état D' (état réduit de $\mathcal{S}$ après l'interaction $\mathcal{S}$-$\mathcal{M}$).

– On voit sur (10.44) que l'entropie moyenne après sélection des résultats, $\langle S(D_\alpha)\rangle$, est plus petite que l'entropie après la mesure non lue $S(D')$

$$S(D') - \langle S(D_\alpha)\rangle = S_A \tag{10.54}$$

L'information moyenne gagnée par la lecture des résultats est l'information S_A associée à la distribution $\{\mathcal{P}(a_\alpha)\}$ des probabilités des résultats de mesure

– Notons enfin qu'on peut montrer que (voir référence 7 et appendice de la référence 6):

$$S(D) \geqslant \langle S(D_\alpha)\rangle \tag{10.55}$$

L'information manquante moyenne après sélection de résultats a_α, $\langle S(D_\alpha)\rangle$, est donc plus petite que l'information manquante $S(D)$ avant la lecture.

Comme, d'après (10.44), on a

$$\langle S(D_\alpha)\rangle = S(D') - S_A \qquad (10.56)$$

on peut réécrire (10.55) sous la forme

$$S_A \geqslant S(D') - S(D) \qquad (10.57)$$

qui exprime que l'information S_A gagnée par la lecture des résultats de mesure est toujours plus grande que la perte d'information $S(D')-S(D)$ sur les observables ne commutant pas avec A, perte d'information qui résulte de la disparition des blocs non diagonaux de A sous l'effet de l'interaction $S-\mathcal{M}$.

Références

1. C.E. Shannon and W. Weaver, The mathematical theory of Communication (Univ. of Illinois Press, Urbana 1949)
2. L. Brillouin, La science et la théorie de l'information (Masson Paris 1959)
3. R. Balian, Du microscopique au macroscopique, cours de physique statistique de l'école polytechnique (Ellipses, Paris 1982)
4. B. Diu, C. Guthmann, D. Lederer, B. Roulet, Physique statistique (Hermann Paris 1989), Complément I.G
5. J. von Neumann, Mathematical foundations of quantum mechanics (Princeton 1955), reproduit dans J.A. Wheeler and W.H. Zurek, Quantum theory and measurement (Princeton, 1983).
6. R. Balian, Eur. J. Phys. 10, 208 (1989)
7. G. Lindblad, Comm. Math. Phys. 28, 245 (1972)