Annales de la Fondation Louis de Broglie - Vol. 3, n° 3, 1978

ÉTUDE DE CERTAINES CONSEQUENCES D'UNE INTERPRÉTATION SUBJECTIVE DE LA NOTION D'ÉTAT

par M. Nicolas HADJISAVVAS
Laboratoire de Mécanique Quantique
Faculté des Sciences de Reims
B.P. 347 51062 Reims Cédex

(manuscrit reçu le 23 Novembre 1977)

Résumé : *Un formalisme est développé, qui permet d'attribuer à un microsystème un état uniquement défini, dans des circonstances où le formalisme traditionnel de la mécanique quantique ne spécifie aucun critère pour une telle attribution. Le formalisme proposé par nous repose sur une généralisation d'un théorème bien connu de Gleason, et il aboutit à une solution ; une tentative avait été faite pour la première fois par Von Neumann, puis reprise par Jaynes. La solution mentionnée s'intègre aux perspectives de la théorie de l'information, et peut-être reliée aux essais récents de définition d'un concept de probabilité "comparative" ou "relative", à l'aide duquel il devienne possible d'élargir la classe des propriétés probabilistes descriptibles.*

INTRODUCTION

Nous proposons dans ce travail une interprétation subjective de la notion d'état en Mécanique Quantique, dans la perspective de la théorie de l'information : nous généralisons en même temps cette notion, de manière à inclure aussi des états non normalisables. Cette généralisation permettra de répondre à certains problèmes concernant l'attribution d'un état à un système physique quand certaines informations sont disponibles, problèmes que la Mécanique Quantique laisse irrésolus.

En effet, en ce qui concerne la définition d'un état en fonction de l'information disponible, la Mécanique Quantique prescrit en général une réponse. Maisil existe des circonstances où l'information -bien qu'elle ne soit pas nulle- est insuffisante pour permettre de définir un état (par exemple lorsque seule la valeur moyenne d'un certain opérateur est connue). Ce problème a été analysé d'abord par von Neumann, puis par Jaynes. La solution proposée a été la suivante : on considère l'ensemble de tous les états compatibles avec l'information donnée, et on cherche parmi ces états celui qui possède la plus grande entropie. Toutefois, cette méthode est inapplicable si l'information disponible est telle que les entropies des états correspondants n'ont pas de maximum, i. e. lorsqu'il existe une suite P_n d'états telle que l'entropie des P_n tend vers l'infini. Nous développons un formalisme capable d'aborder ce problème en le soumettant à l'exigence de donner la même réponse que le formalisme usuel lorsque celui-ci fournit une réponse. Le formalisme que nous proposons repose sur l'axiomatique quantique. Nous sommes amenés à considérer, selon une idée de von Neumann, une généralisation non normalisable de l'opérateur statistique : nous établissons à cet effet une généralisation du théorème de Gleason (5).

Le formalisme de l'Axiomatique Quantique

A chaque système physique, C. PIRON (4) associe un ensemble L, dont les éléments sont toutes les propositions expérimentalement vérifiables qui concernent le système.

On introduit une relation sur L ; pour deux propositions "a" et "b", on écrit $a \subseteq b$ si la vérité de "a" entraîne toujours la vérité de "b". On pose ensuite quelques axiomes.

I) La relation "$\subseteq$" est une relation d'ordre. Elle vérifie donc :

$a \subseteq a \; \forall a \in L$

$a \subseteq b$ et $b \subseteq a$ entraînent $a = b$

$a \subseteq b$ et $b \subseteq c$ entraînent $a \subseteq c$

II) Pour chaque famille $\{a_i\}_{i\in I}$ d'éléments de L, il existe un élément noté $\bigcap_I a_i$ tel que

$$x \subseteq a_i \ \forall_i \Leftrightarrow x \subseteq \bigcap_I a_i$$

La valeur de vérité de $\bigcap_I a_i$ se définit en posant que $\bigcap_I a_i$ est vraie si et seulement si toutes les a_i sont vraies. Bien que cette définition possède une analogie avec la définition logique, elle est en fait différente, parce que les propositions $a_i (i = 1,2 \ldots)$ ne sont pas en général vérifiables simultanément. Ceci est la source d'importantes questions physiques.

L'axiome II entraîne l'existence d'une proposition "absurde" $\emptyset = \bigcap_L a$, telle que $\emptyset \subseteq a$, $\forall a \in L$.

III) A chaque proposition "a" correspond une proposition "a'", telle que $(a')' = a$, $a \cap a' = \emptyset$, $a \subseteq b \Rightarrow b' \subseteq a'$

Physiquement, "a'" représente la proposition qui est fausse quand "a" est vraie, et inversement.

L'axiome III, combiné avec II, entraîne l'existence de l'union pour chaque famille des propositions $\{a_i\}_{i\in I}$, c'est-à-dire d'un élément $\bigcup_I a_i$, qui satisfait aux conditions :

$$\bigcup_I a_i \subseteq x \Leftrightarrow a_i \subseteq x, \forall i$$

ainsi que l'existence d'un élément "identité" $I = \bigcup_L a$

En outre, on définit deux propositions comme compatibles si l'ensemble des éléments de L qui peuvent être écrits à partir de "a", "b", en utilisant les opérations d'intersection $\cap$ et d'orthocomplémentation (') est un treillis de Boole.

IV) Si $a \subseteq b$, alors a,b sont compatibles. Enfin on dit qu'une proposition P est atomique (ou bien que P est un atome ou une proposition minimale) si :

$$\emptyset \subseteq x \subseteq P \qquad \text{entraînent que } x = \emptyset \text{ ou } x = P$$

V) a- Pour chaque $a \in L$, il existe un atome $P \subseteq a$
b- Si P est un atome, alors

$$a \subseteq x \subseteq a \cup P \Rightarrow x = a \text{ ou } x = a \cup P$$

L'ensemble L, équipé par les axiomes I à V, est nommé "système de propositions".

On dit que le système des propositions L est irréductible si $\emptyset$ et I sont les seules propositions compatibles avec toutes les propositions de L.

En utilisant ces axiomes, C. PIRON a pu prouver qu'un système de propositions irréductibles est isomorphe à l'ensemble des sous-espaces fermés d'un espace d'Hilbert. Donc, à chaque élément $a \in L$ correspond un sous-espace R_a d'un espace d'Hilbert H, d'une telle façon que

$R_{a \cap b} = R_a \cap R_b$, $R_{a'} = R_a^{\perp}$ ($R_a^{\perp}$ étant le complément orthogonal de R_a).

La notion d'état

Une définition de l'état donné par l'école de Jauch est la suivante :

Un état est une fonction p, définie sur L, dont les valeurs se trouvent dans l'intervalle [0, 1] et sur laquelle on impose un certain nombre des conditions (Jauch 1). Dans le cas où l'on ferait une mesure de type "oui-non", correspondant à la proposition $a \in L$, on obtiendrait la réponse "oui" avec une probabilité p(a).

Sans nous opposer à cette définition, nous proposons une interprétation de la notion d'état selon laquelle l'état est plus fortement lié qu'en mécanique quantique au degré de connaissances de l'observateur, concernant le système. En utilisant l'information dont il dispose, l'observateur associe à chaque proposition une probabilité. Les probabilités associées par deux observateurs possédant des informations différentes, peuvent être différentes. L'état des connaissances peut changer si l'observateur obtient une information nouvelle*. Alors, afin d'inclure dans le traite-

* Il est bien évident que cette conception utilise la notion de probabilité "subjective" voir Jaynes (2) : "The test of a good subjective probability distribution p(X) is : does it correctly represents our state of knowledge as to the value of X ?"

ment qui suit des situations où l'observateur possède une information insuffisante pour associer au système un vecteur d'état du formalisme quantique, nous modifions la définition de l'état :

Nous appelons état une fonction qui associé à chaque élément de L un nombre positif, tel que le quotient $\frac{\dot{p}(a)}{\dot{p}(b)}$ $(a,b \in L)$ représente la probabilité relative* de "a" et "b". Cela signifie que si l'on effectue le même nombre N d'expériences pour mesurer a et b sur des systèmes pour lesquels on a la même information, on s'attend à constater que le rapport des fréquences de a et de b approche $\frac{\dot{p}(a)}{\dot{p}(b)}$ lorsque N s'accroit.

Nous posons donc la définition suivante :

Définition I) Un état est une fonction $L \xrightarrow{\dot{p}} R \cup \{\infty\}$ telle que :

a) $p(\emptyset) = 0$, $\dot{p}(I) > 0$
b) Si $a \subseteq b'$, alors $p(a \cup b) = p(a) + p(b)$
c) $p(a) = p(b) = 0 \Rightarrow p(a \cup b) = 0$
d) Si $a_1 \subseteq a_2 \subseteq a_3 \ldots$, alors $p(\cup_n a_n) = \lim_{n \to \infty} p(a_n)$

e) Il existe un nombre positif M tel que : $p(a) \leq M$ pour chaque atome $a \in L$.

Remarque :

a) Les exigences a, b, c, que nous avons imposées à l'état p nous semblent raisonnables. En ce qui concerne la condition d, elle est injustifiée par le fait que, comme il découle de la condition c, $p(\cup_n a_n) \geq p(a_n)\ \forall n$, donc $p(\cup_n a_n) \geq \lim p(a_n)$. Il s'agit donc, d'une condition de continuité qui n'est pas trop forte. Enfin, par e) nous exigeons aussi que p(a) soit fini pour les atomes. Cela signifie qu'il n'existe pas d'atomes avec une probabilité infinie par rapport aux autres. Car, s'il existait des atomes avec une probabilité relative infinie, la probabilité relative de chaque atome "a", serait ou bien $p(a) = 0$ ou bien $p(a) = \infty$, étant donné qu'un atome représente un évènement élémentaire. Puisque le treillis est atomique**, ceci sera valable pour toute proposition $a \in L$. Or une telle description représente un état des connaissances trop réduit pour présenter de l'intérêt. Mais par e) nous exigeons encore plus : que l'ensemble de nombres p(a) ("a" atome) soit borné. Cela découle

* Au sens de von Neumann (3)
** Axiome (Va)

probablement des autres axiomes. Nous l'avons cependant imposé indépendamment pour sécurité mathématique. Dans l'espace d'Hilbert, comme nous l'indiquerons plus tard, il peut être exprimé d'une manière dont la signification physique semble acceptable.

b) Lorsque deux états p et p_1 sont tels qu'il existe un nombre positif satisfaisant la relation :

$$p(a) = \lambda p_1(a), \ \forall a \in L$$

le rapport des probabilités relatives pour deux propositions quelconques est le même pour p et p_1. Les fonctions p et p_1 représentent donc le même état des connaissances. On dit que ces états sont équivalents.

Définition II

Un état régulier est un état qui satisfait $p(a) < \infty, \forall a \in L$. On voit qu'un état régulier est équivalent à l'état normalisé $p_1(a) = \frac{p(a)}{p(I)}$, où $p_1(I) = 1$. (Cette définition est équivalente à celle de Jauch).

Nous chercherons maintenant la forme générale des états dans le cas où le système de propositions L est isomorphe à l'ensemble des sous-espaces d'un espace hilbertien. Notamment, nous allons prouver que pour tout état p il existe un opérateur W continu, autoadjoint et positif tel que p(a) = Tr(WPa) où Pa est le projecteur sur le sous-espace correspondant à la proposition "a". Nous établissons d'abord le théorème inverse.

Théorème I

Pour tout opérateur W continu, autoadjoint, positif, la fonction $p_W(a) = Tr(WPa)$★ est un état. Si $Tr(W) < \infty$, alors p_W est un état régulier équivalent à $\frac{p_W}{TrW}$.

Démonstration

La propriété (a) des états étant vérifiée de façon évidente, nous commençons par la propriété (b).

b) si $a \subseteq b'$, alors les sous-espaces R_a, R_b correspondants à a, b sont orthogonaux et $R_{a \cup b} = R_a \oplus R_b$ d'où il vient :

★ P_a étant le projecteur sur R_a

$p(a \cup b) = p(a) + p(b)$.

c) Supposons que $p(a) = p(b) = 0$. Alors $Tr(W\,P_a) = Tr(W\,P_b) = 0$. Pour tout élément $x \in R_a$, on peut trouver une base $\{x_i\}$ de R_a telle que $x_1 = x$. Donc $Tr(WP_a) = \sum_i (x_i, Wx_i) = 0 \Rightarrow (x_i, Wx_i) = 0$★ d'où il vient : $(x, Wx) = 0, \forall x \in R_a$. De la même façon, $(y, Wy) = 0, \forall y \in R_a$. Considérons un élément $Z \in H$ tel que $Z = ax + by$, $a, b \in c$. On a :

$(Z,WZ) = (ax + by, W(ax + by)) = a\,b\,(x, Wy) + b\,a(y,Wx)$. Puisque W est autoadjoint on sait que : $|(Wx,y)|^2 \leqslant (Wx,x)(Wy,y)$ d'où il découle que $(WZ,Z) = 0$.

Mais l'ensemble des éléments Z de la forme $\overline{ax + by}$ où $x \in R_a$ et $y \in R_b$ est dense dans le sous-espace $[\overline{R_a \cup R_b}]$ engendré par R_a et R_b. Donc,

$$\forall Z \in [\overline{R_a \cup R_b}],\ (Z, WZ) = 0$$

Mais ce sous-espace est égal à $R_{a \cup b}$. Nous choisissons une base $\{Z_i\}$ de $R_{a\cup b}$. On obtient alors :

$$p(a \cup b) = Tr(WP_{a\cup b}) = \sum_i (Z_i, WZ_i) = 0$$

Nous avons ainsi démontré que $p(a) = p(b) = 0 \Rightarrow p(a \cup b) = 0$

d) Soit $a_1 \subseteq a_2 \subseteq a_3 \subseteq \ldots$, $P_n = Pa_n$ le projecteur sur R_a, $A_n = \{\phi_{i_n}\}_{i_n \in I_n}$ une base de $R_{a_n} - R_{a_{n-1}}$, $n = 2, 3, \ldots$ et $A_1 = \{\phi_{i_1}\}$ une base de Ra_1. Alors $A_1 \cup A_2 \ \ldots \cup A_n$ est une base de Ra_n et $\bigcup_n A_n$ est une base de $[\overline{\bigcup_n Ra_n}] = R_{\bigcup_n a_n}$.

Il est évident que :

$$Tr(W\,P_{\cup a_n}) = \sum_n \sum_{i_n} (\phi_{i_n}, W\phi_{i_n}) = \sum_n Tr(W\,(P_{n+1} - P_n)) + Tr\,(W\,P_1) =$$

$\lim_{n\to\infty} Tr(WP_n)$, donc $p(\cup a_n) = \lim_{n\to\infty} p(a_n)$

★ Puisque W est positif, $(y,Wy) \geqslant 0, \forall y \in H$

e) Pour tout atome "a", R_a est un espace unidimensionnel. Soit $\psi \in R_a, |\psi| = 1$ alors,

$$p(a) = Tr\ (WP_a) = (\psi, W\psi)$$

Mais W est continu, il possède donc une borne supérieure M, et

$$p(a) = (\psi, W\psi) \leq M \qquad \text{C.Q.F.D.}$$

Nous allons maintenant établir la réciproque du théorème I. Gleason ([5]) a montré que chaque état régulier admet la forme $p_w(a) = Tr(W\ Pa)$ où W est un opérateur positif, autoadjoint, de trace 1. En modifiant quelques théorèmes et définitions qu'il avait donnés nous allons prouver que tous les états -réguliers ou non- admettent la forme ci-dessus où W est un opérateur autoadjoint, positif, continu.

<u>Théorème II</u>

Chaque état admet la forme $p_w(a) = Tr(WP_a)$

<u>Démonstration</u>

Nous posons tout d'abord une définition ;

a) Une fonction de référentiel de poids T positif (réel ou infini) est une fonction $\Omega \xrightarrow{f} [0,\infty]$ (où $\Omega = \{X \in H\ ;\ |X| = 1\}$ et H est un espace d'Hilbert séparable) telle que :

$$\sum_n f(X_n) = T$$

pour toute base $\{X_n\}$ de H.

(Gleason avait considéré seulement le cas $T < \infty$).

Nous citons encore deux définitions données par Gleason :

b) Une fonction de référentiel est régulière s'il existe un opérateur W autoadjoint, continu, positif tel que

$$f(x) = (x, Wx) \quad x \in \Omega$$

c) Un sous-espace réel K de H est dit complètement réel si $(x,y) \in R \quad \forall\, x,y \in K$

<u>Lemme</u> : Soit f une fonction de référentiel sur H, régulière quand

elle est restreinte aux sous-espaces de H de dimension 2, et telle que :

$$\sup_{u\in\Omega} f(u) < \infty.$$ Alors f est régulière sur H.

La démonstration du lemme est celle du lemme 3.4 de Gleason. En effet, la démonstration donnée dans son travail couvre également le cas $T = +\infty$ en supposant que $\sup_{u\in\Omega} f(u) = M < \infty$.

Nous concluons la démonstration d'une manière analogue à celle des théorèmes 3.5 et 4.1 de Gleason.

Etant donné un état p défini sur l'ensemble des sous-espaces de H, $f(x)$★$= p(R_x)$, $|x| = 1$ est une fonction de référentiel sur H, de poids $T = p(H)$. Pour tout sous-espace complètement réel K de dimension finie, f est une fonction de référentiel sur K, de poids p(K). Soit X_i, $i = 1,2, \ldots n$ une base de K. Alors

$p(K) = \sum_{i=1}^{n} f(X_i)$ donc $p(K) < \infty$. Par conséquent, puisque le poids de f sur ces sous-espaces est fini, on peut utiliser les lemmes de Gleason qui les concernent :

Pour chaque sous-espace complètement réel de dimension 3, f est régulière (lemme 2.8). Elle est donc régulière sur chaque sous-espace complètement réel de dimension 2. Ce qui entraîne que f est régulière sur tout sous-espace de dimension 2 (lemme 3.3).

Alors, le lemme que nous avons donné entraîne que f est régulière sur H. Par conséquent, $f(x) = (x, Wx)$ où W est un opérateur autoadjoint, continu et positif.

Si A est un sous-espace, $\{y_i\}$ étant une base de A et $\{x_j\}$ une base de A, alors on trouve :

$$p(A) = \sum_i p(Ry_i) = \sum_i f(y_i) = \sum_i (y_i, Wy_i) = \sum_i (y_i,\ WP_a\ y_i) + \sum_j (x_j,\ WP_a\ x_j) = \mathrm{Tr}(W\ P_a)$$

Le théorème est établi.

★ R_x représente l'espace unidimensionnel qui contient x

Nous pouvons maintenant examiner ce que la condition $p(a) \leq M$ pour tout atome "a" implique : en omettant cette condition (qui est équivalente à $\sup_{u \in \Omega} f(u) < \infty$), on aurait à la place du lemme ci-dessus le suivant :

Lemme : Si f est une fonction de référentiel sur H, régulière quand elle est restreinte aux sous-espaces de dimension 2, il existe une forme hermitienne A(x,y) telle que

$$f(x) = A(x,x),\ \forall x \in \Omega$$

On peut prouver que A(x,y) admet la forme A(x,y) = (x,Wy), si l'on impose soit que $f(x) \leq M\ \forall x$ soit que f(x) est continue : Si $x_n \to x$, alors $f(x_n) \to f(x)$. Cette continuité se traduit par une continuité de p : Si $a_n \to a$★, "a_n" et "a" étant des atomes, alors $p(a_n) \to p(a)$. Nous voyons que dans l'espace de Hilbert, la condition $p(a) \leq M$ peut être remplacée par une condition de continuité qui est raisonnable.

Les théorèmes I et II établissent un élargissement du concept d'état au cas où Tr(W) peut être différente de 1★★. Cela nous permettra maintenant d'aborder certaines questions qui ne peuvent l'être à partir de la définition de Gleason, concernant la détermination de l'opérateur statistique lorsque on dispose peu d'information.

Détermination de l'état en fonction de l'information disponible

Dans ce qui suit nous allons examiner la question de savoir quelle représentation formelle l'on peut en général attribuer à un état en fonction des renseignements disponibles concernant le système.

Cherchons tout d'abord une représentation formelle d'un état exprimant que l'on sait qu'une proposition donnée "a" est vraie. Le problème qui se pose alors est de traduire la vérité

★ On peut définir cette convergence comme suit : $a_n \to a$ lorsqu'il existe dans les espaces R_{a_n}, R_a des vecteurs X_n, X tels que :

$$\frac{|(X_n, X)|}{|X_n| \cdot |X|} \to 1$$

★★ Rényi, Savage et Fine ont analysé des problèmes analogues à l'aide d'une axiomatique différente, voir (6) et (7).

certaine de "a" dans le choix de l'état p. On peut exiger que la proposition est équiprobable avec I : p(a) = p(I). Mais cela n'est pas satisfaisant si p(I) = ∞. En effet, il reste alors possible que l'on ait p(a') ≠ 0 ou même p(a') = ∞ ! Or, la vérité de "a", en termes physiques, signifie que la réponse à une mesure correspondante à "a" sera toujours "oui" ; donc, la réponse pour "a" sera toujours "non". D'où il découle que p(a') = 0. Cette dernière relation est l'exigence de départ dont nous avons besoin: en effet, de la relation p(a') = 0 il découle aussi bien que p(a) = p(I), que le théorème qui donne la réponse à la question générale soulevée plus haut ; nous allons trouver une représentation formelle de W, lorsque dans l'information que l'on possède, on sait qu'une proposition "a" est vraie.

Théorème III

Si la proposition "a" est vraie, l'état de nos connaissances étant $p_W(b) = Tr(WP_b)$, alors W est une intégrale des projecteurs qui projettent sur R_a. Notamment, $W = \int_0^\infty \lambda dF_\lambda$, où F_λ sont des projecteurs sur les sous-espaces de R_a.

Démonstration :

Puisque "a" est vraie, nous posons : p(a') = 0. Pour toute proposition b ⊥a, on trouve :

$$b \perp a \Rightarrow b \subseteq a' \Rightarrow p(b) = 0 \Rightarrow Tr(WP_b) = 0 \qquad (1)$$

Soit ϕ un vecteur normalisé, orthogonal à R_a. Alors on considère comme "b" le sous-espace engendré par ϕ et la relation (1) devient :

$$(\phi,\ W\phi) = 0, \forall \phi \perp R_a \qquad (2)$$

Soit E_λ^W la mesure projective associée à W, telle que $W = \int_{-\infty}^{+\infty} \lambda dE_\lambda^W$.
Puisque $W \geqslant 0$, le spectre de l'opérateur W n'est pas négatif, ce qui donne ;

$W = \int_0^\infty \lambda\ dE_\lambda^W$ et la relation (2) s'écrit :

$$\int_0^\infty \lambda d(\phi,\ E_\lambda^W \phi) = 0, \forall \phi \perp R_a \qquad (3)$$

Dans le domaine d'intégration, on voit que $\lambda \geqslant 0$. Donc (3) entraîne que $(\phi, E_\lambda^W \phi)$ est indépendant de λ, d'où l'on tire :

$$(\phi, E_\lambda^W \phi) = (\phi, E_0^W \phi) \Rightarrow (\phi, (E_\lambda^W - E_0^W)\phi) = 0$$

Mais l'opérateur $F_\lambda = E_\lambda^W - E_0^W$ est aussi un projecteur, donc $F_\lambda\ \phi = 0$, et F_λ projette sur un sous-espace R de R_a.

Enfin, puisque $dE_\lambda^W = d(E_\lambda^W - E_0^W) = d\ F_\lambda$ on trouve que

$W = \int_0^\infty \lambda\ d\ F_\lambda$ où F_λ projettent sur les sous-espaces de R_a.C.Q.F.D.

<u>Remarque</u> :

Lorsque le spectre de W est discret, par exemple $\{a_1, a_2, a_3, \ldots\}$, alors la mesure dF_λ est ponctuelle, et on peut écrire :

$$W = \sum_i a_i\ (Fa_i - Fa_i-) \text{ où } Fa_i- = \lim_{\lambda \to a_i-} F_\lambda$$

On voit alors que $p_W = \sum_i a_i\ p_{(Fa_i - Fa_i-)}$

Dans ce cas là, P_W est un mélange "discret" d' états. Si $\{\phi_j^i\}$ est une base pour le sous-espace correspondant à $Fa_i - Fa_{i-}$, alors $p_W = \sum_{ij} a_i\ p_{\phi_j^i}$. Donc p_W est un mélange des états purs $p_{\phi_j^i}$, tous les vecteurs ϕ_j^i appartenant à Ra.

Le résultat du théorème III nous aidera à répondre au problème suivant : quel opérateur statistique correspond à un état de connaissance consistant <u>seulement dans la certitude</u> que la proposition "a" est vraie ?

On pourrait essayer de résoudre ce problème en utilisant la méthode de von Neumann ([3]), généralisé par Jaynes ([2]) qui est la suivante : en utilisant la définition traditionnelle de l'état p_W (c'est-à-dire, Tr W = 1), on doit choisir parmi tous les états pour lesquels "a" est vraie, l'état possédant l'entropie la plus grande. Comme Tr W < ∞, le spectre W est discret et, suivant la remarque ci-dessus, P_W doit être un mélange des états purs $p_{\phi n}$ avec probabilités a_n, telles que $\phi_n \in Ra$ et $\sum_n a_n = 1$ puisque $Tr\ W = 1$. Alors l'entropie de p_W est définie comme $-\sum_n a_n\ \ell n a_n$. Donc, il faut maximaliser $-\sum_n a_n\ \ell n a_n$ avec la restriction $\sum_n a_n = 1$. Suivant le procédé usuel, on doit trouver un nombre X tel que :

$$\frac{d}{da_n}(-\sum_n a_n\ \ell n\ a_n) - \frac{d}{da_n} X(\sum_n a_n) = 0 \Rightarrow a_n = e^{-1-x}$$

Et la restriction $\sum_n a_n = 1$ donne $\sum_n e^{-1-X} = 1 \Rightarrow N\ e^{-1-x} = 1$

où N est la dimension de l'espace R_a.

Si $N < \infty$, le procédé est correct, on trouve que tous les a_n sont égaux à $\frac{1}{n}$ et que $W = \sum_n \frac{Pa_n}{N} = \frac{P_a}{TrP_a}$. Si $N = \infty$, alors on trouve qu'il n'existe pas de maximum. En outre, le résultat indique l'introduction d'états non normalisés, comme nous l'avons fait. Comme l'entropie n'a pas de sens au cas où $Tr\ W = \infty$, nous cherchons la réponse à ce problème par une voie différente.

Nous supposons donc que la seule information soit la vérité de "a". Il est possible de mesurer la proposition qui correspond à un sous-espace unidimensionnel engendré par le vecteur ψ, où $|\psi| = 1$ et $\psi \in R_a$. La probabilité d'obtenir la réponse "oui" est égale à $Tr(WP_\psi) = (\psi, W\psi)$ où p_W représente l'état correspondant à nos connaissances. Puisque dans le cas envisagé ici nous n'avons pas d'autre information que la vérité de "a", nous sommes obligés d'associer aux propositions concernant de tels cas des probabilités <u>indépendantes</u> de ψ. Donc, nous postulons que l'affirmation "nous savons seulement que "a" est vraie" se traduit par "si l'état est p_W, alors $(y,Wy) = (x,Wx)\ \forall\ x,y \in R_a$". Nous pouvons maintenant énoncer le théorème suivant :

<u>Théorème IV</u>

Lorsque la seule information est la vérité d'une proposition "a", l'état est équivalent à l'état p_E, où E est le projecteur sur R_a.

Démonstration

Le théorème III montre que pour l'état p_W, W admet la forme $W = \int_0^\infty \lambda \, d\, E_\lambda^W = \int_0^\infty \lambda \, d\, F_\lambda$ où $F_\lambda = E_\lambda^W - E_0^W$ et projettent sur les sous-espaces de R_a. Nous allons d'abord montrer que le spectre de W ne contient qu'un seul point sauf 0. En effet, supposons qu'il contiendrait deux points ν, μ et que $\nu > \mu > 0$.

Soit $\varepsilon = \frac{\nu-\mu}{4}$ et ϕ_1 et ϕ_2 : deux vecteurs des sous-espaces correspondant aux projecteurs $F_{\nu+\varepsilon}-F_{\nu-\varepsilon}$, $F_{\mu+\varepsilon}-F_{\mu-\varepsilon}$. On trouve :

$F_\lambda \, \phi_1 = 0, \forall \lambda < \nu-\varepsilon$ $\qquad$ $F_\lambda \, \phi_2 = 0, \forall \lambda < \mu-\varepsilon$

$F_\lambda \, \phi_1 = \phi_1, \forall \lambda > \nu+\varepsilon$ $\qquad$ $F_\lambda \, \phi_2 = \phi_2, \forall \lambda > \mu+\varepsilon$

et :

$$(\phi_1,W\phi_1) = \int_{\nu-\varepsilon}^{\nu+\varepsilon} \lambda \, d(\phi_1,F_\lambda \, \phi_1) \geqslant \int_{\nu-\varepsilon}^{\nu+\varepsilon} (\nu-\varepsilon) \, d(\phi_1,F_\lambda \, \phi_1) =$$

$$= (\nu-\varepsilon) \, [\, (\phi_1,F_{\nu+\varepsilon} \, \phi_1) - (\phi_1,F_{\nu-\varepsilon} \, \phi_1)] = \nu-\varepsilon$$

Donc, $(\phi_1,W\phi_1) \geqslant \nu-\varepsilon$

On peut montrer, de la même façon, que :

$$\nu + \varepsilon \geqslant (\phi_1,W\phi_1) \geqslant \nu - \varepsilon$$
$$\mu + \varepsilon \geqslant (\phi_2,W\phi_2) \geqslant \mu - \varepsilon$$

Puisque $\mu + \varepsilon < \nu - \varepsilon, (\phi_1,W\phi_1) \neq (\phi_2,W\phi_2)$, ce qui contredit l'hypothèse faite, que $(\phi_1,W\phi_1) = (\phi_2,W\phi_2) \; \forall \; \phi_1, \phi_2 \in$ Ra. Donc, le spectre de W ne contient qu'un seul point λ. Mais alors,

$W = \lambda(F_\lambda - F_{\lambda-})$ et comme $F_{\lambda-} = F_0 = E_0^W - E_0^W = 0$, on obtient $W = \lambda F_\lambda$. On sait que $F_\lambda \leqslant E$. Nous allons maintenant montrer que $F_\lambda = E$.

Supposons au contraire que $F_\lambda < E$. Alors $E - F_\lambda \neq 0$ et on peut choisir un vecteur ϕ appartenant à l'espace correspondant à

$E - F_\lambda$ et un vecteur ψ dans l'espace correspondant à F_λ. Tous les deux appartiennent aussi à Ra, car $F_\lambda < E$ et $E - F_\lambda < E$. Donc
$(\psi,W\psi) = (\phi,W\phi)$ (1)

En outre $(\psi,W\psi) = (\psi,\lambda F_\lambda\psi) = \lambda$,

$(\phi,W\phi) = (\phi,\lambda F_\lambda\phi) = 0$ donc $(\psi,W\psi) \neq (\phi,W\phi)$ ce qui contredit (1).

Nous avons montré que $W = \lambda E$. Etant donné qu'en général nous indiquerons par p_W l'état de nos connaissances, il sera en cas particulier, à indiquer par p_E.

Le théorème est une généralisation immédiate du résultat obtenu avec la méthode de Jaynes, dans le cas dim $R_a < \infty$. Il montre l'intérêt d'introduire les états non normalisables et il explique en même temps le sens physique de tels états : ils expriment le fait que nous possédons peu d'information. Dans le cas extrême dans lequel on ne sait absolument rien concernant le système physique, W est l'opérateur identité de H, puisque la seule proposition vraie a priori est la proposition I. Dans ce cas, si R_1, R_2 sont des sous-espaces, le quotient des probabilités relatives des propositions correspondantes sera : $\frac{\dim R_1}{\dim R_2}$.

Le changement de l'état après la mesure

Un problème de la Mécanique Quantique souvent discuté est celui de la réduction du vecteur d'état : lors d'une mesure d'un opérateur A à spectre discret et à espaces propres unidimensionnels, le vecteur d'état initial se transforme en un autre qui est vecteur propre de A. Cette réduction ne pose pas de problème dans une théorie des probabilités subjectives ; elle est une conséquence immédiate d'une remarque que nous avons déjà faite : l'état de nos connaissances change lorsque l'observateur obtient une nouvelle information concernant le système. Puisqu'il n'existe qu'un seul état compatible avec l'information obtenue, on n'a pas de problème de choix.

Nous allons examiner un problème plus général. Supposons que l'état de nos connaissances s'exprime par un état pur p_ψ où $|\psi| = 1$. Supposons également qu'une mesure d'une grandeur, nous indique que la proposition "a" est vraie (par exemple, "a" peut être la proposition : la valeur de la grandeur se trouve dans l'intervalle $[c,d]$). Quel est alors le nouvel état ?

Dans le cas où Ra n'est pas unidimensionnel, il faudra choisir parmi tous les états pour lesquels "a" est vraie, la forme générale des tels états étant donnée par le théorème III.

Si nous ne connaissons que l'état initial p_ψ et la vérité de "a" (par exemple, si nous ne connaissons pas la structure de l'appareil utilisé pour la mesure ou le résultat d'une autre mesure simultanée) nous proposons de choisir parmi les états possibles celui qui donne les probabilités les plus proches des probabilités données par p_ψ. Ainsi, nous ne perdons pas l'information disponible avant la mesure. Il faut donc que l'on donne une notion de "distance" entre les états, qui doit être une mesure de la différence des résultats prévus. Une telle distance, souvent utilisée pour deux fonctions f(x) et g(x) est la suivante : $\sup_x |f(x) - g(x)|$.

<u>Définition III</u> : On appelle distance des deux états réguliers normalisés p_1, p_2, l'expression :

$$d(p_1, p_2) = \sup_{a \in L} |p_1(a) - p_2(a)|.$$

On peut facilement prouver que cette distance donne à l'ensemble des états réguliers la structure d'un espace métrique.

A cause de l'interprétation de $d(p_1, p_2)$ comme mesure de la différence de $p_1(a)$, $p_2(a)$ nous définissons encore :

<u>Définition IV</u> : Lorsque p_1 est un état régulier et lorsque p_2 est non régulier, on a : $d(p_1, p_2) = \infty$

Nous allons déterminer maintenant la forme explicite de la distance de deux états purs.

<u>Théorème V</u> :

Pour deux états purs p_ψ, p_ϕ, on a :

$$d(p_\psi, p_\phi) = \sqrt{1 - |(\phi,\psi)|^2}$$

Démonstration :

Dans ce qui suit, P_X sera le projecteur sur le sous-espace engendré par le vecteur X.

On trouve facilement, si E représente un projecteur quelconque :

$$d(p_\phi,p_\psi) = \sup_E |\mathrm{Tr}\,(EP_\psi) - \mathrm{Tr}\,(EP_\phi)| \qquad (1)$$

La démonstration comprend trois étapes :

a) On prouve qu'il suffit de considérer des projecteurs sur des sous-espaces unidimensionnels.

b) Il suffit que les vecteurs qui engendrent ces espaces appartiennent à l'espace $[\overline{\phi,\psi}]$ engendré par ϕ,ψ.

c) On détermine alors le supremum.

a) On obtient :

$$\sup_E |\mathrm{Tr}(EP_\psi) - \mathrm{Tr}\,(EP_\phi)| = \sup_E |(\psi,E\psi) - (\phi,E\phi)| = \sup_E |(\psi,E(\psi-\phi)) + (E(\psi-\phi),\phi)|.$$

Soit $a = \frac{E(\psi-\phi)}{|E(\psi-\phi)|}$ alors $(a,\psi-\phi) = (\frac{E(\psi-\phi)}{|E(\psi-\phi)|}, \psi - \phi) = |E(\psi-\phi)|$

donc $E(\psi-\phi) = a(a,\psi - \phi) = P_a(\psi-\phi)$

Ce qui montre que pour chaque projecteur E, il existe un projecteur sur un vecteur approprié, a, tel que :

$$(\psi,\ E(\psi-\phi)) + (E(\psi-\phi),\phi) = (\psi,P_a(\psi-\phi)) + (P_a(\psi-\phi),\phi)$$

d'où il découle :

$$\sup_E |(\psi,E(\psi-\phi)) + (E(\psi-\phi)\phi)| = \sup_{\substack{a\in H\\|a|=1}} |(\psi,P_a(\psi-\phi)) + (P_a(\psi-\phi),\phi)| =$$

$$\sup_{\substack{a\in H\\|a|=1}} ||(a,\psi)|^2 - |(a,\phi)|^2|. \qquad (2)$$

b) Evidemment

$$\sup_{\substack{a\in H\\|a|=1}} \left| |(a,\phi)|^2 - |(a,\psi)|^2 \right| \geq \sup_{\substack{a\in[\overline{\phi,\psi}]\\|a|=1}} \left| |(a,\psi)|^2 - |(a,\phi)|^2 \right|$$

Soit $a = a_1 + a_2$, où $a_1 \in [\overline{\phi,\psi}]$, $a_2 \perp [\overline{\phi,\psi}]$. Alors

$$\left| |(a,\psi)|^2 - |(a,\phi)|^2 \right| = \left| |(a_1,\psi)|^2 - |(a_1,\phi)|^2 \right| \leq \left| |(\frac{a_1}{|a_1|},\psi)|^2 - |(\frac{a_1}{|a_1|},\phi)|^2 \right|, \qquad (3)$$

puisque $|a_1| < 1$

Ainsi, pour tout vecteur $a \in H$, $|a| = 1$, il existe un vecteur $\frac{a_1}{|a_1|} \in [\overline{\phi,\psi}]$ et $\left|\frac{a_1}{|a_1|}\right| = 1$, tel que la relation (3) soit valable.

Donc, $$\sup_{\substack{a\in H\\|a|=1}} \left| |(a,\psi)|^2 - |(a,\phi)|^2 \right| = \sup_{\substack{a\in[\overline{\phi,\psi}]\\|a|=1}} \left| |(a,\psi)|^2 - |(a,\phi)|^2 \right| \qquad (4)$$

c) Soit $a = k\phi + \lambda\psi$. Alors des relations (1), (2), (4) on déduit $d(p_\psi,p_\phi) = \sup\limits_{|a|=1} \left| |k|^2 - |\lambda|^2 \right| \; (1 - |(\phi,\psi)|^2)$ (5)

Nous voulons donc trouver le sup $\left| |k|^2 - |\lambda|^2 \right|$ sachant que

$$|a| = 1 \Rightarrow |k|^2 + |\lambda|^2 + \bar{k}\lambda(\phi,\psi) + k\bar{\lambda}(\psi,\phi) = 1 \qquad (6)$$

Si $k = X e^{i\rho}$, $\lambda = We^{i\sigma}$ où $X, W, \rho, \sigma \in R$, on trouve

$$|k|^2 - |\lambda|^2 = X^2 - W^2$$

$$(6) \Rightarrow X^2 + W^2 + XW\,(e^{i(\sigma-\rho)}(\phi,\psi) + e^{i(\rho-\sigma)}(\psi,\phi)) = 1$$

En utilisant la méthode des multiplicateurs de Lagrange, on trouve que $\sup(X^2 - W^2) = \max(X^2 - W^2) = \dfrac{1}{\sqrt{1 - |(\psi,\phi)|^2}}$ (7)

Finalement, de (5), (7) on tire :

$$d(p_\psi, p_\phi) = \sqrt{1 - |(\phi,\psi)|^2} \qquad \text{C.q.f.d.}$$

Nous avons donc prouvé que deux états p_ψ, p_ϕ ont une distance petite -donc, ils donnent des résultats proches- lorsque l'expression $|(\psi,\phi)|$ est voisine de l'unité. Cela signifie que les vecteurs ϕ, ψ sont proches dans l'espace d'Hilbert. Mais l'intérêt du théorème réside dans le fait que nous avons trouvé la forme exacte de $d(p_\psi, p_\phi)$. Cela nous aidera à atteindre notre but initial : choisir parmi <u>tous</u> les états possibles (purs ou mélanges) pour lesquels la proposition "a" est vraie, l'état le plus proche de l'état p_ψ. Le résultat est donné par le théorème suivant :

<u>Théorème VI</u> :

Soit "a" une proposition correspondant au sous-espace R_a et au projecteur P_a. Soit encore S l'ensemble des états pour lesquels "a" est vraie. Alors parmi tous les états de S, l'état le plus proche d'un état P_ψ est celui qui est représenté par la projection de ψ sur R_a, c'est-à-dire :

$$d(p_\psi, p_\phi) = \min_{p\in S} d(p_\psi, p) \quad \text{où } \phi = \frac{P_a\ \psi}{|P_a\ \psi|}$$

<u>Démonstration</u> :

Si Tr W = ∞ alors $d(p, p_W) = \infty$ et

$$d(p_\psi, p_\phi) < d(p_\psi, p_W) \quad \text{où } \phi = \frac{P_a \psi}{|P_a\ \psi|}$$

On peut donc considérer les seuls états p_W pour lesquels Tr W < ∞. Dans ce cas, W a un spectre discret et $p_W = \sum_n a_n\, p_{\phi_n}$, où a_n sont les valeurs propres et ϕ_n les vecteurs propres de W. On peut supposer que $\mathrm{Tr}W = 1 \Rightarrow \sum_n a_n = 1$.

En outre, on a prouvé que si la proposition "a" est vraie pour l'état p_W, alors $\phi_n \in R_a$ (voir la remarque à la suite du théorème III).

Pendant la démonstration du théorème V (étape b), on a vu que pour deux vecteurs quelconques ϕ, ψ on a

$$d(p_\psi, p_\phi) = \sup_{\substack{a \in [\overline{\phi,\psi}] \\ |a|=1}} \left| |(a,\psi)|^2 - |(a,\phi)|^2 \right| \qquad (1)$$

En outre, dans l'étape (c) nous avons prouvé que le supremum est en réalité un maximum. Donc, il existe toujours un vecteur $b \in [\overline{\phi,\psi}]$ tel que

$$d(p_\psi, p_\phi) = \left| |(b,\psi)|^2 - |(b,\phi)|^2 \right|$$

En outre, si $|(b,\phi)|^2 - |(b,\psi)|^2 < 0$ pour le vecteur $b = K\,\phi + \lambda\,\psi$, alors $|(b',\phi)|^2 - |(b',\psi)|^2 = -|(b,\phi)|^2 + |(b,\psi)|^2 > 0$ pour le vecteur $b' = K\psi + \lambda\phi$. On peut donc toujours choisir un vecteur $b \in [\overline{\phi,\psi}]$ tel que $d(p_\psi,p_\phi) = |(b,\phi)|^2 - |(b,\psi)|^2$. (1')

On a : $|(\phi_n, b)|^2 = |(\phi_n, P_a b)|^2 \leqslant |\phi_n|^2 \,.\, |P_a b|^2 = \left|\left(\frac{P_a b}{|P_a b|}, b\right)\right|^2$ (2)

Mais $b = K\,\phi + \lambda\psi = K\,\frac{P_a\psi}{|P_a\psi|} + \lambda\psi$ où $K, \lambda \in C \Rightarrow$

$P_a b = \frac{P_a\psi}{|P_a\psi|}(K + \lambda|P_a\,\psi|) \Rightarrow \frac{P_a\psi}{|P_a\psi|} = \rho\,\frac{P_a b}{|P_a b|}$ où $|\rho| = 1$

Alors (2) donne : $|(\phi_n,b)|^2 \leqslant \left(\frac{P_a\psi}{|P_a\psi|}, b\right)|^2, \quad \forall\ n$ (3)

En outre, si Pb est le projecteur sur b, alors

$$d(p_\psi, p_W) = \max_E |p_\psi(E) - p_W(E)| \geqslant p_\psi(Pb) - p_W(Pb) \qquad (4)$$

Et puisque $p_W(Pb) = Tr\,(WPb) = (b, Wb) = \sum_n a_n\,|(\phi_n, b)|^2 \leqslant$

$\sum_n a_n \left|\left(b, \frac{P_a\psi}{|P_a\psi|}\right)\right|^2 = \left|\left(b, \frac{P_a\psi}{|P_a\psi|}\right)\right|^2$

la relation (4) donne :

$$d(p_\psi,p_W) \geqslant |(b,\psi)|^2 - \left|\left(b,\frac{P_a\psi}{|P_a\psi|}\right)\right|^2$$

Soit, en vertu de (1') :

$$d(p_\psi, p_W) \geqslant d(p_\psi, p_\phi) \quad \forall\ p_W \in S$$

C.q.f.d.

Le théorème VI indique donc la sélection d'un état pur et notamment la projection de ψ sur R_a.

CONCLUSION

La généralisation de la notion d'état proposée plus haut semble indispensable lorsqu'on veut attribuer une représentation à un système physique -donc, lorsqu'on veut pouvoir exprimer certaines prévisions statistiques- cependant que l'information disponible est trop petite pour déterminer un opérateur statistique au sens de la Mécanique Quantique. En outre, l'interprétation donnée à la notion d'état en général permet un raisonnement qui donne la réponse à certains problèmes nouveaux examinés dans le texte.

Note ajoutée à la correction des épreuves :

La généralisation, proposée ici, du théorème de Gleason a été récemment refaite, sous une forme plus élaborée et paraîtra prochainement aux "Nuovo Cimento Letters".

REFERENCES

(1) J. M. Jauch : Foundations of Quantum Mechanics, Addison-Wesley, 1968

(2) E. T. Jaynes : "Information theory and statistical Mechanics" dans Phys. Rev. 106, 620 (1957 et Phys. Rev. 108, 171 (1957)

(3) J. V. Neumann : Mathematical foundations of Quantum Mechanics, Princeton, 1955

(4) C. Piron : Axiomatique Quantique, Thèse, Université de Lausanne

(5) A. Gleason : Journal of Math. and Mechanics : 6, 6, 885 (1957)

(6) A. Renyi : "Foundations of probability" 1970

(7) L. J. Savage : "The foundations of statistics", New-York, Wiley, 1954