UNIVERSITÉ DE REIMS

THESE

présentée à la

FACULTÉ DES SCIENCES

pour obtenir le grade de

DOCTEUR ES SCIENCES PHYSIQUES

par

Nicolas HADJISAVVAS

CRITIQUE DES FORMULATIONS LOGIQUES DE LA THÉORIE QUANTIQUE
INFÉRENCE STATISTIQUE ET DISTANCE HILBERTIENNE ENTRE ÉTATS QUANTIQUES

Soutenue le Mercredi 21 Octobre 1981 devant la commission d'examen :

M. B. DIU
M. J.M. LEVY-LEBLOND
M. B. MISRA
M^{me} M. MUGUR-SCHACHTER
M. R. PAYEN
M. G. PETIAU

Dédiée à

Πλατωνα, Borges, Varèse, et

Γιωργη Jorge Amarous τον εκ Ναξου

Je voudrais exprimer ma profonde reconnaissance à Madame le Professeur M. MUGUR-SCHÄCHTER, qui a bien voulu m'accueillir dans son laboratoire et qui, tout au long des nombreuses et fructueuses discussions que j'ai eues avec elle, m'a formé patiemment dans le sentier difficile de la recherche fondamentale.

Monsieur le Professeur R. PAYEN a bien voulu examiner avec attention la Thèse ainsi que les publications qui l'ont précédée, et m'a fait l'honneur de participer au jury. Je lui exprime ici toute ma gratitude.

Monsieur le Professeur B. MISRA a accepté, malgré les conditions difficiles, d'examiner ce travail et de faire partie du jury. Je lui en suis très reconnaissant.

Je remercie très vivement Monsieur le Professeur J.M. LEVY-LEBLOND qui m'a fait l'honneur de s'intéresser à mon travail et a accepté de participer au jury.

Monsieur le Professeur G. PETIAU a bien voulu examiner mes publications et ce mémoire. Je suis heureux de lui exprimer toute ma gratitude.

Monsieur le Professeur B. DIU m'a fait l'honneur de participer au Jury. Je l'en remercie très vivement.

Je tiens aussi à remercier le Docteur D. EVRARD pour les intéressantes et fructueuses discussions qu'on a eues ensemble, et Monsieur F. THIEFFINE, qui a collaboré à la publication des articles dont est tirée la première partie de la Thèse. J'associe également Madame PADET pour le climat amical qu'elle m'a réservé.

Madame PETIT et Messieurs DIETISCHIM, FLEISCHHACKER et Y. GRONET ont collaboré à la réalisation pratique de cette Thèse. Je leur exprime toute ma gratitude.

TABLE DES MATIERES

DEUXIEME PARTIE

CONTRIBUTIONS A L'ETUDE DU PROBLEME DE L'INFERENCE STATISTIQUE POUR LE CAS CLASSIQUE ET POUR LE CAS QUANTIQUE

PRÉFACE

Cette thèse se compose de deux parties entièrement distinctes, à ceci près qu'elles concernent toutes les deux la Théorie Quantique.

Un demi-siècle après l'établissement de la Mécanique Quantique, des nombreux problèmes concernant les fondements même de cette Théorie restent à résoudre. Ainsi, si dès 1932, la Théorie a été dotée d'un formalisme mathématique rigoureux et d'une interprétation plus ou moins cohérente de ses principaux éléments descriptifs (surtout de la fonction d'onde), ce formalisme était posé *ad hoc*, et était justifié seulement *a posteriori* par l'exactitude de ses prédictions. De même, l'interprétation statistique de la Théorie proposée à l'époque et largement admise aujourd'hui,a toujours laissé insatisfaits un certain nombre de physiciens qui concevait des exigences d'un autre ordre pour une théorie physique. Cette insatisfaction -ainsi que le désir d'augmenter la capacité descriptive de la Théorie- ont donné naissance à de nombreuses questions, dont plusieurs sont restées sans réponse jusqu'à ce jour. Ainsi, on se demande toujours s'il est possible d'arriver "déductivement" au formalisme Hilbertien de la Mécanique Quantique en se fondant sur l'expérience ; ou encore, si la théorie est compatible avec une conception réaliste et locale des faits physiques ; ou quel est l'opérateur position pour la M.Q.relativiste ; enfin, si la théorie quantique est compatible avec la Relativité Générale, etc.

L'étude de ses questions pendant toutes ces années a ouvert plusieurs voies de recherche. Parmi les plus nouvelles, citons, d'une part, les essais de "déduire" le formalisme Hilbertien en construisant d'abord une "logique Quantique", et d'autre part, l'étude du problème de l'inférence statistique (i.e. détermination d'un état à partir de l'information disponible). C'est à ces deux voies de recherche que seront consacrées les deux parties distinctes dont est constituée notre thèse.

La première partie, dont l'essence est contenue dans trois articles {1-3} , est vouée à une étude critique des formalisations dites logiques de la Mécanique Quantique, tout particulièrement celles de J.M. Jauch et de C. Piron.

La deuxième partie se compose de deux chapitres, consacrés respectivement à l'étude des versions classique et quantique de l'inférence statistique, et des deux annexes mathématiques qui nous sont utiles dans les chapitres précédants,mais qui constituent aussi des contributions à part. Cette deuxième partie a été publiée en quatre articles {4-7} .

Afin de faciliter la lecture, nous avons séparé, dans la mesure du possible, les considérations physiques du traitement mathématique. Ainsi, chacun des chapitres II à V est complété d'un appendice qui regroupe la plupart des démonstrations.

CONVENTIONS

La numérotation des relations et des théorèmes recommence à chaque chapitre. Le renvoi à une relation se fait par son numéro d'ordre, si la relation appartient au chapitre où se trouve ce renvoi, par exemple: relation (3); sinon le numéro de la relation est précédé par le numéro du chapitre auquel elle appartient, par exemple (II, 3). La fin des démonstrations est signalée par le symbole $\blacksquare$. Un nombre au dessus d'un symbole d'égalité ou d'implication renvoie à la relation en vertu de laquelle cette égalité ou implication est valable.

PREMIERE PARTIE

ETUDE CRITIQUE DE L'APPROCHE LOGIQUE DE LA THEORIE QUANTIQUE

CHAPITRE I

LES PREMIERES TENTATIVES DE DEVELOPPEMENT D'UNE "LOGIQUE QUANTIQUE"

1. INTRODUCTION

L'histoire de l'approche "logique" de la Théorie Quantique est presque aussi longue que l'histoire de la Théorie elle-même. En effet, dès 1932 J. Von Neumann {8} remarquait qu'à chaque proposition expérimentalement vérifiable correspond un projecteur dans l'espace d'Hilbert utilisé pour la description du système. Il a alors proposé une voie de recherche qui aurait pour but d'élaborer un "calcul propositionnel" quantique, par analogie au calcul propositionnel classique, en traduisant les propriétés bien connues des projecteurs en termes logiques.

Depuis cette époque, le développement de l'approche logique a franchit trois étapes importantes. La première date de 1936, lorsque G. Birkhoft et J. Von Neumann se sont proposé de découvrir "quelles structures logiques peut-on espérer trouver dans des Théories physiques qui, comme la Mécanique Quantique, ne s'adaptent pas à la logique classique" {9}. Ainsi, pour la première fois, on envisage une étude du problème qui ne serait pas strictement liée à la Mécanique Quantique. Dans le même article, les deux auteurs avancent l'idée que la structure qui serait apte à décrire ces nouvelles logiques pourrait être une sorte de géométrie projective.

A cet article fondamental succède une longue période de tâtonnements. On essaye de relier la "logique quantique" aux logiques polyvalentes etc. {10,11}. L'intérêt à propos des logiques est ravivé aux environs de 1960 par les travaux de J.M. Jauch et de son école {12-16}. La raison en est peut-être l'apparition en 1955 d'un article de Loomis sur la théorie de certains treillis {17} , qui va fournir la base mathématique pour l'avancement de l'approche logique. Finalement, en 1964 la deuxième étape du développement de cette approche a été franchie, lorsque C. Piron {18} arrive à trouver un système de postulats qui implique pour la "logique quantique" une structure bien connue, i.e. celle de l'ensemble des projecteurs d'un espace d'Hilbert. Ainsi, non seulement on proposait un "calcul propositionnel" bien établi, mais de surcroît le formalisme Hilbertien de la Mécanique Quantique <u>en découlait</u>.

Cela est apparu tout de suite comme une réussite formidable. En effet, l'approche logique semblait élucider certaines questions qui intriguait les physiciens, comme la descriptibilité des états par des vecteurs d'un espace d'Hilbert et celle des observables par des opérateurs, la linéarité de l'équation d'évolution etc. Qui plus est, une légère généralisation des postulats permettait à cette approche d'englober la Mécanique Classique, les règles de supersélection etc.

Cependant, des ombres persistantes sur le tableau ont modéré cet enthousiasme. Plusieurs des postulats proposés par C. Piron n'avaient aucune interprétation physique. La réussite donc restait caduque : il n'était même pas sûr que l'axiomatique proposée eût une parenté quelconque avec une Théorie physique, autrement que sur le niveau formel. Ainsi, on s'est vite aperçu qu'une telle axiomatique serait tout au plus un outil mathématique commode, et ne pourrait aucunement prétendre fournir une interprétation physique aux divers postulats du formalisme quantique, puisqu'elle-même n'en possédait aucune.

L'étape finale du développement de l'approche logique a été franchi pendant la dernière décennie, lorsque C. Piron a proposé dans une suite d'articles {19-23} une axiomatique enrichie, qui non seulement engendrait le formalisme quantique, mais en plus, ses axiomes semblaient avoir une signification physique claire. Si cela était vrai, alors la recherche sur la voie logique aurait pratiquement atteint son point final. Cependant, nous allons bientôt démontrer que cette nouvelle tentative d'établir une axiomatique sur des bases physiques a, elle aussi, échoué.

Les apports de Von Neumann { 8} et de Birkhoff et Von Neumann { 9} - qui ne sont en fait que des tâtonnements - ont été longuement commentés par d'autres auteurs {24-26}. Ainsi, nous nous contenterons ici d'un bref rappel situant le problème. Notre critique commencera au chapitre suivant par l'étude des travaux de Jauch et de son école {12-16,27}. Nous finirons par une étude beaucoup plus approfondie du nouveau système de Piron.

2. L'APPORT DE J. VON NEUMANN

Rappelons d'abord une partie du formalisme mathématique de la théorie quantique, ainsi qu'il a été établi par Von Neumann : tout système microscopique est décrit à l'aide d'un espace d'Hilbert H complexe et séparable. A chaque observable A on associe un opérateur auto-adjoint dans H (nous identifierons les symboles correspondants). Selon la théorie spectrale, à chaque opérateur auto-adjoint on associe une "mesure spectrale", i.e. une appliquation

$$\mathcal{B}(R) \ni \Delta \longrightarrow E^A(\Delta)$$

de l'ensemble des Boréliens de $\mathbb{R}$ dans l'ensemble des projecteurs de H, telle que :

$$1)\ E^A(\emptyset) = \{0\}, \quad E^A(R) = H,$$

$$2)\ \Delta_1 \cap \Delta_2 = \emptyset \implies E^A(\Delta_1) \cdot E^A(\Delta_2) = 0$$

$$3)\ \text{Si}\ \ \Delta_n \in \mathcal{B}(R)\ \ (n \in N)\ \text{et}\ \Delta_n \cap \Delta_m = \emptyset$$

pour $n \neq m$, alors $\quad E^A(\bigcup_n \Delta_n) = \sum_n E^A(\Delta_n)$

Les prévisions statistiques de la Théorie Quantique sont toutes basées sur le postulat suivant : si l'état du microsystème est représenté par un vecteur normalisé $\psi \in H$, alors la probabilité $p(A,\psi,\Delta)$ qu'en mesurant A on trouve une valeur appartenant à Δ est

$$p(A,\psi,\Delta) = \| E^A(\Delta)\psi\|^2 \tag{1}$$

Considérons maintenant une quelconque proposition concernant le système et qui peut être vérifiée ou falsifiée par une expérience. De telles propositions sont, par exemple, les suivantes : "La valeur de l'observable A est λ_1", " La valeur de l'observable A appartient à l'ensemble Δ", " La valeur de A est λ_1 et la valeur de B est λ_2", " La somme des carrés de A et de B est plus grande que 1" etc. Bien entendu, une proposition comme la troisième que nous venons de donner n'est vérifiable expérimentalement que si A et B sont simultanément mesurables.

Pour toute telle proposition , Von Neumann ({ 8 }, section III.5) a définit une nouvelle observable comme suit : l'observable peut prendre seulement les valeurs 1 et 0. Pour le mesurer, on effectue une expérience pour vérifier la proposition, et on lui donne la valeur 1 ssi la proposition a été vérifiée, sinon on lui donne la valeur 0. Selon le formalisme quantique , cette observable est représentée par un opérateur autoadjoint, et il est évident que ce dernier est,en fait, un projecteur. Ainsi s'établit une application $a \longrightarrow E_a$ qui à chaque proposition "a" associe un projecteur E_a. Puisque d'autre part, tout projecteur E correspond de façon biunivoque au sous-espace fermé R(E) sur lequel il projette, on a aussi une application $a \longrightarrow R_a \equiv R(E_a)$ qui à chaque proposition "a" associe un sous-espace fermé R_a.

Ces applications ainsi définies possèdent quelques propriétés dont nous retenons les plus frappantes :

a) Si b est une proposition du type : "La valeur de l'observable A appartient à l'ensemble $\Delta \in \mathcal{B}(R)$", alors $E_b = E^A(\Delta)$.

b) A la négation b' de b correspond $E_{b'} = 1-E_b$. En termes de sous-espaces, $R_{b'} = R_b^{\perp}$ (le complément orthogonal de R_b).

c) Si a et b sont simultanément mesurables, alors à la proposition "a et b" (qui est, bien sûr, une proposition mesurable) correspond le sous-espace $R_a \cap R_b$.

Comme le remarque Von Neumann, en s'inspirant des propriétés (b) et (c), on pourrait établir une sorte de "calcul propositionnel" au sens de la logique classique, mais enrichi par le concept de mesurabilité (ou "décidabilité") simultanée.

L'apport de V. N. peut donc être défini ainsi : *Primo*, en utilisant le formalisme quantique, il a établit une correspondance entre les propositions mesurables concernant un microsystème et les projecteurs. *Secundo*, il a commencé une étude de la structure de l'ensemble de ces propositions.

Nous ne nous attarderons pas sur l'article de Birkhoff et Von Neumann de 1936, non pas qu'il soit sans importance - au contraire - mais pour la simple raison que ce travail a été repris et considérablement développé par Jauch. Nous ferons donc une étude globale du travail de ce dernier, en nous basant surtout sur son livre "Foundations of Quantum Mechanics" {27} .

CHAPITRE II

ETUDE DU SYSTEME DE JAUCH.

1. LE BUT DU SYSTEME

L'objectif de Jauch était beaucoup plus ambitieux que celui de Von Neumann : il ne s'agissait plus d'édifier un calcul propositionnel à partir du formalisme Quantique et de son interprétation, mais plutôt de déduire le formalisme Quantique du calcul propositionnel.

Pour réaliser cet objectif, Jauch a construit son système sur la base d'une axiomatique. Son opinion sur ce que devrait être une axiomatisation d'une théorie physique est clairement exprimée dans ce court extrait de {28} : "In physics the interpretations of the axioms and the undefined terms is an essential part of the process (of the construction of an axiomatics)... The axioms therefore cannot be freely chosen, subject only to completeness and consistency, but should be inductively underpinned by a considerable body of empirical data. Without such a support on the solid ground of experience, physical axiomatics is sterile".

Nous nous proposons de montrer dans ce chapitre que l'axiomatique de Jauch ne répond nullement aux critères qu'il a lui-même imposés. Bien sûr cela était déjà connu : de nombreux auteurs ont critiqué, par exemple, le fameux postulat qui concerne l'existence d'une proposition "a et b" pour toute paire de propositions "a" et "b". Néanmoins, notre critique va beaucoup plus loin. Nous allons montrer, en particulier, que le système de Jauch est incohérent du point de vue sémantique. D'autre part nous examinerons les récentes tentatives { 29} de rémédier au problème que pose le postulat mentionné, et nous prouverons qu'elles échouent.

2. LE FORMALISME

Nous commencons en exposant sommairement le formalisme que nous allons examiner. Nous transcrivons les définitions de façon presque littérale de {27} , lorsque cela est possible.

DJ_1 (*définition de l'expérience "oui-non", proposition*). Une expérience "oui-non" est une observation qui admet l'un des deux termes d'une alternative comme résultat. Les expériences oui-non sont appelées aussi "propositions. L'ensemble de toutes les propositions sera noté par $\mathcal{L}$.

DJ_2 (*Relation $\subseteq$ entre propositions*). Il existe une relation entre certaines paires de propositions, exprimée par $a \subseteq b$, et qui signifie : chaque fois que a est vrai, b est vrai aussi.

(Note : Jauch ne définit pas le concept de vérité pour une proposition. D'après l'exemple illustrant sa définition pour $\subseteq$ dans {27}, p. 73, $a \subseteq b$ signifie que a et b sont simultanément mesurables et chaque fois que la réponse à a est "oui", la réponse à b l'est aussi).

DJ_3 (*Egalité des propositions*). Par définition, $a = b$ ssi $a \subseteq b$ et $b \subseteq a$.

Axiome AJ1. La relation $\subseteq$ est une relation d'ordre partiel dans $\mathcal{L}$. En d'autres termes, on a :

$$a \subseteq a, \forall a \in \mathcal{L}$$

$$(a \subseteq b \text{ et } b \subseteq a) \Longrightarrow a = b$$

$$(a \subseteq b \text{ et } b \subseteq c) \Longrightarrow a \subseteq c$$

DJ_4 (*plus grand minorant*). Soit $\{a_j\}_{j \in J}$ une famille quelconque d'éléments de $\mathcal{L}$. S'il existe une proposition b telle que

$$\forall x \in \mathcal{L} \quad : (x \subseteq a_j, \forall j \in J) \Longleftrightarrow x \subseteq b$$

alors nous appelerons b "le plus grand minorant de $\{a_j\}_{j \in J}$ " et nous le noterons par $\bigcap_J a_j$.

Axiome AJ2. (*Existence du plus grand minorant*). Pour toute famille de propositions $\{a_j\}_{j \in J}$, son plus grand minorant $\bigcap_J a_j$ existe*.

---

* Cet axiome sera longuement commenté dans la partie critique.

DJ_5 (*Proposition absurde*). L'axiome AJ2 implique l'existence d'une proposition ϕ avec la propriété $\phi \in a$, $\forall a \in \mathcal{L}$. Elle est définie par $\phi = \bigcap_{a \in \mathcal{L}} a$ et appelée proposition absurde, car elle est toujours fausse. ({27} , p. 76)

DJ_6 (*négation d'une proposition*). Pour toute proposition $a \in \mathcal{L}$, sa négation (ou son "orthocomplément") est une proposition $a' \in \mathcal{L}$ qui peut être mesurée par le même appareil que a. La seule différence est que les résultats sont inversés : si a est vraie, alors a' est fausse et vice versa.

Axiome AJ3 (*orthocomplémentation*). L'application $a \longrightarrow a'$ définie par DJ_6 est une orthocomplémentation, i.e. on a

$$(a')' = a$$

$$a' \cap a = \phi$$

$$a \subseteq b \Longrightarrow b' \subseteq a'.$$

DJ_7 (*orthogonalité*). Nous appelerons deux propositions a,b orthogonales (notation : $a \perp b$) ssi $a \subseteq b'$.

DJ_8 (*plus petit majorant*). Des trois axiomes précédents nous déduisons l'existence d'un plus petit majorant pour toute famille $\{a_j\}_{j \in J}$, $a_j \in \mathcal{L}$. Il est défini par $\bigcup_J a_j = (\bigcap_J a_j')'$ et on vérifie aisément qu'il obéit à la relation suivant

$$\forall x \in \mathcal{L} \quad : \quad (a_j \subseteq x, \forall j \in J) \Leftrightarrow \bigcup_J a_j \subseteq x$$

DJ_9 (*Proposition triviale*). La proposition $I = \bigcup_{a \in \mathcal{L}} a$ est appelée proposition triviale parce qu'elle est toujours vraie.

Rappelons la définition bien connue de treillis complets orthocomplémentés {30} :

DJ_{10} (*Treillis*). Tout ensemble possédant une relation d'ordre partiel par rapport à laquelle il existe le plus grand minorant et le plus petit majorant pour toute famille d'éléments, est appelé un treillis complet. S'il possède en plus une orthocomplémentation (cf. AJ3), il est appelé treillis complet orthocomplémenté.

DJ_{11} (*propositions compatibles*). Un sous-ensemble $S \subset \mathcal{L}$ est un ensemble de propositions compatibles ssi le treilli orthocomplémenté engendré par S est Booléen (i.e. distributif). Si {a,b} est un ensemble compatible, nous écrirons $a \leftrightarrow b$.

Axiome AJ4 (*modularité faible*). Nous postulons :

$$a \subseteq b \Longrightarrow a \leftrightarrow b$$

DJ_{12} (*proposition atomique*).Une proposition q est appelée proposition minimale ou atomique, si

$$q \neq \emptyset \quad \text{et} \quad (x \subseteq q,\ x \neq q) \Longrightarrow x = \emptyset$$

Axiome AJ5 (*Atomicité, loi de couverture*).

a) Pour toute proposition $a \neq \emptyset$, il existe une proposition atomique q telle que $q \subseteq a$.

b) Si q est une proposition atomique, alors :

$$a \subsetneq x \subseteq a \cup q \Longrightarrow x = a \text{ ou } x = a \cup q$$

DJ_{13} (*Système de propositions*). Nous appelerons tout ensemble $\mathcal{L}$ satisfaisant aux axiomes AJ1-5 un"système de propositions".

Les définitions DJ_1 - DJ_{13} et les axiomes AJ1 - 5 constituent le calcul propositionnel, ainsi qu'il a été établi par Jauch. Son contenu peut être résumé dans le postulat suivant :

Postulat PJ : Pour tout système physique, soit $\mathcal{L}$ l'ensemble des propositions qui lui correspond, d'après DJ1. Alors $\mathcal{L}$, muni de la relation d'ordre partiel définie par DJ_2 et l'orthocomplémentation définie par DJ_6, est un système de propositions, i.e. satisfait aux AJ1 - 5.

Le formalisme de Jauch se complète par la définition des concepts physiques de base : état et observable.

DJ_{14} (*Etat, définition physique*). "L'état est le résultat d'une série de manipulations physiques sur le système, qui constituent la préparation de l'état. Deux états sont identiques si les conditions qui entrent en ligne de compte pour la préparation de l'état sont identiques" ({27} , p. 22).

Ensuite, Jauch remarque que l'état physique définit une fonction p(a), $a \in \mathcal{L}$, p(a) étant la probabilité de la réponse "oui" pour l'expérience oui-non "a". Il a donc posé la définition mathématique suivante :

DJ_{15} (*Etat, définition mathématique*) . Un état est mathématiquement déterminé par une fonction réelle $p(a)$ sur $\mathcal{L}$. Nous allons imposer les exigences suivantes pour les états $p(a)$:

1) $0 \leqslant p(a) \leqslant 1$

2) $p(\emptyset) = 0$, $p(I) = 1$

3) Si $\{a_n\}_{n \in N}$ est une suite de propositions orthogonales, alors $p(\bigcup_N a_n) = \sum_N p(a_n)$

4) Pour toute suite de propositions $\{a_n\}_{n \in N}$, $p(a_n) = 1$ implique $p(\bigcap_N a_n) = 1$

5) (a) Si $a \neq \emptyset$, alors il existe un état p tel que $p(a) \neq 0$

(b)* Si $a \neq b$, alors il existe un état p tel que $p(a) \neq p(b)$

Jauch ne donne pas de définition physique pour les observables. Toutefois, la définition mathématique qu'il donne (p.98) découle des propriétés du concept physique bien connu. En effet, si A est une observable quelconque, on peut définir pour tout ensemble Borélien Δ une proposition $x^A(\Delta)$ ainsi $x^A(\Delta)$ est une expérience permettant de vérifier si la valeur de x appartient ou non à Δ. On voit que la fonction $x^A(\Delta)$ ainsi définie satisfait aux conditions contenues dans la définition suivante :

DJ_{16} (*Observable*). Un observable, ou σ - homomorphisme, est une application de $\mathcal{B}(R)$ (les boréliens de l'ensemble des réels) dans $\mathcal{L}$: $\mathcal{B}(R) \ni \Delta \xrightarrow{x} x(\Delta) \in \mathcal{L}$, qui satisfait aux conditions suivantes :

1) $x(\emptyset) = \emptyset$, $x(R) = I$

2) $\Delta_1 \cap \Delta_2 = \emptyset$ implique $x(\Delta_1) \perp x(\Delta_2)$

3) Pour toute suite Δ_n $(n = 1,2,...)$ de Boréliens Δ disjoints, on a : $x(\bigcup_N \Delta_n) = \bigcup_N x(\Delta_n)$.

On pose enfin une dernière définition :

---

* Apparemment Jauch n'a pas remarqué que 5(a) découle de 5(b)

DJ_{17} (*Principe de superposition*).* Nous dirons que $\mathcal{L}$ satisfait le principe de superposition si et seulement si pour toute paire de propositions atomiques e_1 et e_2 ($e_1 \neq e_2$) il en existe une troisième, e_3, telle que : $e_3 \neq e_1$, $e_3 \neq e_2$ et

$$e_1 \cup e_2 = e_1 \cup e_3 = e_2 \cup e_3$$

La construction du formalisme quantique habituel à partir du postulat PJ et des définitions DJ_{1-17} se fait à l'aide des théorèmes suivants :

Théorème TJ1 {18,27} : Pour tout système de propositions $\mathcal{L}$ qui satisfait au principe de superposition (et aussi à un autre ensemble de conditions assez générales) il existe un espace Hilbertien complexe et séparable H tel que $\mathcal{L}$ soit isomorphe au treillis des sous-espaces fermés (ou de manière équivalente, au treillis des projecteurs) de H.

Théorème TJ2 {31,32} : Pour tout état p il existe un opérateur W positif et de trace 1, tel que :

$$p(a) = \mathrm{Tr}(WE_a) \ , \ \forall a \in \mathcal{L}$$

où E_a est le projecteur correspondant à "a" via le Théorème TJ1.

---

* Cette formulation du principe de superposition peut surprendre ; elle est pourtant mathématiquement équivalente à celle habituellement posée en Mécanique Quantique. En effet : principe de superposition en M. Q. $\Longleftrightarrow$ il n'existe pas de règle de supersélection $\Longleftrightarrow$ (par définition) les seules observablesqui commutent avec toutes les autres observables, sont celles qui correspondent aux multiples de l'opérateur identité $\Longleftrightarrow$ les seules propositions compatibles avec toutes les autres sont $\emptyset$ et I $\Longleftrightarrow$ principe de superposition de Jauch (pour la démonstration de la dernière équivalence, voir {27} p. 107).

Théorème TJ3 : Pour toute observable x (cf. définition DJ_{15}) il existe un opérateur autoadjoint A tel que

$$E_{x(\Delta)} = E^{A}(\Delta) \quad , \forall \Delta \in \mathcal{B}(R)$$

$E^{A}(\Delta)$ étant le projecteur correspondant à Δ via la décomposition spectrale de A.

3. ETUDE CRITIQUE

3.1. *Remarques préliminaires*

Le système proposé par Jauch a eu un impact considérable parmi les physiciens qui s'occupent des fondements physiques et mathématiques de la Théorie Quantique, pour la bonne raison qu'il a conduit à quelques réussites indéniables. Tout d'abord, on a pu déduire, sur la base de ce système, la représentabilité des concepts physiques fondamentaux d'état et d'observable par des concepts mathématiques définis à l'aide d'un certain espace d'Hilbert. En fait , la totalité du formalisme Quantique en découl (nous avons ommis dans l'exposé précédant plusieurs "détails" comme le principe de supersélection, la loi d'évolution de l'état etc. qui ne nous interessent pas ici).

Un des défauts du système est l'absence de précision de certaines définitions donnant la valeur sémantique des concepts primaires. Par exemple, Jauch parle souvent des propositions "vraies" et pourtant, comme nous l'avons déjà noté pour la définition DJ_2, il n'a pas défini la notion de vérité pour des propositions. Toutefois, l'usage du terme implique que l'auteu appelle une proposition "vraie" ssi une mesure a été faite et la réponse "oui" a été trouvée.

Mais le système a aussi d'autres défauts, beaucoup plus graves. Nous allons en discuter aussitôt.

3.2. *Cohérence syntaxique et sémantique*

Nous avons déjà exposé dans la section II.1 l'opinion de Jauch sur les axiomatisations des théories physiques, opinion qui rejoint la nôtre. Nous pouvons l'exprimer par un ensemble de quatre conditions que toute telle axiomatisation doit remplir :

C1) (Cohérence syntaxique) Ses axiomes doivent être formellement compatibles au sens de la logique formelle.

C2) (Interprétabilité) Chaque axiome doit posséder une interprétation physique. En d'autres termes, il devrait exprimer un postulat physique.

C3) (Cohérence sémantique) L'axiomatique et son interprétation doivent être sémantiquement cohérentes.

C4) Les postulats qu'elle contient ne doivent pas être contredits par l'expérience.

Nous avons effectué quelques changements mineurs : la condition de complétude syntaxique est omise car elle serait en contradiction avec le théorème de Gödel {33} . Nous ne demandons pas que les postulats soient soutenus par un ensemble considérable de données expérimentales, mais nous imposons la condition beaucoup plus faible qu'ils ne soient pas contredits par l'expérience. Ainsi, le postulat de l'existence du positron n'était pas basé sur des données expérimentales lorsqu'il a été proposé. De même, le postulat de l'existence des quarks pourrait ne jamais être vérifié directement ; toutefois on le conserve parce qu'il fournit un modèle physique simple, avec des conséquences utiles. C'est pourquoi nous avons afaibli les conditions imposées par Jauch en ne demandant que l'interprétabilité physique et la non-contradiction avec l'expérience.

Enfin, nous avons ajouté la condition supplémentaire C3 dont la nécessité est évidente. Elle signifie tout simplement qu'il n'y a pas de contradiction interne du formalisme avec son interprétation. Un exemple de non-cohérence sémantique est fourni par les variables cachées locales et la Mécanique Quantique. Si le théorème de Bell est vrai, alors toute théorie à variables cachées locales implique une certaine inégalité. D'autre part, la mécanique quantique ne vérifie pas cette inégalité. De même, l'expérience semble aussi la contredire. Nous avons donc les implications:

variables cachées locales	$\Longrightarrow$	inégalité
Mécanique Quantique	$\Longrightarrow$	inégalité fausse
Expérience	$\Longrightarrow$	inégalité fausse

Puisque le postulat des variables cachées locales a une implication qui est contredite par l'expérience, il ne satisfait pas à la condition C4. Mais indépendamment de ce fait, les deux premières implications mettent déjà en contradiction la M.Q.

et les variables cachées locales. Par conséquent, la M. Q. avec des variables cachées locales est une théorie sémantiquement incohérente et ne satisfait donc pas à C3. (Nous disons sémantiquement car la démonstration du théorème de Bell n'est pas entièrement formelle, mais elle repose aussi sur la sémantique).

Il est évident que si l'une quelconque des conditions C1, C3 et C4 n'était pas remplie, l'axiomatique devrait être immédiatement rejetée. Par contre, si la condition C2 n'est pas remplie, on peut espérer de trouver dans l'avenir une interprétation physique convenable pour les axiomes qui n'en possèdent pas (c'est du moins la tactique suivie par plusieurs physiciens).

On peut vérifier facilement que l'axiomatisation exposée dans la section précédente satisfait à la condition 1. En effet, on sait que pour qu'un ensemble d'axiomes soit formellement compatible il suffit qu'il existe un modèle quelconque pour lequel ses axiomes sont des affirmations vraies { 34,35 }. Or, les axiomes AJ 1-5 sur lesquels repose le système de Jauch possèdent de tels modèles ; l'ensemble des sous-espaces de R^n en est un. Donc AJ1-5 sont compatibles.

La condition C2 n'est pas remplie. En effet, nous allons voir dans la section suivante que certains des axiomes n'ont aucune interprétation physique et en ce sens, l'axiomatique de Jauch n'est que partiellement interprétée.

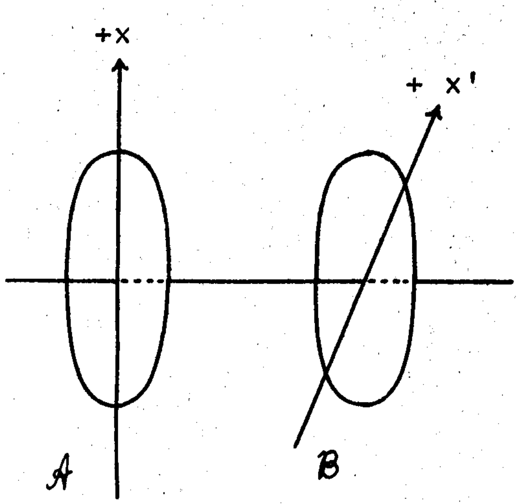
Mais ce qui est beaucoup plus grave, est que la condition C3 n'est pas remplie non plus. En effet, nous avons le théorème :

<u>Théorème 1</u> *Soit $\mathcal{L}$ l'ensemble des "propositions", défini par DJ_1, pour un système physique microscopique. Soit encore $\subseteq$ la relation d'ordre partiel définie par DJ_2 + DJ_3*

sur $\mathcal{L}$. Alors $\mathcal{L}$ ne peut pas être isomorphe à l'ensemble des sous-espaces fermés de l'espace d'Hilbert décrivant le système en Mécanique Quantique.

<u>Preuve</u>. Considérons, à titre d'exemple, une particule de spin $\frac{1}{2}$ (nous ne tenons pas compte des variables spatiales). Soit $\mathcal{A}$ et $\mathcal{B}$ deux appareils mesurant le spin le long des axes x et x' respectivement, et laissant passer la particule ssi on trouve $+\frac{1}{2}$ (un tel appareil est appelé "filtre"). La particule après passage par $\mathcal{A}$ ou $\mathcal{B}$ est supposée être dans l'état propre correspondant. Les appareils $\mathcal{A}$ et $\mathcal{B}$ définissent deux propositions a et b au sens de DJ_1, comme suit : a est la proposition "la particule passe par $\mathcal{A}$ " et b "la particule passe par $\mathcal{B}$ ".

Définissons maintenant une nouvelle expérience oui-non (i.e. proposition) a o b, qui consiste à faire passer la particule successivement par les deux filtres $\mathcal{A}$ et $\mathcal{B}$, la réponse "oui" étant donnée si la particule arrive effectivement à passer.



$\sphericalangle (x, x') \neq 0, \frac{\pi}{2}$

Figure 1

On a évidemment $a \circ b \subseteq a$, puisque $a \circ b$ et a sont d'une part mesurées simultanément et d'autre part si la réponse à $a \circ b$ est oui, la réponse à a l'est aussi. En plus, $a \not\subseteq a \circ b$ (a peut donner "oui" et b "non"). S'il y avait isomorphisme entre $\mathcal{L}$ et l'ensemble des sous-espaces fermés, on aurait pour les sous-espaces correspondants : $R_{a \circ b} \subseteq R_a$, $R_{a \circ b} \neq R_a$.

L'espace de Hilbert H décrivant la particule en Mécanique Quantique est de dimension 2. Puisque $a \neq \emptyset$ et $a \neq I$, (a n'étant ni toujours "vraie" ni toujours "fausse") on doit avoir $\dim (R_a) = 1$. Mais alors on devrait avoir $\dim (R_{a \circ b}) = 0$, ce qui implique $a \circ b = \emptyset$. Cela est faux, puisque $a \circ b$ n'est pas toujours fausse. Donc il ne peut pas y avoir isomorphisme entre $\mathcal{L}$ et l'ensemble des sous-espaces de H. ∎

Ainsi, nous avons démontré qu'une partie de l'interprétation -déjà partielle- du formalisme de Jauch, est suffisante pour entraîner des incohérences. Cette partie peut être facilement localisée en reformulant l'énoncé du théorème ainsi : Si l'on veut établir un "calcul propositionnel" tel qu'il y ait isomorphisme entre l'ensemble des propositions et l'ensemble des sous-espaces fermés de l'espace d'Hilbert du formalisme Quantique, il faut changer la définition DJ_1 du concept de proposition et/ou la définition $DJ_2 + DJ_3$ de la relation d'ordre.

Nous croyons qu'un changement au niveau de DJ_1 serait plus conforme à l'esprit de l'auteur. Il suffirait, probablement, d'appeler "proposition" toute expérience "simple", i. e. mesurant une grandeur dynamique ou un ensemble de grandeurs commensurables (comme c'était le cas chez Von Neumann), en excluant les expériences "oui-non" du type $a \circ b$ que nous avons utilisé dans la démonstration du théorème 1. Mais un tel procédé présupposerait que l'on ait défini le concept d'observable <u>avant</u> celui de proposition, et non pas après comme l'a fait Jauch. Il y aurait donc un changement de concepts primaires.

Une autre possibilité consiste à changer la définition de la relation d'ordre. En effet, nous allons voir que dans le système proposé par C. Piron où l'ordre est défini de façon différente,la démonstration du théorème 1 n'est plus valable. Cependant, nous verrons aussi que ce changement entraîne des graves difficultés au niveau de l'interprétation.

3.3. *Examen de l'interprétabilité physique des axiomes*.

Notre tâche dans cette section sera d'examiner dans quelle mesure le système de Jauch remplit la condition C2 d'existence d'une interprétation physique des axiomes. Bien sûr, cette question a été déjà dicutée par d'autres auteurs ; toutefois, nous allons voir que cette discussion se base souvent sur des idées fausses. En outre, notre analyse nous sera utile dans l'examen ultérieur du système de Piron.

Avant de passer à l'inspection des axiomes un par un, faisons une remarque : d'après le théorème 1, le système est sémantiquement incohérent, et cette incohérence provient d'une petite partie de son interprétation. Par conséquent, il semblerait que toute discussion sur l'interprétation du reste du formalisme serait superflue. Cependant, nous allons supposer qu'une des définitions (par exemple DJ_1) a été convenablement modifiée, de façon à ne pas entrainer d'incompatibilité. Nous commençons maintenant notre examen :

AJ1 Son interprétation physique est claire.

AJ2 Il est bien compris aujourd'hui que cet axiome est le talon d'Achille pour le système de Jauch, mais du reste une grande confusion règne en ce qui concerne l'emplacement exact du problème. Cette confusion est due, pour une grande partie, aux commentaires de Jauch sur cet axiome (cf. {27} ;P. 75).

Avant de commencer notre discussion sur l'axiome, nous allons d'abord clarifier quelques points délicats concernant la définition DJ_4 du "plus grand minorant". Considérons à titre d'exemple deux propositions $b,c \in \mathcal{L}$ et notons leur plus grand minorant, s'il existe, par $b \cap c$. Jauch a appelé cette proposition "b et c". Cela donne à croire que $\cap$ correspond tout à fait à l'opération de conjonction logique et que $b \cap c$ peut être mesurée en mesurant simultanément "b" et "c". C'est du moins l'avis de certains physiciens {26} . Cette interprétation de la définition DJ_4 est fausse, puisqu'elle mettrait en contradiction l'axiome AJ2 qui affirme l'existence de $b \cap c$ pour toute paire de propositions avec le fait bien connu qu'en théorie Quantique il y a des propositions qui ne peuvent pas être mesurées simultanément. En réalité "b et c" n'est qu'une appellation, et la proposition $b \cap c$ (à supposer qu'elle existe) pourrait n'avoir rien en commun avec le contenu physique de b et de c. En particulier, l'expérience oui-non $b \cap c$ (à supposer toujours qu'elle existe) n'est pas en général une mesure simultanée de b et de c. Une situation analogue se présente avec les observables : $\frac{p^2}{2m} + V(x)$ est formellement la somme des $\frac{p^2}{2m}$ et $V(x)$ et pourtant sa mesure n'est pas une mesure simultanée des $\frac{p^2}{2m}$ et $V(x)$.

Afin de justifier notre affirmation, rappelons la définition de $\bigcap_J a_j$: c'est l'élément de $\mathcal{L}$ qui rend vraie la double implication :

$$\forall x \in \mathcal{L} : \quad x \subseteq \bigcap_J a_j \iff x \subseteq a_j, \forall j \in J \qquad (1)$$

Il est facile de montrer que (1) est équivalente à la paire de relations :

$$\bigcap_J a_j \subseteq a_j \;, \; \forall j \in J \qquad (2a)$$

$$x \subseteq a_j \;, \forall j \in J \Longrightarrow x \subseteq \bigcap_J a_j \qquad (2b)$$

La relation (2a) signifie que $\bigcap_J a_j$ est un minorant commun des a_j , tandis que (2b) signifie que si x est un autre minorant commun, alors $\bigcap_J a_j$ est plus grand que x. Cela montre pourquoi Jauch appelle $\bigcap_J a_j$ "le plus grand minorant". Il montre aussi que $\bigcap_J a_j$ ne dépend pas seulement des a_j, mais de toute la structure du treillis, puisque il faut connaître tous les minorants des a_j pour connaître aussi $\bigcap_J a_j$. C'est pourquoi la signification de la proposition $b \cap c$ pourrait n'avoir rien en commun avec les significations de b et de c.

En conclusion, la définition purement mathématique de $b \cap c$ n'implique nullement que $b \cap c$ est une mesure simultanée de b et de c. Mais d'autre part, elle n'implique pas non plus l'existence de $b \cap c$. En effet, il y a des ensembles ordonnés tels que le plus grand minorant de deux éléments peut ne pas exister, comme dans l'exemple ci-dessous :

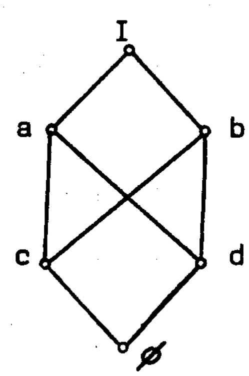


Figure 2

Ici deux éléments sont dans une relation d'ordre s'ils sont unis par une ligne, celui qui est en bas étant plus petit que l'autre. Les éléments a,b ont trois minorants communs : c , d et $\emptyset$. Il n'y a évidemment pas de plus grand entre eux.

Venons maintenant à l'étude de l'axiome AJ2. Cet axiome affirme que le plus grand minorant existe pour toute famille de propositions, et impose ainsi sur l'ensemble des propositions expérimentales $\mathcal{L}$ une structure de treillis complet. Cette affirmation peut être justifiée physiquement dans le cas spécial d'une famille de propositions commensurables. En effet, supposons que toutes les propositions a_j, $j \in J$ puissent être mesurées simultanément. Alors on peut définir une nouvelle proposition, notée b, comme suit : pour mesurer b on mesure simultanément toutes les a_j, $j \in J$ et on donne la réponse "oui" ssi toutes les réponses aux a_j sont "oui". On peut montrer que b ainsi définie satisfait aux conditions (2 a-b). En effet , on a: $b \subseteq a_j$, $\forall$ $j \in J$, car si b est trouvée vraie alors toutes les a_j sont vraies. D'autre part, $x \subseteq a_j$, $\forall$ $j \in J$, implique que $x \subseteq b$, car si x est vraie, alors de $x \subseteq a_j$ découle que les a_j sont vraies donc b est vraie, ce qui démontre $x \subseteq b$.

Puisque (2a-b) est équivalente à (1), nous concluons que le plus grand minorant $\bigcap_J a_j$ des a_j existe et qu'il est égal à b. Par conséquent, dans le cas d'une famille de propositions simultanément mesurables, l'existence de $\bigcap_J a_j$ peut être déduite sur des bases physiques.

Nous sommes maintenant en mesure de localiser le problème posé par l'axiome AJ2. Selon les propres termes de Jauch, les axiomes d'une théorie physique devrait "être soutenus par un ensemble considérable de donnés expérimentales". Or l'existence du plus grand minorant imposé par l'axiome AJ2 n'est justifiée physiquement que pour des familles $\{a_j\}_{j \in J}$ de propositions simultanément mesurables. En effet, on n'a jamais pu proposer une constructio[n] valable de $\bigcap_J a_j$ pour des a_j non simultanément mesurables.* Cela signifie que "l'ensemble considérable de données expérimentales"

---

* L'idée de construire $\bigcap_J a_j$ en utilisant des filtres infinis {27, 36 } est aujourd'hui abandonnée. D'autre part, C. Piron a pu construire $\bigcap_J a_j$ en changeant complètement les définitions du concept de proposition et de relation d'ordre, mais il a créé des nouveaux problèmes dans d'autres axiomes, comme nous allons voir dans le chapitre suivant.

sur lequel se base l'affirmation de l'existence de la conjonction pour de telles propositions est vide. Pis encore, l'axiome AJ2 n'a aucune signification physique.

Il y a eu des tentatives de remédier à cette défaillance du système de Jauch. Leur examen fera l'objet de la section suivante. Pour le moment, poursuivons par l'étude des trois autres axiomes.

AJ3. Son interprétation est évidente.

AJ4. L'interprétation decet axiome aussi est solide. En effet, d'après la définition de la relation d'ordre DJ2, si $a \subseteq b$ alors a et b sont simultanément mesurables. D'après la définition DJ6, les 4 propositions a,b,a',b' sont simultanément mesurables. Il s'ensuit, d'après notre discussion de l'axiome AJ2, que les propositions a,b,a',b', $a \cap b$, $a' \cap b$, $a \cup b$ etc sont toutes simultanément mesurables (il suffit de mesurer a et b et d'appliquer les tableaux de vérité de la logique habituelle pour obtenir la valeur de vérité de toutes les autres). Par conséquent, la loi distributive sera vérifiée, ou en d'autres termes : $a \longleftrightarrow b$ (cf. DJ_{11}). Cela justifie l'axiome AJ4.

AJ5 . Jauch lui même admet que cet axiome n'a pas de signification physique ({27} , p. 87).

3.4. *Etude des tentatives d'amélioration du système*

3.4.1. Un nouveau programme de recherche.

Nous venons d'exhiber les principaux handicaps du système de Jauch : son incohérence sémantique et le manque total de signification physique pour les axiomes AJ2 et AJ5. Le premier de ces handicaps qui, à notre connaissance, n'a pas été jusqu'à maintenant remarqué, est sans doute le plus grave. Toutefois, on peut espérer le surmonter par une modification adéquate des définitions des concepts de propositions et de relation d'ordre. En ce qui concerne les deux axiomes, la recherche d'une signification physique acceptable a suivi deux voies distinctes. Parmi elles, celle de C. Piron{ 19-23} a conduit à un formalisme très riche qui a convaincu plusieurs physiciens {28,37 } par son efficacité apparente. Ceux qui ont suivi l'autre voie ont essayé d'arriver au formalisme Hilbértiensans faire usage des axiomes AJ2 et AJ5.

Cette dernière tentative se base sur l'argument suivant : on sait que les propositions de la forme " la valeur de l'observable A appartient à l'ensemble Δ " correspondent à des sous-espaces fermés de l'espace de Hilbert qui décrit le système (Cf. section I.2). On sait aussi que l'ensemble de tous les sous-espaces fermés est un treillis complet, orthocomplémenté, faiblement modulaire, atomique et satisfaisant à la loi de couverture (c.à d. il satisfait aux axiomes AJ1-5). D'autre part, il y a des raisons physiques qui donnent à penser que JA2 et JA5, qui affirment l'existence d'éléments qui possèdent certaines propriétés, ne sont probablement pas vrais pour l'ensemble des propositions. On en déduit que l'ensemble des propositions pourrait ne pas être isomorphe à l'ensemble de <u>tous</u> les sous-espaces fermés, mais seulement à un sous-ensemble de celui-ci.

De là on est arrivé au programme de recherche suivant : on impose à l'ensemble des propositions $\mathcal{L}$ seulement les axiomes AJ1, AJ3, AJ4 qui ont une justification physique. Ensuite on peut essayer de montrer que l'on peut compléter $\mathcal{L}$ avec des éléments supplémentaires de façon que le nouvel ensemble $\mathcal{L}'$ qui en résulte satisfasse tous les axiomes AJ1-5. Dans un langage plus formel : on devrait montrer qu'il existe un ensemble $\mathcal{L}'$ satisfaisant aux AJ1-5 et un isomorphisme entre $\mathcal{L}$ et un sous-ensemble de $\mathcal{L}'$. Les éléments supplémentaires de $\mathcal{L}'$ par rapport à $\mathcal{L}$ n'auraient probablement pas de correspondant physique, et seraient simplement une sorte d'intermédiaire de calcul. Ensuite, en utilisant le théorème TJ1 pour ce $\mathcal{L}'$, on en déduirait* l'isomorphisme de $\mathcal{L}$ avec un sous-ensemble de l'ensemble de tous les sous-espaces fermés d'un espace de Hilbert.

Si ce nouveau programme avait abouti, on serait effectivement arrivé à la construction du formalisme quantique sur une base axiomatique dont tous les axiomes ont une signification physique. Parmi les physiciens qui ont suivi une voie de recherche analogue, W. Guz {40,41} a utilisé le formalisme de G. Mackey {42}; K. Bugajska et S. Bugajski {43} ont aussi utilisé un formalisme plus riche que celui de Jauch. Par contre, J. Cyranski est resté à l'intérieur du formalisme de Jauch et il a suivi presque exactement le programme que nous venons d'exposer. C'est pourquoi nous allons étudier sa proposition à part.

---

* Sous l'hypothèse supplémentaire que $\mathcal{L}'$ satisfasse encore deux conditions, Cf. {38,39}.

3.4.2. *Travail de J. Cyranski* {44}

L'auteur a utilisé un théorème fondamental de la théorie des ensembles ordonnés :

<u>Théorème 2</u>({45} ,p. 14): *Soit* E *un ensemble partiellement ordonné quelconque, contenant deux éléments* I, ϕ *tels que* $\phi \leq a \leq I$, $\forall a \in E$. *Alors il existe un treillis complet* $\tilde{E}$ *et un isomorphisme entre* E *et un sous-ensemble* E_1 *de* $\tilde{E}$ *qui engendre* $\tilde{E}$ *en ce sens, que pour chaque* $a \in \tilde{E}$ *il existe une famille d'éléments de* E_1, $\{x_j\}_{j \in J}$, *telle que* $a = \bigwedge_J x_j$.

Nous ne donnons pas ici la démonstration de ce théorème qui se trouve dans toutes les monographies sur la théorie des treillis. Mais nous exposons la méthode de construction de $\tilde{E}$ et de l'isomorphisme, car nous en ferons bientôt usage. Pour chaque $A \subset E$, soit $s(A)$ l'ensemble de ses majorants et $i(A)$ l'ensemble de ses minorants. Alors $\tilde{E}$ est défini par :

$$\tilde{E} = \{A \subset E : i(s(A)) = A\} \tag{3}$$

et l'isomorphisme entre E et un sous-ensemble de $\tilde{E}$ par l'application biunivoque :

$$E \ni x \longrightarrow i(x) \in \tilde{E}. \tag{4}$$

$\tilde{E}$ s'appelle "extension de Dedekind" de E.

Etant donné que l'ensemble des propositions $\mathcal{L}$ satisfait à l'axiome AJ1 et qu'il possède deux éléments extrêmes I et ϕ , Cyranski a suggéré que l'axiome AJ2 ne constitue pas un véritable obstacle. En effet, si AJ2 n'est pas vrai pour $\mathcal{L}$, on peut remplacer $\mathcal{L}$ par son extension de Dedekind $\tilde{\mathcal{L}}$ qui satisfait aussi AJ2 et dont l'existence est garantie par le théorème.

Malheureusement, ce procédé n'est pas correct. L'ensemble $\mathcal{L}$ satisfait aussi aux axiomes AJ3 et AJ4. Si l'on veut donc le remplacer par une extension $\mathcal{L}' = \tilde{\mathcal{L}}$, il faudra que l'extension aussi satisfasse aux axiomes AJ3 et AJ4 (sinon, elle ne pourrait pas être isomorphe à l'ensemble des sous-espaces fermés d'un espace Hilbértien,comme on le souhaite). Or nous avons le théorème suivant, dont la démonstration est donnée en appendice :

<u>Théorème 3</u>. *Soit* E *un ensemble satisfaisant aux*

axiomes AJ1, AJ3, AJ4. Alors son extension de Dedekind $\tilde{E}$ satisfait les axiomes AJ1, AJ2, AJ3 mais pas nécessairement AJ4.

On voit donc qu'un essai d'escamoter le problème posé par l'axiome AJ2 en utilisant l'extension de Dedekind, créerait de nouvelles difficultés avec l'axiome AJ4. On ne peut donc pas accepter la suggestion de Cyranski.

3.4.3 Travail de Griechie

Griechie {46,47} a démontré un théorème très puissant, qui affirme qu'il existe des ensembles qui satisfont aux axiomes AJ1, AJ3, AJ4 et qui ne peuvent pas être immergés dans l'ensemble des sous-espaces fermés d'un espace d'Hilbert. Ainsi le théorème de Griechie donne une réponse définitivement négative à la question de savoir si l'on peut retrouver le formalisme Hilbérien en se basant uniquement sur les axiomes AJ1, AJ3, et AJ4 qui ont une signification physique.

4. LA NECESSITE D'UN AUTRE SYSTEME

Le programme de Jauch n'a pas abouti. Le système qu'il a construit souffre d'une incohérence sémantique, et possède deux axiomes qui n'ont aucune signification physique. Cette déficience du système est d'autant plus grave que l'un de ces axiomes, le AJ2, constitue le pilier de sa construction.

D'autre part, nous avons vu que les tentatives de retrouver le formalisme Hilbértiende la Mécanique Quantique en utilisant seulement les axiomes possédant une signification physique, ont échoué. Il apparait donc que pour arriver à la formalisation Hilbértienneil faudra transgresser le formalisme de Jauch. Une approche prometteuse dans cette direction est celle de W. Guz {40 } qui utilise le formalisme de G. Mackey {42} .

Néanmoins malgré ses échecs, le formalisme "logique" a eu une répercussion considérable chez les physiciens qui s'occupent des fondements conceptuels et mathématiques de la théorie Quantique. Cela s'explique tant par le but de l'approche logique, qui était extrêmement ambitieux, que par certains résultats (succès indéniables en ce qui concerne l'introduction physiquement acceptable des concepts d'état et d'observable et leur représentation mathématique qui en découle, possibilité d'avoir un langage universel pour traiter à la fois la Mécanique Quantique et la Mécanique Classique etc). Ces succès, ainsi que l'importance de l'enjeu, ont incité plusieurs physiciens à essayer d'améliorer le système à travers un plus grand raffinement des concepts. Parmi les nouveaux systèmes qui ont été ainsi construits, celui de C. Piron est de loin le plus élaboré. Plusieurs physiciens ont été convaincus que ce nouveau système possède des bases physiques solides, et constitue donc l'aboutissement heureux de l'approche logique. Le but du chapitre suivant est de montrer que cela n'est pas vrai.

APPENDICE AU CHAPITRE II

Preuve du théorème 3 : Par construction $\tilde{E}$ satisfait AJ1 et AJ2. Il faut donc montrer qu'il satisfait AJ3, mais pas toujours AJ4.

a) Pour montrer AJ3, nous allons utiliser la relation:

$$\forall A \in \tilde{E} : A = \bigcup_{x \in A} i(x) = \bigcap_{y \in s(A)} i(y) \qquad (5)$$

(Cf. {45} p. 16), où $\bigcap$ et $\bigcup$ sont l'union et intersection ensembliste. De (5) il découle

$$\forall A \in \tilde{E} : \bigcap_{x \in A} i(x') = \bigcup_{y \in s(A)} i(y') \qquad (6)$$

En effet, on a les implications :

$$(y \in s(A) \text{ et } x \in A) \Longrightarrow y \geqslant x \Longrightarrow y' \leqslant x' \Longrightarrow i(y') \subset i(x')$$

qui montrent que $\left(\bigcap_{x \in A} i(x')\right) \supset \left(\bigcup_{y \in s(A)} i(y')\right)$. D'autre part,

$$z \in \bigcap_{x \in A} i(x') \Longrightarrow z \leqslant x', \forall x \in A \Longrightarrow z' \geqslant x, \forall x \in A \Longrightarrow$$
$$\Longrightarrow z' \in s(A) \Longrightarrow z \in \bigcup_{y \in s(A)} i(y')$$

montre l'inclusion inverse, donc (6) est vraie. Nous pouvons donc définir

$$\forall A \in \tilde{E} : \quad A' := \bigcap_{x \in A} i(x') = \bigcup_{y \in s(A)} i(y') \qquad (7)$$

Montrons que $A' \in \tilde{E}$. On a, d'après (7) :

$$u \in s(A') \Longleftrightarrow u \geqslant z, \forall z \in A' \Longleftrightarrow u \geqslant y', \forall y \in s(A) \Longleftrightarrow$$
$$\Longleftrightarrow u' \leqslant y, \forall y \in s(A) \Longleftrightarrow u' \in A. \qquad (8)$$

Donc

$$z \in i(s(A')) \Longrightarrow z \leqslant u, \forall u \in s(A') \Longrightarrow z \leqslant u', \forall u \in A \Longrightarrow z \in A'$$

ce qui montre que $A' \in \tilde{E}$ (Cf. la relation (3)).

Il reste à démontrer que l'application $A \longrightarrow A'$ ainsi définie satisfait aux 3 conditions de l'axiome AJ3. En combinant (7) et (8) avec (5) on trouve :

$$(A')' = \bigcup_{y \in s(A')} i(y') = \bigcup_{y' \in A} i(y') = A.$$

D'autre part,

$$A \subset B \Longrightarrow \Big(\bigcap_{x \in B} i(x')\Big) \subset \Big(\bigcap_{x \in A} i(x')\Big) \Longrightarrow B' \subset A'$$

Et enfin,

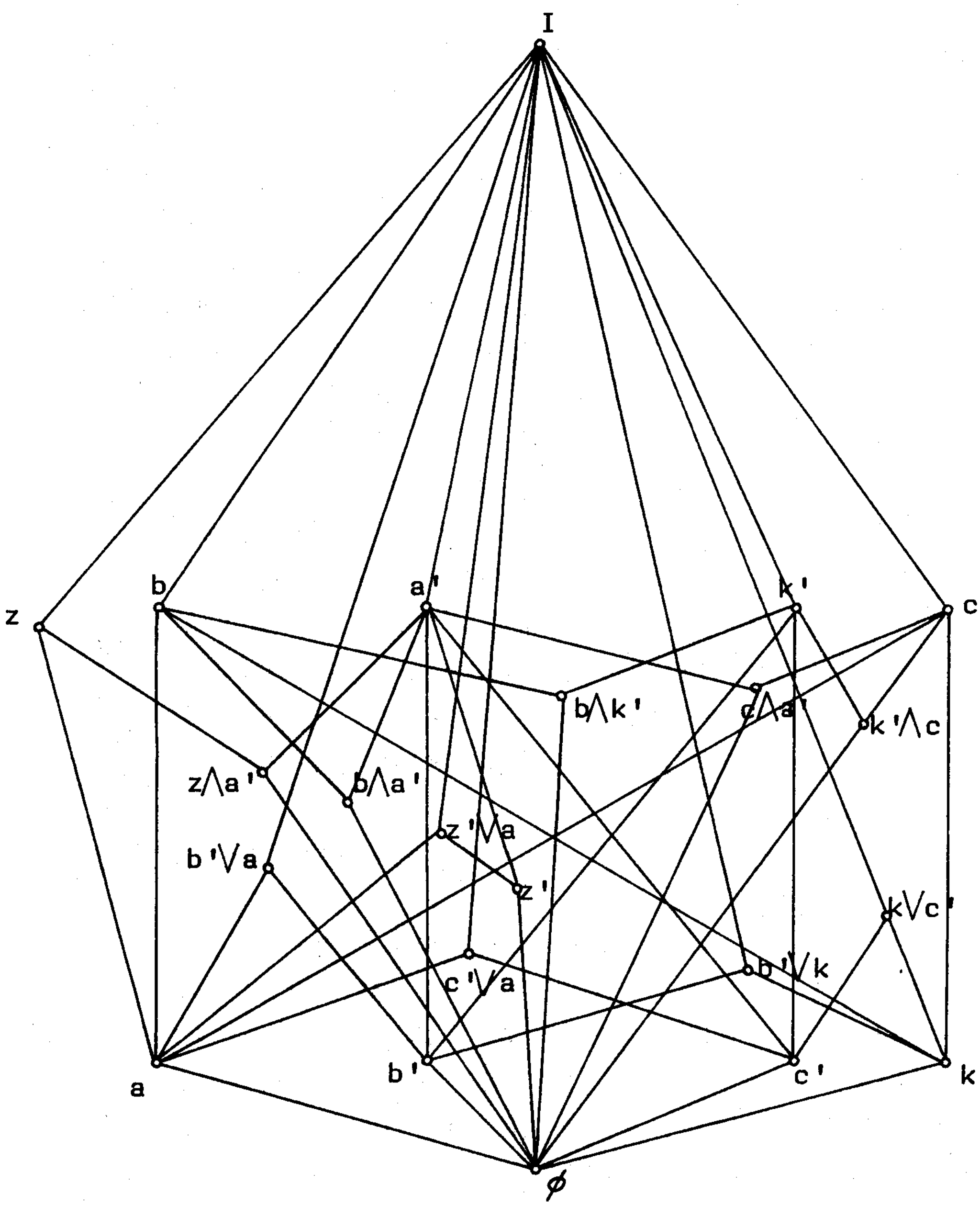
$$y \in A \cap A' \Longrightarrow y \in A \text{ et } (y \in i(x'),\ \forall x \in A) \Longrightarrow y \leqslant y' \Longrightarrow$$
$$\Longrightarrow y = \emptyset$$

et donc $A \cap A' = i(\emptyset)$. Par conséquent, $\tilde{E}$ satisfait l'axiome AJ3, i.e. il est orthocomplémenté.

b) Montrons qu'il existe un E qui satisfait AJ1, AJ3 et AJ4 tel que $\tilde{E}$ ne satisfait pas AJ4. Un tel E est donné par la représentation graphique de la page suivante (figure 3). Considérons son extension de Dedekind $\tilde{E}$ et l'isomorphisme i définis par les relations (3) et (4). Nous utiliserons le symbole $\cap$ pour noter l'intersection ensembliste et $\wedge$ pour le plus grand minorant. On a :

$$i(a) = \{a, \emptyset\}$$
$$i(b) = \{b, a, b \wedge a', k, b \wedge k', \emptyset\}$$
$$i(c) = \{c, a, c \wedge a', k, k' \wedge c, \emptyset\}$$
$$i(b) \wedge i(c) = i(b) \cap i(c) = \{a, k, \emptyset\}$$
$$i(a') = \{a', z', z \wedge a', c', c \wedge a', b', b \wedge a', \emptyset\}$$
$$i(a') \wedge (i(b) \wedge i(c)) = \{\emptyset\}$$
$$i(a) \vee \big[i(a') \wedge (i(b) \wedge i(c))\big] = i(a) \neq i(b) \wedge i(c)$$

On voit donc que bien que $i(a) \subset i(b) \wedge i(c)$, on n'a pas $i(a) \longleftrightarrow i(b) \wedge i(c)$. Donc $\tilde{E}$ ne satisfait pas l'axiome AJ4. $\blacksquare$



$E = \{I, \emptyset, a, b, c, k, z, a', b', c', k', z', k \vee c', k' \wedge c, b' \vee k,$
$c \wedge a', b \wedge k', c' \vee a, z' \vee a, b \wedge a', b' \vee a, z \wedge a'\}$

Figure 3. Un ensemble ordonné orthocomplémenté et faiblement modulaire dont l'extension de Dedekind n'est pas faiblement modulaire.

CHAPITRE III

ETUDE DU SYSTEME DE PIRON

1. PREALABLES

Pendant la dernière décennie, C. Piron a proposé un nouveau système axiomatique. Son but déclaré était double : d'une part, il voulait que tous les axiomes aient une justification empirique, d'autre part, le système devrait fournir une description réaliste au sens défini par Einstein { 48} de la nature.

Afin de parvenir à ce but, l'auteur a été amené à une plus grande précision dans les définitions et à un raffinement des concepts. Son système est donc plus riche que celui de Jauch. Le concept primaire est ici le concept de "question", tandis que les "propositions" sont des classes d'équivalence de questions. C'est pourquoi nous appellerons souvent le nouveau système qp-s (système des questions-propositions).

Notre objectif dans ce chapitre est de montrer que, malgré ce plus grand raffinement, le système de Piron n'a pas atteint son but. En particulier, nous allons prouver à l'aide de 4 théorèmes que qp-s a des vices syntaxiques cachés qui enlèvent toute possibilité d'interprétation pour deux de ses trois axiomes. Le manque de support sémantique est si grave, que nous serons amenés à nous demander si le système possède une cohérence quelconque,ne cerait-ce qu' au niveau de la syntaxe. Nous montrerons que cette cohérence syntaxique est acquise ; qp-s est consistant au sens de la théorie abstraite des modèles. C'est donc seulement son aptitude de servir comme base d'une théorie physique qui est contestable.

2. LE FORMALISME

La plus grande précision des définitions du système de Piron nous permet de les transcrire ici de façon presque littérale, de {22} . Chaque fois qu'une définition est tirée d'un autre article, nous en ferons explicitement mention.

DP_1 (*Système physique*). Un système physique est une partie du monde réel, existant en espace-temps et extérieur au physicien.

DP_2 (*question*). Une question est toute expérience qui amène à une alternative dont les termes sont "oui" et "non".

DP_3 (*question opposée ou inverse*). Si α est une question, $\alpha^{\sim}$ est la question obtenue en permutant les termes de l'alternative.

DP_4 (*produit des questions*). Si $\{\alpha_j\}_{j \in J}$ est une famille de questions, $\prod_J \alpha_j$ est la question définie ainsi : on mesure une quelconque des α_j et on attribue à $\prod_J \alpha_j$ la réponse ainsi obtenue.

Règle R_1 (*opposée de la question produit*). Les définitions impliquent la règle suivante : $(\prod_J \alpha_j)^{\sim} = \prod_J \alpha_j^{\sim}$

DP_5 (*question triviale*). Il existe une question triviale I qui consiste à mesurer n'importe quoi (ou ne rien faire) et donner chaque fois la réponse "oui".

DP_6 (*question certaine ou vraie*). Lorsque le système a été préparé de façon telle que le physicien peut affirmer que dans l'éventualité d'une mesure correspondant à la question α le résultat sera "oui", nous dirons que la question α est "certaine" ou "vraie".

DP_7 (*relation de préordre entre les questions*). Si la question γ est vraie chaque fois que la question β est vraie, on dira que β est plus forte que γ et on le notera par $\beta < \gamma$. La relation $<$ est transitive.

DP_8 (*questions équivalentes*). Si l'on a $\beta < \gamma$ et $\gamma < \beta$, alors β et γ sont des questions équivalentes, ce qui sera noté par $\beta \simeq \gamma$.

DP_9 (*proposition*). La classe d'équivalence qui contient la question β est appelée proposition et elle est notée par b. L'ensemble de toutes les propositions pour un système physique est symbolisé par $\mathcal{L}$.

DP_{10} (*proposition vraie, Cf.* {20} *p.291*). La proposition b est vraie ssi la question β dont elle est la classe d'équivalence est vraie (dans le sens de DP_6).

DP_{11} (*relation d'ordre entre propositions*, {20} p. 291). Si l'on a $\forall \beta \in b, \forall \gamma \in c : \beta < \gamma$, alors la proposition

DP_1 (*Système physique*). Un système physique est une partie du monde réel, existant en espace-temps et extérieur au physicien.

DP_2 (*question*). Une question est toute expérience qui amène à une alternative dont les termes sont "oui" et "non".

DP_3 (*question opposée ou inverse*). Si α est une question, $\alpha^{\sim}$ est la question obtenue en permutant les termes de l'alternative.

DP_4 (*produit des questions*). Si $\{\alpha_j\}_{j \in J}$ est une famille de questions, $\prod_J \alpha_j$ est la question définie ainsi : on mesure une quelconque des α_j et on attribue à $\prod_J \alpha_j$ la réponse ainsi obtenue.

<u>Règle</u> R_1 (*opposée de la question produit*). Les définitions impliquent la règle suivante : $(\prod_J \alpha_j)^{\sim} = \prod_J \alpha_j^{\sim}$

DP_5 (*question triviale*). Il existe une question triviale I qui consiste à mesurer n'importe quoi (ou ne rien faire) et donner chaque fois la réponse "oui".

DP_6 (*question certaine ou vraie*). Lorsque le système a été préparé de façon telle que le physicien peut affirmer que dans l'éventualité d'une mesure correspondant à la question α le résultat sera "oui", nous dirons que la question α est "certaine" ou "vraie".

DP_7 (*relation de préordre entre les questions*). Si la question γ est vraie chaque fois que la question β est vraie, on dira que β est plus forte que γ et on le notera par $\beta < \gamma$. La relation $<$ est transitive.

DP_8 (*questions équivalentes*). Si l'on a $\beta < \gamma$ et $\gamma < \beta$, alors β et γ sont des questions équivalentes, ce qui sera noté par $\beta \simeq \gamma$.

DP_9 (*proposition*). La classe d'équivalence qui contient la question β est appelée proposition et elle est notée par b. L'ensemble de toutes les propositions pour un système physique est symbolisé par $\mathcal{L}$.

DP_{10} (*proposition vraie, Cf.* {20} *p.291*). La proposition b est vraie ssi la question β dont elle est la classe d'équivalence est vraie (dans le sens de DP_6).

DP_{11} (*relation d'ordre entre propositions*, {20} p. 291). Si l'on a $\forall \beta \in b, \forall \gamma \in c : \beta < \gamma$, alors la proposition b est plus forte que la proposition c, ce qui est noté par $b < c$.

DP$_{12}$ (*"produit" ou "conjonction" de propositions**). Pour toute famille de propositions $\{b_j\}_{j \in J}$, $\bigwedge_J b_j$ est la classe d'équivalence qui contient la question $\prod_J \beta_j$, où $\beta_j \in b_j$.

DP$_{13}$ (*"somme" ou "disjonction" de propositions**). Pour toute famille de propositions $\{b_j\}_{j \in J}$, $\bigvee_J b_j$ est le produit $\bigwedge_a x_a$ de toutes les propositions $x_a \in \mathcal{L}$ telles que $b_i < x_a$, $\forall i$.

<u>Théorème T$_1$</u> L'ensemble des propositions $\mathcal{L}$ est un treillis complet, i.e. il existe un plus grand minorant et un plus petit majorant (au sens de la théorie des treillis, {30}) pour toute famille de propositions. Le plus grand minorant et le plus petit majorant sont donnés, respectivement, par la conjonction et la disjonction.

DP$_{14}$ (*proposition absurde, proposition triviale*). Le théorème T$_1$ implique l'existence d'une proposition absurde $0 = \bigwedge_{b \in \mathcal{L}} b$. La classe d'équivalence de la question triviale I définit la proposition triviale I (même notation pour la question et la proposition triviales).

DP$_{15}$ (*proposition complémentaire pour b*). La proposition c est une proposition complémentaire pour b ssi $b \vee c = I$ et $b \wedge c = 0$.

DP$_{16}$ (*complément compatible de b*). La proposition c est un complément compatible $c = b'$ d'une proposition donnée b ssi elle est une proposition complémentaire pour b et si en plus il existe une question β telle que $\beta \in b$ et $\beta^{\sim} \in c$.

<u>Axiome C</u> (*existence d'un complément compatible*). Pour toute proposition $b \in \mathcal{L}$ il existe au moins un complément compatible.

<u>Axiome P</u> (*modularité faible*). Si $b < c$ sont des propositions dans $\mathcal{L}$ et b' est le complément compatible de b, et c' le complément compatible de c, alors le sous-treillis engendré par $\{b, c, b', c'\}$ est distributif.

L'axiome P implique l'unicité du complément compatible, et que $\mathcal{L}$ est orthocomplémenté :

$$\forall b \in \mathcal{L} \; : \; (b')' = b, \; b \vee b' = I, \; b \wedge b' = 0$$

$$\forall (b,c) \in \mathcal{L}^2 \; : \; b < c \Longrightarrow c' < b'.$$

* Dénommé ainsi par nous-mêmes.

DP_{17} (*proposition atomique*). Si $q \in \mathcal{L}$ est telle que

$$q \neq 0 \quad \text{et} \quad 0 < x < q \implies x = 0 \text{ ou } x = q,$$

alors q est appelée atomique.

<u>Axiome A</u> (*Atomicité, loi de couverture*) :

a) Si $b \in \mathcal{L}$, $b \neq 0$ alors il existe une proposition atomique q telle que $q < b$.

b) Si q est atomique et si $q \wedge b = 0$, alors

$$b < x < b \vee q \implies x = b \text{ ou } x = b \vee q.$$

DP_{18} (*système propositionnel*). Un treillis complet qui satisfait les axiomes C, P et A est un système propositionnel.

Les définitions DP_{1-18} et les Axiomes A, P, C forment le système proposé par Piron. Les définitions qui suivent nous seront utiles dans notre étude :

DP_{19} (*propositions orthogonales*). $b \in \mathcal{L}$ est orthogonale à $c \in \mathcal{L}$ (notation $b \perp c$) ssi $b < c'$.

DP_{20} (*mesure de première espèce*) : la question β est une mesure de première espèce si, chaque fois que la réponse est "oui", on peut affirmer que la proposition b définie par β est vraie immédiatement après la mesure ({22} , p. 68)

DP_{21} (*question idéale*). Une question β, appartenant à la proposition b, est appelée idéale, si toute proposition compatible avec b et qui est vraie avant une mesure de β, reste vraie après la mesure chaque fois que la réponse à β a été "oui".

DP_{22} (*succession de mesures*). Etant données deux questions idéales et de première espèce β et γ, nous noterons par $\gamma \circ \beta$ la question définie comme suit : on effectue la mesure β et, si le système n'a pas été détruit, on effectue γ immédiatement après. On attribue à $\gamma \circ \beta$ la réponse "oui" ssi les deux réponses ont été "oui".

DP_{23} (*propositions compatibles*). On appelle compatibles deux propositions b et c, si le treillis engendré par {b, c, b', c'} est distributif.

3. ETUDE CRITIQUE

3.1. *Remarques préliminaires*

Par rapport au système de Jauch, le qp-s a sans doute l'avantage de la précision. Aucune définition de concept n'a été laissée dans le vague, comme c'était le cas pour la relation d'ordre dans le système précédent. D'autre part, la grande réussite du qp-s vient sans doute du fait que l'axiome AJ2 devient ici un théorème : l'existence du plus grand minorant ne pose plus de problème.

Cependant, cette précision dans les définitions et cette réussite concernant la conjonction se payent par une plus grande complexité et, surtout, par l'introduction d'éléments très artificiels comme les expériences "produit" $\alpha \Pi \beta$, qui n'ont rien à faire avec la pratique expérimentale. Bien sûr, l'introduction de tels éléments n'est pas à *priori* contestable. Elle pourrait même sembler souhaitable puisqu'elle conduit à une déduction de la structure de treillis pour l'ensemble des propositions. Toutefois, nous allons montrer qu'en réalité elle ne fait que transposer le problème d'interprétation d'un axiome à l'autre.

3.2. *Objections d'autres auteurs*

Le qp-s a été déjà critiqué à deux reprises, par W. Ochs {49} et B. Mielnick { 50 }, mais ces critiques ne sont pas concluantes. W. Ochs a opposé deux objections à qp-s : (a) la relation de préordre $<$ entre les questions définie par Piron n'est pas identique à celle que l'on a habituellement introduite dans les approches logiques {22}. (b) les probabilités de vérité pour certaines questions ne sont pas données par la formule habituelle Tr(WE) de la Mécanique Quantique.

Les deux remarques d'Ochs sont correctes, mais ne sont pas significatives en tant que critiques du qp-s. Piron n'a jamais affirmé, par exemple, que la probabilité de vérité pour toutes les questions avait la forme Tr(WE), ({ 22}, p. 73). (b) est une propriété caractéristique du qp-s et de la Mécanique Quantique et ne peut avoir valeur d'objection. De même, Piron n'a jamais affirmé que sa relation de préordre était identique à celle de Jauch, Mackey, etc...

La critique de Mielnick se porte sur l'axiome C et elle est plus significative. Mais en fait elle ne fait qu'effleurer l'énorme problème que présente cet axiome. En outre, l'exposé de Mielnick souffre d'une grave erreur dont nous allons bientôt parler, ce qui enlève une grande partie de l'importance de sa critique.

3.3. *Examen du support physique des axiomes*

Le qp-s contient trois axiomes , C, P, et A. Nous allons centrer notre attention sur deux d'entre eux, i.e. C et A.

Axiome C. Cet axiome, qui affirme l'existence d'un"complément compatible" (Cf. DP_{16}) pour toute proposition, n'a jamais été sérieusement critiqué. Nous croyons que cette acceptation de la part des physiciens provient de deux raisons, sans doute reliées. En premier lieu, l'axiome correspondant du système de Jauch(i.e. AJ3) avait une justification physique très simple, ce qui a donné à croire que c'était aussi le cas dans qp-s. Et en second lieu, on a très souvent supposé que le complément compatible a' d'une proposition $a \in \mathcal{L}$ est tout simplement la classe $\{\alpha^{\sim}\}$ des "négations" $\alpha^{\sim}$ des questions α (voir, par exemple : Griechie et Gudder {37} p. 559, Mielnick { 50 } p. 120, et dans une première version du qp-s, Jauch et Piron {51} ,p. 170). Si cela était vrai, l'axiome C serait satisfait de façon si triviale, que la nécessité de son assertion comme un axiome serait contestable. En fait, cette supposition est erronée, et ce qui est contestable est l'existence même du complément compatible. Nous allons montrer notre affirmation par une succession de quatre théorèmes de puissance croissante. Nos démonstrations se feront entièrement à l'intérieur du système qp-s, dont nous utiliserons les axiomes.

Théorème 1. A l'intérieur du pq-s il existe au moins une proposition telle que son complément compatible n'est pas l'ensemble des "négations" $\alpha^{\sim}$ des questions α dont cette proposition est la classe d'équivalence.

Preuve. Soit $a \in \mathcal{L}$ une proposition quelconque. D'après l'axiome C, il existe $\alpha \in$ a telle que $\alpha^{\sim} \in$ a'. Par conséquent, on a pour cette question α :

$$\alpha \,\Pi\, \alpha^{\sim} \in a \wedge a' = 0$$

D'autre part, la règle R_1 donne :

$$(\alpha \,\Pi\, \alpha)^{\sim} = \alpha^{\sim} \,\Pi \alpha \in a' \wedge a = 0 \neq 0' = I.$$

ce qui montre que $0' \neq \{\beta^{\sim} : \beta \in 0\}$. ∎

Le contenu du théorème 1 peut être illustré par le diagramme "non-commutatif" suivant :

$$\exists a \in \mathcal{L}, \exists \alpha \quad : \begin{array}{ccc} \alpha & \in & a \\ \updownarrow & & \updownarrow \\ \alpha^{\sim} & \notin & a' \end{array}$$

L'exemple donné dans la démonstration du théorème 1 introduit la proposition particulière 0. Mais nous allons maintenant démontrer le même résultat pour <u>toute</u> proposition du qp-s :

<u>Théorème</u> 2. Pour toute $b \in \mathcal{L}$ différente de la proposition triviale, le complément compatible b' est différent de l'ensemble de négations de toutes les questions dont b est la classe d'équivalence.

<u>Preuve</u>. Pour toute $b \in \mathcal{L}$, $b \neq I$, soit c une proposition orthogonale à b mais différente de b'. Nous montrerons par construction qu'il existe une question $\alpha \in b$ telle que $\alpha^{\sim} \in c$. Puisque toute question appartient à une proposition et une seule, cela prouvera que $\alpha^{\sim} \notin b'$ et établira le théorème.

D'après l'axiome C, il existe des questions β et γ telles que $\beta \in b$, $\beta^{\sim} \in b'$, $\gamma \in c$, $\gamma^{\sim} \in c'$. Considérons la question $\alpha = \beta \,\Pi\, \gamma^{\sim}$. D'après la définition DP_{19} on a:$b < c'$, et donc $c < b'$. Par conséquent

$$\alpha = \beta \,\Pi \gamma^{\sim} \in b \wedge c' = b$$

tandis que

$$\alpha^{\sim} = (\beta \Pi \gamma^{\sim})^{\sim} = \beta^{\sim} \,\Pi\, \gamma \in b' \wedge c = c$$

On a donc $\gamma \in b$ et $\gamma^{\sim} \in c \neq b'$. ∎

Nous pouvons représenter graphiquement le contenu du théorème 2 comme suit :

$$\forall a \in \mathcal{L} \ , \ \exists \alpha : \begin{array}{ccc} \alpha & \in & a \\ \updownarrow & & \updownarrow \\ \alpha^{\sim} & \notin & a' \end{array}$$

On peut illustrer d'une autre façon le contenu du théorème 2, à l'aide d'un exemple : Considérons une particule dans un espace unidimensionnel représenté par la droite réelle. Soit β , γ deux questions définies ainsi : β est définie par un appareil capable de vérifier si la position de la particule appartient ou non à l'ensemble (0, 1), la réponse "oui" correspondant au cas où la particule est trouvée dans (0, 1) ; de manière analogue, γ est définie par un appareil capable de vérifier si la position de la particule appartient à l'ensemble (2,3). Notons que, selon la définition DP_3, $\beta^{\sim}$ et $\gamma^{\sim}$ correspondent respectivement au même appareil, mais les réponses "oui" et "non" sont permutées pour ces questions. Soient b, c, les propositions auxquelles appartiennent β et γ . Nous constatons que dans cet exemple précis, les propositions b et c possèdent un contenu sémantique : b, par exemple, est la proposition " la position de la particule appartient à l'intervalle (0,1)", qui peut être vérifiée ou falsifiée par l'appareil β . Par construction, nous avons $\beta \in b$, $\gamma \in c$, $\beta^{\sim} \in b'$, $\gamma^{\sim} \in c'$. En outre, $b < c'$. Par conséquent, selon la définition DP_{19}, b et c sont orthogonales.

Si l'on définit maintenant la question $\alpha = \beta \; \Pi \; \gamma^{\sim}$, alors comme nous l'avons démontré dans la preuve du théorème 2, on a : $\alpha \in b$, $\alpha^{\sim} \notin b'$. Nous sommes donc, du point de vue syntaxique, devant le fait curieux, que la question opposée à une question α donnée ne se trouve pas dans le complément de la proposition à laquelle α appartient. Mais venons maintenant à la sémantique. Supposons que l'on mesure $\alpha \in b$ et que l'on trouve la réponse "oui". En général, aucune conclusion ne peut en être tirée en ce qui concerne la vérité de la proposition b, à savoir " la position appartient à (0,1)", et cela en dépit du fait que $\alpha \in b$. En effet, d'après la définition DP_4, afin d'effectuer une mesure de α , on doit choisir l'une de deux questions β , $\gamma^{\sim}$ .Si le hasard nous conduit à choisir $\gamma^{\sim}$, alors la réponse "oui" pour α signifie, selon DP_3 et DP_4, que la réponse à γ est "non", i.e. que la particule a été trouvée dans $R \setminus (2,3)$. Cependant , on ne peut pas décider si b est vraie ou pas sur la base de cette information. En d'autres termes, $\alpha \in b$ n'implique pas que α peut mesurer b ! Cela met en évidence le fait que la caractéristique formelle exprimée par le théorème 2 implique des conséquences sémantiques bizarres dans l'interprétation du qp-s.

Nous avons donc montré que, contrairement à une opinion très répandue, le complément compatible d'une proposition ne peut pas être formé par la classe de toutes les *"négations" des questions qui appartiennent* à cette proposition. Toutefois, cela ne suffit pas pour nous amener à nier l'existence de a'. En effet, il pourrait y avoir une possibilité que pour chaque $a \in \mathcal{L}$ il existe une méthode de construction d'un complément compatible non vide a', même si a' ne contient pas toutes les "négations" $\alpha^{\sim}$des $\alpha \in a$. Nous allons cependant démontrer un théorème qui semble exclure cette possibilité aussi.

Considérons deux propositions distinctes b, $c \in \mathcal{L}$, $b \neq c$, $b = \{\beta_i\}_{i \in I}$, $c = \{\gamma_j\}_{j \in J}$. Considérons encore toutes les questions-produits $\beta \Pi \gamma$, $\beta \in b$, $\gamma \in c$. D'après la définition DP_{12}, la classe d'équivalence de $\beta \Pi \gamma$définit la proposition $b \wedge c$. Nous affirmons maintenant que le complément compatible ne contient aucune des "négations" $(\beta \Pi \gamma)^{\sim}$ des questions $\beta \Pi \gamma$à l'aide desquelles la proposition $b \wedge c$ peut être exclusivement définie.

Théorème 3. Soit dans $\mathcal{L}$ la proposition $b \wedge c$, où $(b,c) \in \mathcal{L}^2$, $a \neq b$, définie comme la classe d'équivalence de toutes les questions $\beta \Pi \gamma$, où $\beta \in b$, $\gamma \in c$.
Alors on a : $(\beta \Pi \gamma)^{\sim} \notin (b \wedge c)', \forall \beta \Pi \gamma$.

Preuve. Par réduction à l'absurde. Supposons qu'il existe au moins une paire de questions $\beta \in b$ et $\gamma \in c$ telles que $(\beta \Pi \gamma)^{\sim} \in (b \wedge c)'$. Soient d et e les propositions qui contiennent $\beta^{\sim}$ et $\gamma^{\sim}$, respectivement :

$\beta^{\sim} \in d$, $\gamma^{\sim} \in e$. Alors

$$(\beta \Pi \gamma)^{\sim} = \beta^{\sim} \Pi \gamma^{\sim} \in d \wedge e.$$

Puisque $(b \wedge c)' = b' \vee c'$, on devrait donc avoir

$$d \wedge e = b' \vee c'$$

ce qui entraîne

$$d > b', \; d > c', \; e > b', \; e > c'. \qquad (1)$$

L'inégalité $d > b'$ et l'axiome P impliquent que le sous-treillis engendré par $\{b, d, b', d'\}$ est distributif.
Par conséquent,

$$(b \wedge d) \vee b' = (b \vee b') \wedge (d \vee b') = I \wedge d = d \qquad (2)$$

Etant donné que $\beta \,\Pi\, \beta^{\sim}$ n'est jamais "certaine", on a

$$\beta \,\Pi\, \beta^{\sim} \in 0$$

mais aussi

$$\beta \,\Pi\, \beta^{\sim} \in b \wedge d$$

et donc $b \wedge d = 0$. Par conséquent,

$$(b \wedge d) \vee b' = 0 \vee b' = b' \qquad (3)$$

La comparaison des (2) , (3) donne

$$b' = d \qquad (4)$$

D'une manière analogue on peut démontrer

$$c' = e \qquad (5)$$

En utilisant (1) , (4) , et (5) on déduit que $b' = c'$, et donc $b = c$, contrairement à l'hypothèse faite que $b \neq c$. Par conséquent, l'hypothèse selon laquelle il existe au moins une paire de questions $\beta \in b$, $\gamma \in c$ telles que $(\beta \,\Pi\, \gamma)^{\sim} \in (b \wedge c)'$ est fausse. ■

Le contenu du théorème 3 peut être représenté comme suit :

$$\begin{array}{l} \forall\, (a,b) \in \mathcal{L} \times \mathcal{L}\;,\; a \neq b, \; : \\ \forall\; \alpha \,\Pi\, \beta \;\in\; a \wedge b \end{array} \qquad \begin{array}{ccc} \alpha\Pi\beta & \in & a \wedge b \\ \updownarrow & & \updownarrow \\ (\alpha\Pi\beta)^{\sim} & \notin & (a \wedge b)' \end{array}$$

Les conséquences sémantiques du théorème seront mieux illustrées par un exemple. Considérons un système microscopique S de masse <u>non nulle</u>*. Considérons encore deux

---

* Ce choix est dicté par le fait que Piron, pour illustrer son axiome, a toujours considéré un système classique ou alors un système quantique de masse nulle, pour lequel le treillis de propositions est fortement dégénéré, le produit de toute paire de propositions non triviales étant nul.

appareils $\mathcal{A}(Q)$ et $\mathcal{A}(P)$ qui peuvent être utilisés pour mesurer, au sens de la Théorie Quantique, l'observable position Q et l'observable quantité de mouvement P, respectivement. A l'intérieur du qp-s, pour tout système et tout appareil capable de mesurer quelque chose sur ce système, il existe des questions qui correspondent à ces mesures possibles. Pour le système S et l'appareil $\mathcal{A}(Q)$ en particulier, il existe à l'intérieur du qp-s des questions β , tandis que pour S et $\mathcal{A}(P)$ il existe des questions γ . En outre, chaque question définie dans qp-s une proposition. Soient donc b et c deux propositions définies par deux questions β_0 et γ_0. Puisque $[Q,P] \neq 0$, b et c ne sont pas compatibles. Soit enfin $b \wedge c$ leur conjonction, définie d'après la définition DP_{12}. L'axiome C affirme l'existence d'un complément compatible $(b \wedge c)'$ de $b \wedge c$ et donc, d'après DP_{16} , d'une question $\alpha \in b \wedge c$ telle que $\alpha^{\sim} \in (b \wedge c)'$. Or d'après le théorème 3, $(b \wedge c)'$ ne contient aucune des "négations" $(\beta \Pi \gamma)^{\sim}$ des questions $\beta \Pi \gamma$ à l'aide desquelles $b \wedge c$ a été définie. En d'autres termes, l'axiome affirme l'existence dans $b \wedge c$ de questions autres que celles de la forme $\beta \Pi \gamma$.

Examinons donc quelles peuvent être ces autres questions, et s'il est possible d'en produire en utilisant la syntaxe du qp-s. Le qp-s possède seulement deux algorithmes de construction syntaxique de questions : le produit (définition DP_4) et la succession dans le temps (déf. DP_{22}). Nous avons vu que le produit nous donne par définition des questions $\beta \Pi \gamma$ appartenant à $b \wedge c$, mais qui sont telles que $(\beta \Pi \gamma)^{\sim} \notin (b \wedge c)'$. En ce qui concerne la succession temporelle, on peut montrer que <u>si</u> b et c contiennent des questions β et γ idéales et de première espèce, alors on a en effet $\gamma \circ \beta \in$ $\in b \wedge c$ ({22} ,p. 72). Or on peut aussi montrer que dans ce cas $(\gamma \circ \beta)^{\sim} \notin (b \wedge c)'$, puisque b et c ont été choisies non-compatibles (la démonstration est donnée en appendice, voir corollaire du théorème 4).

Récapitulons : l'axiome C affirme l'existence d'un complément compatible pour toute proposition dans $\mathcal{L}$. En particulier il affirme, d'après la définition DP_{16}, qu'il existe une question $\alpha \in b \wedge c$ telle que $\alpha^{\sim} \in (b \wedge c)'$. D'autre part, les seules questions appartenant à $b \wedge c$ que l'on peut construire syntaxiquement ont la forme $\beta\Pi\gamma$ et $\gamma o \beta$. Or, comme nous l'avons montré, les questions opposées $(\beta\Pi\gamma)^{\sim}$ et $(\gamma o \beta)^{\sim}$ n'appartiennent jamais à $(b \wedge c)'$. L'axiome C affirme donc subrepticement l'existence dans $b \wedge c$ de questions (i.e. expériences oui-non) que l'on ne peut pas construire syntaxiquement, ni spécifier sémantiquement.

Nous sommes en face exactement du même problème que celui que nous avons rencontré dans le système de Jauch. La seule différence est qu'ici le problème intervient à travers l'axiome qui introduit la négation lorsqu'on applique à des propositions de la forme $b \wedge c$, b et c n'étant pas compatibles. Ainsi, dans les deux systèmes, on affirme l'existence de certaines expériences oui-non qui définiraient la conjonction $b \wedge c$ (axiome AJ2 de Jauch) ou son complément $(b \wedge c)'$ (axiome C de Piron) sans que cette affirmation se base sur une justification empirique quelconque.

Le système qp-s est donc un système formel dont la prétention est d'engendrer par son interprétation, la <u>théorie physique</u> dénommée Mécanique Quantique mais dont le <u>premier</u> des axiomes fondamentaux n'a déjà aucune signification physique. Bien sûr, cette analyse de l'axiome C est déjà concluante en ce qui concerne la signifiance physique du qp-s. Nous complétons quand même notre étude sémantique par une brève analyse de l'autre axiome.

<u>Axiome A</u>. A notre connaissance le seul contenu sémantique à avoir été spécifié pour cet axiome, à l'intérieur du qp-s, se trouve dans un article commun de Jauch et de Piron {19} . Les auteurs ont pu trouver une justification de l'axiome en faisant l'hypothèse qu'il existe, à l'intérieur de <u>toute</u> proposition b, une question idéale et de première espèce.

Considérons de nouveau la proposition $b \wedge c$, où b et c sont les propositions non compatibles que nous avons définies précédemment dans l'illustration du théorème 2. Comme nous l'avons déjà dit, les questions qui appartiennent à $b \wedge c$ ont la forme $\beta\Pi\gamma$ ou $\gamma o \beta$. Aucun exemple de question qui n'a pas cette forme et qui appartient à $b \wedge c$ n'a jamais été trouvé. Mais Piron lui-même a démontré les théorèmes suivants ({22} , p. 72) :

(1) Si $\gamma \in c, \beta \in b$ et $b \neq c$, alors la question $\beta\Pi\gamma$ n'est jamais une mesure idéale et de première espèce.

(2) La question $\gamma o \beta$ est une mesure idéale et de première espèce ssi les propositions b et c corresondant à β et γ sont compatibles.

Par conséquent, on ne peut pas accepter l'hypothèse qu'il existe des questions idéales et de première espèce dans toute proposition, et en particulier pour les propositions $b \wedge c$ pour b et c non compatibles. Et avec elle, on ne peut pas accepter la seule justification connue de l'axiome A*.

3.4. *Consistence syntaxique du qp-s*.

Le formalisme du qp-s est construit sur deux niveaux qui sont en connexion, le niveau des questions et le niveau des propositions. La partie du formalisme qui concerne seulement les propositions conduit, comme dans le système de Jauch, à la structure d'un treillis complet, orthocomplémenté, faiblement modulaire et atomique. De telles structures ont été plusieurs fois étudiées et leur consistence syntaxique est garantie. Mais la structure globale introduite par les deux niveaux, celui des propositions et celui des questions, est une structure nouvelle qui inclut plus d'affirmations que celles qui sont en usage dans la théorie des treillis. Une telle structure formelle n'a jamais été étudiée. En particulier, sa consistence syntaxique n'est pas du tout évidente.

Il semblerait même, qu'au contraire, la consistence d'une telle structure serait surprenante, et cela pour la raison suivante :
Si b et c sont deux propositions <u>non compatibles</u> (au sens tout à fait formel de la définition DP_{23}) alors, comme nous l'avons déjà remarqué, les seules questions que l'on peut spécifier à l'intérieur de $b \wedge c$ ont la forme $\beta\Pi\gamma$ et $\gamma o \beta$, et nous avons démontré que $(\beta\Pi\gamma)^{\sim} \notin (b \wedge c)'$, $(\gamma o \beta)^{\sim} \notin (b \wedge c)'$. D'autre part, l'axiome C affirme qu'il existe une question $\alpha \in b \wedge c$ telle que $\alpha^{\sim} \notin (b \wedge c)'$. On est donc devant une apparente contradiction, qui nous porte à poser deux questions :

---

* Bien sûr, il y a une différence de qualité entre nos critiques de deux axiomes ; nous avons critiqué <u>une interprétation</u> de l'axiome A, mais dans le cas de l'axiome C nous avons critiqué l'axiome <u>lui-même</u>.

a) Le système qp-s est-il syntaxiquement inconsistant ? En d'autres termes, est-il posiible, en utilisant ses axiomes et les relations entre les éléments de son langage d'arriver à démontrer une absurdité ?

b) Dans le cas où le qp-s serait consistant, serait-il possible de démontrer qu'il n'existe pas de propositions incompatibles ?

Il y a aussi d'autres raisons qui renforcent l'impression que qp-s pourrait être inconsistant, comme par exemple la curieuse règle pour la "négation" d'une question-produit $(\prod_J \alpha_j)^{\sim} = \prod_J \alpha_j^{\sim}$ qui semble contredire la loi de De Morgan. Toutefois, malgré les apparences, la réponse aux questions (a) et (b) est négative. Nous montrerons en effet que le système qp-s est consistant, et que sa syntaxe n'exclut pas l'existence de propositions incompatibles. En fait, la "contradiction" que nous avons mise en évidence entre l'axiome C et les deux théorèmes 3 et 4, ne se situe pas au niveau de la logique formelle ; il n'est pas <u>logiquement</u> exclu qu'il y ait dans $b \wedge c$ des questions autres que $\beta \Pi \gamma$ et $\gamma \circ \beta$. C'est seulement notre expérience physique qui nous dit le contraire.

Pour démontrer la consistence d'un système formel il suffit de produire un modèle (Cf. {34,35} , théorie des modèles). Ce concept se base sur celui de "réalisation" d'un langage. Par définition, une réalisation du langage d'un système formel consiste en la donnée d'un ensemble non vide quelconque D, et de certaines relations et opérations sur D, qui sont considérées comme interprétant les symboles fonctionnels et les prédicats du système. On dit qu'une réalisation est un modèle pour le système formel, si elle rend valides ses axiomes.

Dans le cas présent, un modèle du qp-s consisterait en la donnée d'un ensemble Q muni d'une relation de préordre et de deux opérations :

$$\alpha \in Q \longrightarrow \alpha^{\sim} \in Q$$

$$\{\alpha_j\}_{j \in J} \in P(Q) \longrightarrow \prod_J \alpha_j \in Q$$

P(Q) étant l'ensemble des sous-ensembles de Q, qui satisfont les conditions suivantes :

$$(A) \quad (\alpha^{\sim})^{\sim} = \alpha$$

$$(B) \quad (\prod_J \alpha_j)^{\sim} = \prod_J \alpha_j^{\sim}$$

(C) La relation de préordre définit, par passage aux classes d'équivalence, un treillis complet et atomique $\mathcal{L}$, satisfaisant à la loi de couverture.

(D) $\mathcal{L}$ est muni d'une orthocomplémentation $a \longrightarrow a'$ telle que

$$\forall a \in \mathcal{L}, \exists \alpha \in a : \alpha^{\sim} \in a'$$

et qui fait de $\mathcal{L}$ un treillis faiblement modulaire.

(E) Pour toute famille d'éléments $\{\alpha_j\}_{j \in J}$ et pour tout choix $\alpha_j \in a_j$, on a :

$$\prod_J \alpha_j \in \bigwedge_J a_j.$$

On peut maintenant établir le théorème suivant , dont la démonstration est donnée en appendice, et qui donne la réponse aux questions (a) et (b).

<u>Théorème 5</u>. *Le qp-s possède un modèle. De plus on peut choisir ce modèle de façon que le treillis $\mathcal{L}$ défini par la condition (C) soit non commutatif.*

Il s'ensuit donc que le qp-s est formellement consistant.

<u>4. CONCLUSION</u>

Nous avons fait une étude complète du système de Piron, des points de vue syntaxique et sémantique. Nous avons démontré la consistence syntaxique du système. Mais d'autre part, nous avons aussi démontré à l'aide des théorèmes 1-4 que l'aptitude de ce système comme base d'une théorie physique est contestable. L'un de ses axiomes affirme l'existence d'éléments d'un certain type pour lesquels aucune méthode de construction n'est spécifiée à l'intérieur du système, et aucune définition sémantique ne peut être trouvée. De plus, il a été montré que ce qu' affirme cet axiome est exactement de nature identique à l'axiome AJ2 de Jauch. Malgré la plus grande sophistication du système de Piron, le problème posé par les propositions du type $b \wedge c$ n'a pas été résolu.

Par rapport au système de Jauch, la plus grande nouveauté a été l'introduction dans le formalisme d'éléments artificiels du type $\beta \Pi \gamma$ qui ont servi à démontrer l'existence de $b \wedge c$. Nous avons montré que cela ne constitue pas un avantage, puisque ce qui a été gagné pour la conjonction, a été perdu pour la négation. Il constitue même un inconvénient. En effet, l'exemple qui a servi à illustrer le contenu du théorème 2 montre bien que l'introduction de $\beta \Pi \gamma$ conduit à des bizarreries qui nuisent à la capacité descriptive du système et provoquent une distorsion considérable du langage. Nous avons vu ainsi que le fait qu'une expérience oui-non β appartient à la proposition b n'implique pas que β peut mesurer b !

Il est intéressant de remarquer que notre étude du qp-s est faite de l'intérieur de celui-ci plutôt qu'en le comparant constamment à la Théorie Quantique. Nous avons seulement examiné si le qp-s pourrait fournir, par interprétation, une théorie physique quelconque, et nous avons trouvé que cela est loin d'être établi. Dès lors, une comparaison des résultats concrets du qp-s à ceux de la Mécanique Quantique serait superflue.

De notre analyse des systèmes de Jauch et de Piron, il découle clairement que le but de l'approche logique i. e. la déduction du formalisme Hilbértien de la Mécanique Quantique à partir d'un "calcul propositionnel" fondé sur des considérations empiriques, n'est pas atteint. Reste que l'apport de Von Neumann en ce domaine est certain : le formalisme Hilbértien implique une certaine structure de l'ensemble des propositions. Nous en ferons usage dans la deuxième partie de la thèse, ainsi que de l'idée de représenter les états par des mesures de probabilité sur l'ensemble des propositions.

APPENDICE AU CHAPITRE III

Théorème 4. Soient b et c deux propositions dans $\mathcal{L}$ et $\beta \in b$, $\gamma \in c$ deux questions idéales et de première espèce. Alors $(\gamma \circ \beta)^{\sim} \in b' \vee (b \wedge c')$

Preuve. Chaque $a \in \mathcal{L}$ est l'union de toutes les propositions atomiques q telles que $q < a$ ({18} ,p. 458). Pour démontrer donc le théorème, il suffit de montrer que si $(\gamma \circ \beta)^{\sim} \in d$ alors pour toute proposition atomique q on a :

$$q < d \Longleftrightarrow q < b' \vee (b \wedge c'). \qquad (6)$$

Montrons que

$$q < b' \vee (b \wedge c') \Longleftrightarrow (q \vee b') \wedge b < c'. \qquad (7)$$

On a évidemment :

$$q < b' \vee (b \wedge c') \Longrightarrow (q \vee b') \wedge b < (b' \vee (b \wedge c')) \wedge b \qquad (8)$$

D'après l'axiome P on a $b \leftrightarrow (b \wedge c')$. Donc (Cf. {22} , théorème 2.21) :

$$(b' \vee (b \wedge c')) \wedge b = (b' \wedge b) \vee (b \wedge c') = b \wedge c' < c' \qquad (9)$$

En combinant (8) avec (9) on trouve l'implication $\Longrightarrow$ de (7). La démonstration de l'implication $\Longleftarrow$ est analogue. La relation (7) est donc vraie, et montre que pour achever la démonstration du théorème il suffit de montrer que

$$q < d \Longleftrightarrow (q \vee b') \wedge b < c' \qquad (10)$$

pour toute proposition atomique q.

Montrons l'implication $\Longrightarrow$. Soit donc $q < d$, q atomique. On peut supposer, sans nuire à la généralité, que $q \wedge b' = 0$ (Si $q \wedge b' \neq 0$, alors $q < b'$ puisque q est atomique et donc $(q \vee b') \wedge b = 0 < c'$). Supposons que le système ait été préparé de façon que q soit "vraie" (au sens de DP_{10}). Alors $q < d$ implique que d est vraie, i.e. $(\gamma \circ \beta)^{\sim}$ est vraie. D'autre part b' n'est pas vraie, puisque $q \wedge b' = 0$. Puisque

β est idéale et de première espèce, on a $\beta^{\sim} \in b'$ ({22} , p. 71). Donc, si l'on mesure β, la réponse "oui" <u>est</u> possible. Dans une telle éventualité, les propositions b, $q \vee b'$, c' seront vraies après la mesure. En effet : b sera vraie car β est de première espèce ; $b' \vee q$ était vraie avant la mesure (puisque q était vraie et $q < q \vee b'$) et donc elle restera vraie puisque $b' \vee q \longleftrightarrow b$ et β est idéale ; enfin c' sera vraie car d'une part $\gamma^{\sim} \in c'$ (γ étant idéale et de première espèce, Cf. {22} p. 71) et d'autre part la réponse à $\gamma^{\sim}$ sera certainement "oui" (car <u>avant</u> la mesure β , la proposition $(\gamma \circ \beta)^{\sim}$ était vraie). Donc $[(q \vee b') \wedge b] \wedge c'$ sera vraie après la mesure, ce qui implique $[(q \vee b') \wedge b] \wedge c' \neq 0$. Or $(q \vee b') \wedge b$ est atomique ({19} ,p. 848). Donc $(q \vee b') \wedge b < c'$, ce qui montre l'implication $\Longrightarrow$ dans la relation (10).

Montrons enfin l'implication $\Longleftarrow$. Soit $(q \vee b') \wedge b < c'$. Il faut montrer que si q est vraie, alors $(\gamma \circ \beta)^{\sim}$ est vraie (c.à d. la réponse à $\gamma \circ \beta$ est certainement "non"). Supposons donc que q est vraie et que l'on mesure $\gamma \circ \beta$. Si la réponse à β est non, alors c'est aussi le cas pour $\gamma \circ \beta$. Si la réponse à β est oui, alors après la mesure de β , $(q \vee b') \wedge b$ est vraie (même démonstration que celle du paragraphe précédent) et donc c' est vraie. Il s'ensuit que la réponse à γ est "non". Par conséquent on aura pour $(\gamma \circ \beta)^{\sim}$ "non" dans tous les cas, ce qui achève la démonstration de la relation (10) et du théorème. ∎

<u>Corollaire 1</u>. *Soient b,c deux propositions non compatibles, et $\beta \in b$, $\gamma \in c$ deux questions idéales et de première espèce. Alors on a toujours $(\gamma \circ \beta)^{\sim} \notin (b \wedge c)'$.*

<u>Preuve</u>. D'après le théorème 4 on a $(\gamma \circ \beta)^{\sim} \in b' \vee (b \wedge c')$. Or $b' \vee (b \wedge c') = (b \wedge c)'$ ssi b et c sont compatibles ({22} théorème 2.19). ∎

<u>Théorème 5.</u> *Le qp-s possède un modèle. En particulier, il possède un modèle dans lequel $\mathcal{L}$ est isomorphe à l'ensemble des sous-espaces fermés d'un espace de Hilbert.*

<u>Preuve</u>. Soit H un espace Hilbértien séparable et soit L l'ensemble de ses sous-espaces fermés. Il est bien connu (Cf. chapitre II) que L est un treillis complet, orthocomplémenté, atomique, faiblement modulaire et satisfaisant à

la loi de couverture. Nous noterons la relation d'ordre dans L par $\sqsubseteq$, l'orthocomplémentation par une prime, et nous utiliserons $\sqcap$ pour le plus grand minorant. Soit Q l'ensemble des sous-ensembles de L,

$$Q := P(L) \tag{11}$$

Pour toute $\gamma \in Q$ nous définissons $A(\gamma) \in L$ par la relation

$$A(\gamma) := \bigsqcap_{x \in \gamma} x. \tag{12}$$

Nous définissons encore une relation de préordre sur Q :

$$\gamma < \delta \Longleftrightarrow A(\gamma) \sqsubseteq A(\delta) \tag{13}$$

et deux applications $\alpha \longrightarrow \alpha^{\sim}$ et $\{\alpha_j\}_{j \in J} \longrightarrow \prod_J \alpha_j$

$$\alpha^{\sim} := \{x' : x \in \alpha\} \tag{14}$$

$$\prod_J \alpha_j := \bigcup_J \alpha_j \tag{15}$$

où $\bigcup$ est l'union ensembliste. Les définitions (11) et (13), (14), (15) forment une réalisation de la théorie. Nous allons maintenant démontrer que cette réalisation est un modèle, i.e. que les conditions (A) $\longrightarrow$ (E) sont remplies (section III.3.4).

a) La condition (A) est une conséquence immédiate de (14).

b) La condition (B) peut être démontrée facilement :

$$(\prod_J \alpha_j)^{\sim} \overset{(14)}{=\!=} \{x' : x \in \prod_J \alpha_j\} \overset{(15)}{=\!=} \{x' : x \in \bigcup_J \alpha_j\}$$

$$\bigcup_J \{x' : x \in \alpha_j\} \overset{(14)}{=\!=} \bigcup_J \alpha_j^{\sim} \overset{(15)}{=\!=} \prod_J \alpha_j^{\sim}$$

c) Afin de démontrer la condition (C), nous procédons comme suit : tout d'abord, la relation < est réflexive et transitive, donc une relation de préordre. Par conséquent, la relation $\simeq$ définie sur Q par

$$\gamma \simeq \delta \Longleftrightarrow \gamma < \delta \quad \text{et} \quad \delta < \gamma \tag{16}$$

est une relation d'équivalence. Soit $\mathcal{L}$ l'ensemble des classes d'équivalence d'éléments de Q et $\subsetneq$ la relation d'ordre sur $\mathcal{L}$ définie par

$$a \subsetneq b \Leftrightarrow (\alpha < \beta \ , \quad \forall \ \alpha \in a, \forall \beta \in b) \tag{17}$$

Les relations (12), (13) et (16) montrent que $\gamma \simeq \delta$ ssi $A(\gamma) = A(\delta)$. En d'autres termes, $A(\gamma) = A(\delta)$ ssi γ et δ appartiennent à la même classe a. On peut donc définir

$$\forall a \in \mathcal{L} : \quad A(a) := A(\alpha) \ , \quad \alpha \in a \tag{18}$$

L'application $\mathcal{L} \xrightarrow{A} L$ est biunivoque :

$$A(a) = A(b) \overset{(18)}{\Longrightarrow} A(\alpha) = A(\beta), \forall \alpha \in a, \forall \beta \in b \Longrightarrow$$

$$\Longrightarrow \alpha \simeq \beta \Longrightarrow a = b.$$

En outre, si $x \in L$, alors $\{x\} \in Q$ et nous avons en vertue de la relation(12),

$$A(\{x\}) = x.$$

Si a est la classe d'équivalence qui contient $\{x\}$, alors $A(a) = A(\{x\}) = x$. Donc

$$\forall x \in L, \exists a \in \mathcal{L} : A(a) = x$$

i.e. A est aussi surjective. Enfin, la relation (17) implique

$$A(a) \subset A(b) \Longrightarrow A(\alpha) \subset A(\beta), \forall \alpha \in a, \forall \beta \in b \Rightarrow \alpha < \beta \Longrightarrow a \subsetneq b$$

i.e. l'application A est une bijection qui conserve la relation d'ordre entre $\mathcal{L}$ et L (un isomorphisme). Puisque L est un treillis complet, atomique et obéissant à la loi de couverture, $\mathcal{L}$ possède aussi cette structure. De plus , nous pouvons définir une orthocomplémentation sur $\mathcal{L}$ par la relation :

$$\forall a \in \mathcal{L} : \quad a' := A^{-1} \left[(A(a))' \right] . \tag{19}$$

On peut facilement vérifier que $\mathcal{L}$ devient ainsi orthocomplémenté et faiblement modulaire. La condition (C) est donc satisfaite.

d) Pour démontrer la condition (D) il suffit de montrer que

$$\forall\, a \in \mathcal{L},\ \exists \alpha \in a : \alpha^{\sim} \in a'$$

Notons d'abord que pour tout $a \in \mathcal{L}$, les relations suivantes sont valables :

$$A(a') \overset{(19)}{=\!=} A\left[A^{-1}\,(A(a))'\right] = \left[A(a)\right]' \tag{20}$$

$$A(a) \in L \Longrightarrow \{A(a)\} \in Q \overset{(12)}{\Longrightarrow} A(\ \{A(a)\}\ ) = A(a) \Longrightarrow$$
$$\Longrightarrow \{A(a)\} \in a \tag{21}$$

$$\{A(a)\}^{\sim} \overset{(14)}{=\!=} \{(A(a))'\} \overset{(20)}{=\!=} \{A(a')\} \tag{22}$$

Puisque (21) est vraie pour $\forall\, a \in \mathcal{L}$, on a

$$\{A(a')\} \in a' \tag{23}$$

et les relations (22) et (23) impliquent

$$\{A(a)\}^{\sim} \in a' \tag{24}$$

Enfin, (21) et (24) montrent que pour tout $a \in \mathcal{L}$ il existe au moins un $\alpha \in Q$, à savoir $\alpha = \{A(a)\}$, tel que $\alpha^{\sim} \in a'$: la condition (D) est satisfaite.

e) En ce qui concerne la condition (E), pour toute famille $\{a_j\}_{j \in J}$ d'éléments de $\mathcal{L}$ et tout $\alpha_j \in a_j$ on trouve, puisque A est un isomorphisme :

$$A\,(\bigwedge_J a_j) = \sqcap_J A(a_j) \tag{25}$$

et

$$A\,(\prod_J \alpha_j) \overset{(15)}{=\!=} A\,(\bigcup_J \alpha_j) = \sqcap_{x \in \bigcup_J \alpha_j} x = \sqcap_J \sqcap_{x \in a_j} x = \tag{26}$$

$$\sqcap_J A(\alpha_j) = \sqcap_J A(a_j).$$

Les deux égalités (25) et (26) donnent

$$A\,(\bigwedge_J a_j) = A\,(\prod_J \alpha_j)$$

et donc $\prod_{J} \alpha_j \in \bigwedge_{J} a_j$, ce qui montre que (E) est satisfaite, et achève la démonstration. ■

DEUXIEME PARTIE

CONTRIBUTIONS A L'ETUDE DU PROBLEME DE L'INFERENCE STATISTIQUE EN PHYSIQUE CLASSIQUE ET EN MEQUANIQUE QUANTIQUE.

INTRODUCTION

Dans cette deuxième partie nous présentons certains travaux portant sur le problème de l'inférence statistique, en théorie classique des probabilités et en mécanique quantique. Nous commençons donc par un bref exposé du problème.

Selon la formalisation de la théorie des probabilités introduite par Kolmogorov, à chaque expérience aléatoire correspond un "espace de probabilité" $[S,\mathcal{B},\rho]$, où S est l'ensemble de tous les évènements élémentaires, $\mathcal{B}$ est une σ-algèbre de sous-ensembles de S et ρ une mesure de probabilité sur $\mathcal{B}$ telle que $\rho(S) = 1$. Le contenu physique de la mesure $\rho(A)$ pour $A \in \mathcal{B}$ est donné par l'interprétation fréquentielle* : Si n(A) est le nombre de fois qu'apparaît un évènement élémentaire dans A lors de N répétitions de l'expérience, alors $\rho(A)$ est la limite de la fréquence relative $\frac{n(A)}{N}$ lorsque N s'accroît vers l'infini.

La mesure ρ qui donne les limites des fréquences relatives des divers évènements n'est pas toujours connue. Plus souvent, on possède sur l'expérience aléatoire quelques informations partielles, qui ne sont pas suffisantes pour fixer ρ. Dans ce cas, on essaie d'utiliser l'information disponible pour déterminer une mesure de probabilité ρ' qui est censée représenter "le mieux" cette information. La mesure ρ' est décrite aussi comme la "plus probable" parmi toutes les mesures de probabilité possibles, ou celle qui "minimise le risque d'erreur". La recherche de cette mesure ρ' constitue l'objet du problème de l'inférence statistique.

---

* Cette interprétation rencontre plusieurs obstacles, mathématiques et conceptuels {52}. Cependant, nous ne nous interesserons pas à ces questions ici.

Ce problème se rencontre sous divers aspects dans plusieurs branches de la physique. Ainsi, en Mécanique Statistique Classique, dans laquelle l'état d'un système est représenté par une mesure de probabilité sur l'espace des phases, le problème de l'inférence statistique consiste à estimer l'état sur la base de l'information disponible. Le même problème se pose en mécanique quantique, dans laquelle l'état peut être représenté, via le théorème de Gleason {32,53} par une mesure de probabilité sur l'ensemble de sous-espaces fermés d'un espace de Hilbert.

Les méthodes proposées pour la solution du problème de l'inférence statistique varient en fonction de la nature de l'information disponible. Elles se fondent en général sur des principes, plus ou moins justifiés intuitivement : principes de Bayes, de Jaynes, de Kullback, principe de maximum de vraisemblance, de moindres carrés {54,55} etc. Parmi tous ces principes, les plus appropriés pour les besoins de la Mécanique Statistique Classique sont ceux de Jaynes et de Kullback. Pour la Mécanique Quantique, le problème reste encore ouvert. (Nous en parlerons plus en détail au chapitre V).

Dans cette deuxième partie de la thèse, nous apportons principalement deux contributions à l'étude du problème de l'inférence statistique*:

a) Dans le chapitre suivant, nous élucidons, à l'aide de certains théorèmes que nous établissons, la signification des principes de Jaynes et de Kullback. Nous démontrons que le principe de Jaynes est une <u>conséquence</u> du principe bien connu d'équiprobabilités de Laplace { 4 }.

b) Dans le chapitre V, nous abordons le problème de l'inférence statistique en Mécanique Quantique. Nous définissons pour cela une *distance* entre les états. Ce concept est ensuite utilisé pour résoudre l'analogue quantique du problème étudié par Kullback dans le cas classique. En particulier, notre méthode conduit à la *déduction*, dans un cadre précis, de la proposition connue sous la dénomination de "postulat de projection" { 5 }.

---

* Il s'agit de travaux distincts, qui sont en cours de publication, et dont le premier concerne la version classique du problème, et le second sa version quantique.

Enfin, la thèse est complétée par deux annexes mathématiques qui concernent les mélanges d'états non orthogonaux et la "distance" de V. Cantoni. Ces annexes se trouvent en relation avec le contenu du chapitre V, mais elles consituent aussi des contributions originales qui apportent la solution à deux problèmes précis { 6., 7 }.

CHAPITRE IV

LE PRINCIPE D'ENTROPIE INFORMATIONNELLE MAXIMALE COMME UNE CONSEQUENCE DU PRINCIPE DE LAPLACE.

I. LE PRINCIPE DE JAYNES EN MECANIQUE CLASSIQUE.

Depuis sa proposition en 1957 par Jaynes le principe d'entropie maximale (PEM) a été un sujet de controverses. Un rapport exhaustif du "débat" entre les défenseurs(les plus nombreux)et les adversaires du PEM a été donné par Cyranski. {56}. Dans le chapitre présent nous prenons place parmi les défenseurs et nous allons développer un argument simple et quantitatif qui élucide la signification du PEM.

Comme nous l'avons déjà dit, ce principe permet de déterminer une distribution de probabilité "subjective" sur la base d'une information insuffisante. Il affirme que la distribution de probabilité qui représente le mieux notre connaissance est celle qui maximise l'entropie en tenant compte de l'information disponible (problème du maximum sous contrainte). L'objection la plus courante de ses adversaires est basée sur l'application du PEM à des situations délicates, ce qui conduit parfois à des résultats qui semblent paradoxaux. Des réponses à de telles objections {57} existent {56}, mais nous ne nous en préoccuperons pas: car nous nous interesserons exclusivement au cas simple auquel le PEM est appliqué le plus souvent, i.e. lorsque l'information disponible consiste en la donnée des valeurs moyennes de certaines variables aléatoires.

Le problème peut être formulé explicitement ainsi: Soit $\Omega=\{e_1,..,e_k\}$ un ensemble fini d'évènements élémentaires qui s'excluent mutuellement. Les e_i peuvent être, par exemple, les résultats possibles d'une expérience. Supposons que l'apparition des e_i soit régie par une loi de probabilité μ inconnue, et que nous ne connaissions que les valeurs moyennes $\hat{f}_\ell$ de certaines variables aléatoires $f_\ell(e_i)$,$\ell=1,...m$ définies sur Ω

En général, la connaissance des $\hat{f}_\ell$ ne fixe pas la probabilité $\mu(e_i)= \mu_i$, car il peut y avoir une infinité de mesures de probabilités différentes $p(e_i)=p_i$ donnant les mêmes valeurs moyennes aux $f_\ell(e_i)$, de façon à avoir

$$\sum_i p_i \; f_\ell(e_i) = \hat{f}_\ell, \quad \ell=1,2,...m. \tag{1}$$

Des situations de ce genre se rencontrent souvent en Mécanique Statistique, où on peut connaître la valeur moyenne de l'énergie sans connaître la distribution de probabilité sur l'espace des phases.

Puisque la probabilité véritable μ ne peut pas être déterminée sur la base de notre connaissance des $\hat{f}_\ell$, Jaynes a posé le problème de définir un choix, parmi toutes les mesures de probabilité satisfaisant à (1), choix posé comme représentant "le mieux" l'information disponible, en ce sens qu'il évite au maximum de favoriser un évènement e_i plutôt qu'un autre.

Un cas spécial de ce problème se rencontre souvent dans la pratique. On ne connaît que l'ensemble des résultats possibles e_i de l'expérience. On attribue alors souvent à tous les e_i des probabilités égales $p_i=\frac{1}{k}$. Ce choix est suggéré fortement par l'intuition, puisque tout autre attitude favoriserait certains évènements, sans que cela soit dicté par l'information disponible. Laplace a élevé ce choix au rang d'un postulat qui porte son nom.

Pour le cas général dans lequel l'information disponible est donnée par la relation (1), Jaynes a proposé le PEM comme une généralisation du principe de Laplace: la mesure de probabilité qui représente le mieux cette information est, selon le PEM, celle qui maximise l'entropie informationnelle de Shannon $-\sum p_i \log p_i$ sous les contraintes (1). Jaynes a avancé certains arguments afin de justifier ce choix {58 - 61} . Malheureusement ces arguments ont tous un caractère qualitatif et n'ont pas convaincu certains physiciens. Dans ce qui suit nous proposons un argument nouveau et quantitatif en faveur du PEM. Nous montrerons que si on accepte le principe de Laplace, alors le PEM en découle comme un cas particulier plutôt qu'une généralisation. Nous commençons la discussion sur un niveau heuristique qui fera transparaître plus facilement l'argument physique.

2. CONSIDERATIONS HEURISTIQUES.

Supposons que l'on répète l'expérience n fois. Le résultat de cette opération sera une suite $e_{j_1} e_{j_2} \ldots e_{j_N} \in \Omega^N$, $(j = 1,2,\ldots k)$. Puisque nous savons seulement que les valeurs moyennes des $f_\ell(e_i)$ sont $\hat{f}_\ell$, nous ignorons quelle séquence sera effectivement réalisée, et avec quelles fréquences relatives pour les évènements élementaires e_i. Nous savons seulement que si N est suffisament grand, nous aurons presque sûrement:

$$\forall \ell = 1,2,\ldots m: \quad \sum_{i=1}^{k} \frac{n_i}{N} f_\ell(e_i) \simeq \hat{f}_\ell \qquad (2)$$

et que l'approximation $\simeq$ peut devenir aussi bonne que l'on désire en augmentant N.

Soit S l'ensemble de toutes les suites dans Ω^N qui satisfont à (2). Puisque notre seule information est que la suite qui sera effectivement réalisée appartiendra presque sûrement à S, le choix de la loi de probabilité p_i s'affirme de façon simple: selon le principe de Laplace on devrait choisir, si possible, une p_i qui attribue la même probabilité à tous les éléments de S. On peut maintenant démontrer qu'une telle loi de probabilité existe, qu'elle est unique et qu'elle maximise l'entropie sous les contraintes (1). C'est là notre justification du PEM.

Il est en effet facile de voir que si p_i maximise l'entropie sous les contraintes (1), alors les probabilités de tous les éléments de S sont les mêmes. En ce cas, comme il est bien connu, il doit y avoir des constantes X, Y_ℓ (multiplicateurs de Lagrange) telles que:

$$\frac{\partial}{\partial p_i}\left(-\sum_i p_i \log p_i\right) + X \frac{\partial}{\partial p_i}\left(\sum_i p_i\right) + \sum_\ell Y_\ell \frac{\partial}{\partial p_i}\left(\sum_i p_i f_\ell(e_i)\right) = 0$$

ce qui implique:

$$\log p_i = -1 + X + \sum_\ell Y_\ell f_\ell(e_i) \qquad (3).$$

Les constantes X et Y_ℓ peuvent être déterminées à partir des contraintes $\sum_i p_i=1$ et (1). Maintenant, tout élément $C = e_{j_1} e_{j_2} \dots e_{j_N}$ de S a la probabilité $p(C) = \prod_i p_i^{n_i}$, où n_i est le nombre d'apparitions de e_i dans C. Si l'on prend un $C \in S$, alors les fréquences relatives $\frac{n_i}{N}$ obéissent à (2), et nous aurons en vertu de la relation (3):

$$\frac{\log p(C)}{N} = \sum_i \frac{n_i}{N} \log p_i \simeq -1+X+\sum_\ell Y_\ell \hat{f}_\ell$$

On constate donc effectivement que les probabilités p(C) des éléments de S ne dépendent pas des $\frac{n_i}{N}$ (elles dépendent uniquement des $\hat{f}_\ell$ et N) et elles sont donc toutes (presque) égales. Il est possible de démontrer aussi la réciproque de cette proposition. Si p(C) (ou $\frac{\log p(C)}{N}$) est indépendant des $\frac{n_i}{N}$ pour tout $C \in S$ et pour N suffisament grand, alors la probabilité p_i maximise l'entropie. Nous donnons la démonstration de cette affirmation dans la section qui suit, où notre étude est reprise de façon rigoureuse.

3. DEVELOPPEMENT RIGOUREUX.

Traduisons dans le language habituel des epsilon-delta tout ce qui vient d'être dit dans la section précédente. Ainsi, par exemple, la relation (2) devrait être une affirmation du type $|\sum_i \frac{n_i}{N} f_\ell(e_i) - \hat{f}_\ell| < \delta$ pour un δ positif donné. De même l'ensemble S que nous avons introduit devrait dépendre de N et de δ. Définissons donc, pour tout $N \in \mathbb{N}$ et $\delta>0$, l'ensemble

$$S_{N,\delta} = \{C \in \Omega^N : |\sum_i \frac{n_i}{N} f_\ell(e_i) - \hat{f}_\ell| < \delta, \forall \ell\} \qquad (4)$$

où, comme d'habitude, n_i représente le nombre de fois que e_i apparaît dans C. Comme il est bien connu, une distribution de probabilité p sur Ω satisfait (1) si et seulement si

$$\forall \delta > 0: \quad p(S_{N,\delta}) \xrightarrow[N \to \infty]{} 1 \qquad (5)$$

Nous voulons choisir la plus impartiale (unbiased) des lois de probabilité satisfaisant la relation (1) (ou, de façon équivalente, la relation (5)). Nous répétons maintenant l'argument de la section précédente. Puisque presque toutes les suites C qui peuvent effectivement se réaliser appartiennent à $S_{N,\delta}$, et que cela est <u>tout</u> ce que nous savons, nous devrions choisir, selon le principe de Laplace, une loi de probabilité p qui assigne approximativement les mêmes probabilités à tous les éléments de $S_{N,\delta}$. De plus, cette approximation devrait être aussi bonne que l'on veut si l'on prend δ suffisamment petit et N suffisamment grand. La traduction exacte de cette affirmation dans le language formel est la suivante:

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ et } \exists N_0 \in \mathbb{N} \text{ tels que:}$$

$$\forall N > N_0, \forall C, C' \in S_{N,\delta} : \left| \frac{\log p(C)}{N} - \frac{\log p(C')}{N} \right| < \varepsilon \qquad (6)$$

Nous utilisons $\frac{\log p(C)}{N}$ à la place de p(C) pour deux raisons. La première est que nous sommes guidés par l'approche heuristique qui a précédé. La seconde est que la comparaison des p(C) serait inutile. En effet, il est évident que p(C) tend vers zéro uniformément quand N s'accroît, pour tout $C \in S_{N,\delta}$ et pour toute loi de probabilité p_i. Par conséquent, $|p(C) - p(C')| < \varepsilon$ serait satisfaite indépendamment du choix des p_i.

La condition (6) est une affirmation mathématique plutôt compliquée qui possède une interprétation simple. Inversement, le PEM est une affirmation mathématique simple dont l'interprétation physique exacte est inconnue. Si nous arrivons à établir l'équivalence de la condition (6) et du PEM, nous aurons montré que la "maximalisation de l'entropie" est une façon mathématiquement simple pour dire "assignation de probabilités (presque) égales à toutes les suites qui ont des chances d'apparaître".

L'équivalence de (6) et du PEM sera établie par le théorème suivant. Mais faisons d'abord une convention: si notre information (1) implique que certaines des probabilités des e_i sont nécessairement nulles, nous exclurons ces évènements de l'espace Ω, puisqu'ils ne peuvent pas être réalisés. C'est pourquoi nous allons faire l'hypothèse que pour tout e_j il existe une mesure de probabilité $p^j(e_i) = p_i^j$ sur Ω satisfaisant (1) et telle que $p_j^j = 0$. Il s'ensuit que la mesure de probabilité qui maximise l'entropie est partout non nulle (La démonstration de cette proposition intuitivement évidente se trouve en appendice, lemme 1). Nous avons maintenant:

<u>*Théorème 1*</u> *Soit* $\Omega = \{e_1 \ldots e_k\}$ *un ensemble,* $f_\ell(e_i)$, $\ell = 1, 2, \ldots, m$ *des variables aléatoires sur* Ω, *et* $\hat{f}_\ell$ *des nombres réels. Soit encore p une mesure de probabilité sur* Ω *satisfaisant*

$$\sum_i p_i f_\ell(e_i) = \hat{f}_\ell, \quad \ell = 1, 2, \ldots m \tag{1}$$

et telle que $p_i \neq 0, \forall i$. *Alors p obéit à la condition (6) si et seulement si elle maximise l'entropie sous les contraintes (1).*

La démonstration du théorème est donnée dans l'appendice de ce chapitre. Ce théorème démontre ce que nous avons voulu établir: en effet, il est bien connu {58} que la mesure de probabilité qui maximise l'entropie existe toujours et est unique; par conséquent, d'après le théorème (1), la probabilité qui satisfait (6) + (1) existe, elle est unique, et elle maximise l'entropie.

On pourrait se poser la question suivante: comment le théorème (1) peut-il être compatible avec la loi des grands nombres? Car supposons que la loi de probabilité p_i qui régit effectivement l'expérience soit celle qui maximise l'entropie sous les contraintes (1). Alors la loi des grands nombres affirme que la suite C qui apparaîtra après N répétitions de l'expérience aura, presque sûrement, des fréquences relatives $\frac{n_i}{N} \simeq p_i$

D'autre part, le théorème affirme que toutes les suites possèdant des fréquences relatives telles que $\sum_i \frac{n_i}{N} f_\ell(e_i) \simeq \hat{f}_\ell$ ont approximativement la même probabilité, même celles pour lesquelles $\frac{n_i}{N}$ est complètement différent des p_i. Mais alors pourquoi la réalisation d'une suite telle que $\frac{n_i}{N} \simeq p_i$ est-elle plus probable? La réponse à cette question est simple: toutes les suites qui satisfont à $\sum_i \frac{n_i}{N} f_\ell(e_i) \simeq \hat{f}_\ell$ ont, en effet, la même probabilité, mais le nombre des suites pour lesquelles on a en plus $\frac{n_i}{N} \simeq p_i$ est écrasant par rapport au nombre de toutes les autres suites pour lesquelles (2) est vraie. La démonstration de ce fait est très simple et elle se trouve en essence dans plusieurs articles sur le PEM (cf., par exemple, {60} page 231).

4. LE "GAIN D'INFORMATION" DE S. KULLBACK.

Quelques années après l'article fondamental de Jaynes, S. Kullback {62,63} a proposé une généralisation du PEM. Cette généralisation concerne les cas où/~~l'arr~~ avant l'arrivée de l'information (1) il existait une information préalable qui avait permis de fixer une mesure de probabilité a priori. En termes plus précis, le problème étudié par Kullback peut être ennoncé ainsi: soit de nouveau $\Omega = \{e_1, e_2, \ldots, e_k\}$ l'ensemble de tous les résultats d'une expérience aléatoire dont nous ignorons la loi de probabilité exacte. Comme nous l'avons dit, lorsqu'on possède certaines informations sur l'expérience aléatoire, on pose sur Ω une mesure de probabilité qui représente "le mieux" cette information. Supposons donc qu'à un certain moment nos informations soient correctement représentées par une mesure de probabilité $q(e_i) \equiv q_i$ sur Ω. Cette mesure implique une certaine valeur moyenne $\hat{f}'_\ell$ pour les variables aléatoires $f_\ell(e_i)$:

$$\sum_{i=1}^{k} q_i f_\ell(e_i) = \hat{f}'_\ell \qquad (7)$$

Bien sûr, ces valeurs moyennes n'ont pas un caractère objectif, mais représentent seulement des estimations, étant donné que q exprime simplement un certain niveau des connaissances. Supposons maintenant qu'une mesure nous ait permis de déterminer expérimentalement, objectivement, les valeurs moyennes $\hat{f}_\ell$ sans pour autant fixer la vraie mesure de probabilité (nous rappelons que ce cas se présente couramment, par exemple, en Mécanique Statistique). Il y a alors deux cas possibles: Si $\hat{f}_\ell=\hat{f}'_\ell$, les nouvelles informations nous renforcent dans notre choix de la probabilité q_i. Si, au contraire, $\hat{f}_\ell \neq \hat{f}'_\ell$ pour certains indices ℓ, alors on devrait choisir une autre mesure $p(e_i) = p_i$ pour représenter nos nouvelles connaissances. Le problème posé par Kullback est celui de choisir la mesure de probabilité *a posteriori* p_i en fonction de la nouvelle information, représentée par la relation (1), <u>et</u> de l'ancienne information, représentée par la mesure de probabilité *a priori* q_i.

Le choix proposé par Kullback est le suivant: p_i est la mesure de probabilité qui, tout en satisfaisant à (1), minimise le "gain d'information" ou "information de discrimination" $\sum_i p_i \log \frac{p_i}{q_i}$. De nouveau, les divers arguments avancés en faveur de ces principes ont un caractère qualitatif et ne sont pas concluants.

Par une méthode similaire à celle que nous avons utilisée pour examiner le principe d'entropie maximale, nous allons maintenant élucider les conséquences du choix de Kullback. Il nous semble que cette élucidation apporte en même temps un argument suffisamment fort en faveur de ce choix.

Considérons de nouveau l'ensemble $S \subset \Omega^N$ de toutes les suites $C \in \Omega^N$ telles que $\sum_i \frac{n_i}{N} f_\ell(e_i) \simeq \hat{f}_\ell$. L'une de nos deux données, à savoir l'information (1), implique que la nouvelle mesure de probabilité p sur Ω sera telle que $p(S) \simeq 1$, si N est assez grand. Cela signifie que les suites C qui pourraient effectivement se réaliser appartiennent presque sûrement à S. Avec cette affirmation nous avons <u>entièrement</u> utilisé la donnée exprimée par la relation (1) car, comme nous l'avons remarqué, (1) est équivalente à (5).

Considérons maintenant la deuxième donnée, i.e. la connaissance de la probabilité a priori q. Comme $p(S) \simeq 1$, p s'annule pratiquement hors de S. En outre, on voudrait respecter au maximum l'information préalable exprimée par q, et donc de modeler p en conséquence à l'intérieur de S. Cela serait réalisé s'il était possible de trouver une loi de probabilité $p_i = p(e_i)$ parmi toutes celles qui satisfont à (1), telle que l'on ait sur S:

$$\forall\, C, C' \in S \; : \qquad \frac{p(C)}{p(C')} = \frac{q(C)}{q(C')} \qquad (8)$$

Le choix indiqué par la relation (8) peut se justifier ainsi: en l'absence de toute information préalable, i. e. s'il n'y avait pas de probabilité a priori, on serait tout a fait dans le cas étudié par Jaynes. Le meilleur choix est alors celui qui satisfait $p(C) = p(C')$ pour tout $C, C' \in S$, car ce choix évite de favoriser un élément de S par rapport à un autre. Mais lorsqu'il existe une information préalable qui nous dit que la probabilité a priori de C est, par exemple, le double de la probabilité de C', le meilleur choix pour la probabilité a posteriori est celui qui conserve ce rapport, si ce choix est possible. Une représentation très schématique de ce choix est donnée par la figure 4.

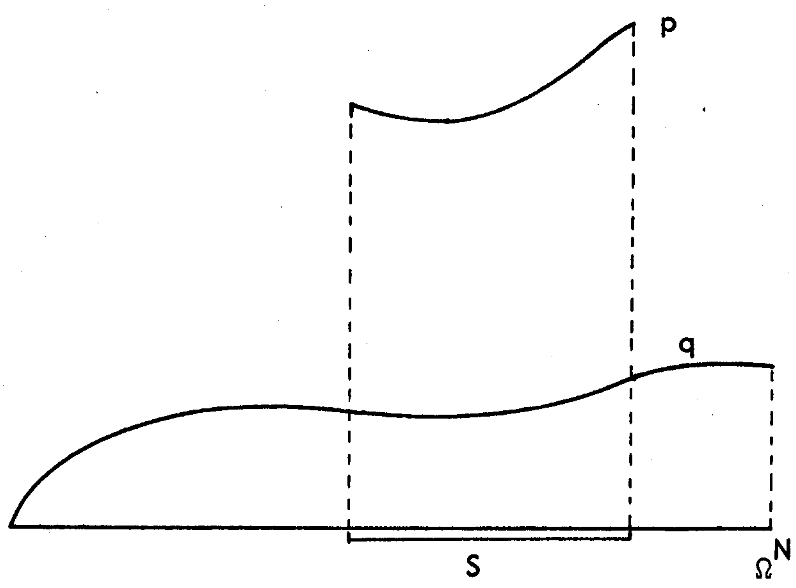


Figure 4.

Nous allons maintenant démontrer que ce choix est possible, qu'il est unique et que la loi de probabilité p qui en résulte minimise le gain d'information de Kullback. Mais faisons encore deux remarques avant d'énoncer le théorème: la comparaison des rapports se fera à travers la fonction $\frac{\log}{N}$, comme dans le cas de Jaynes. En plus, nous excluerons içi aussi les évènements élémentaires $e_i \in \Omega$ qui ont une probabilité certainement nulle d'après l'information (1). Ainsi, soit Ω' l'ensemble de tous les éléments $e_i \in \Omega$ tels qu'il existe une mesure de probabilité p^i sur Ω satisfaisant à (1) et telle que $p^i(e_i) \neq 0$. Nous avons:

<u>Théorème 2</u> *Soit* $\Omega = \{e_1, e_2 \dots e_k\}$ *un ensemble,* $f_\ell(e_i)$ $(\ell=1,\dots m)$ *des fonctions sur* Ω *et* $\hat{f}_\ell$ *des nombres réels. Soit encore q une mesure de probabilité sur* Ω *telle que* $q_i \neq 0, \forall q_i$. *Soit enfin* Ω' *l'ensemble défini plus haut et* $S_{N,\delta}$ *l'ensemble*

$$S_{N,\delta} = \{C \in \Omega'^N : \left| \sum_i \frac{n_i}{N} f_\ell(e_i) - \hat{f}_\ell \right| < \delta \} \qquad (9)$$

où $N \in \mathbb{N}$ *et* $\delta > 0$. *Alors il existe une et une seule mesure de probabilité p sur* Ω, *satisfaisant à (1) et à la condition:*

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ et } \exists N_0 \in \mathbb{N} \text{ tels que: } \forall N \geq N_0,$$

$$\forall C, C' \in S_{N,\delta} : \left| \frac{\log \frac{p(C)}{q(C)}}{N} - \frac{\log \frac{p(C')}{q(C')}}{N} \right| < \varepsilon \qquad (10)$$

et c'est exactement la mesure de probabilité qui minimise le "gain d'information" $\sum_i p_i \log \frac{p_i}{q_i}$ *sous la contrainte (1).*

<u>Preuve</u>. C' est un fait connu {62,63} qu'il existe toujours une et une seule mesure de probabilité p qui minimise le gain d'information sous les contraintes (1). Pour démontrer donc le théorème, il suffit de démontrer que la minimisation du gain d'information est équivalente à (1) + (10). La démonstration est identique à celle du théorème (1) (il suffit de remplacer partout Ω par Ω', p (C) par $\frac{p(C)}{q(C)}$ et p_i par $\frac{p_i}{q_i}$) et donc nous l'omettons. $\blacksquare$

5. <u>COMMENTAIRES</u>.

Notre but était de démontrer que le PEM est une conséquence du principe de Laplace. Afin d'y parvenir, nous sommes passés de l'espace Ω à l'espace Ω^N et de là au sous-ensemble

$S \subsetneq \Omega^N$ défini dans la section 2. La raison en est évidente: on ne peut appliquer le principe de Laplace sur Ω puisque l'information disponible contient plus que la seule connaissance de l'ensemble des résultats possibles. D'autre part, toute l'information disponible peut être exprimée par le fait que, presque sûrement, $C \in S$, et par conséquent le principe de Laplace est applicable à S. Nous avons atteint notre but en démontrant que cela implique le PEM.

Le théorème de la section 4 donne aussi un sens précis à l'affirmation selon laquelle la distribution de probabilité qui rend l'entropie maximale est plus "étalée" que toutes les autres {60} . En fait, cette distribution est uniforme sur S.

Il est bien sûr légitime de désirer une élucidation encore plus profonde des racines conceptuelles du PEM. De toute manière, notre résultat montre qu'il suffit pour cela de concentrer son attention sur le principe de Laplace, qui est conceptuellement et mathématiquement beaucoup plus simple.

L'élucidation des principes de Jaynes et de Kullback peut se faire aussi par d'autres biais. Une approche prometteuse est celle qui utilise le nouveau concept d'opacité d'une statistique face à une loi de probabilité, mis au point par M. Mugur-Schächter {64, 65} .

APPENDICE AU CHAPITRE IV

Lemme 1 *Soit Ω, $f_\ell(e_i)$ et $\hat{f}_\ell$ comme dans l'ennoncé du théorème IV.1. Supposons que pour tout e_j il existe une mesure de probabilité $p^j(e_i) = p^j_i$ sur Ω satisfaisant (1) et telle que $p^j_j \neq 0$. Soit enfin p la mesure de probabilité qui maximise l'entropie sous les contraintes (1). Alors $p_i \neq 0, \forall i$.*

Preuve. Soit q la mesure de probabilité définie par $q = \sum_j \frac{p^j}{k}$. On voit immédiatement que $q(e_i) \equiv q_i \neq 0$ $\forall i$, et que q satisfait (1).

Considérons maintenant pour chaque $a \in [0,1]$ la mesure de probabilité $p(a) = ap + (1-a)q$. Son entropie est donnée pour chaque a par la fonction:

$$g(a) = -\sum_i (ap_i + (1-a)q_i)\log(ap_i + (1-a)q_i)$$

Il est facile de voir que p(a) satisfait (1) et donc,par la définition de la mesure p, la fonction g(a) a un maximum absolu pour a=1. Si l'on avait $p_i=0$ pour certains i, alors on aurait aussi, comme le montre un calcul élementaire, $g'(a) \xrightarrow[a \to 1]{} -\infty$ et donc g(1) ne pourrait être un maximum. Le lèmme est donc démontré.

Démonstration du théorème 1 :

a) Supposons que p_i maximise l'entropie. Alors, comme nous le savons déjà, il existe des constantes X, Y_ℓ telles que (3) soit vraie. Pour $\delta > 0$ et $N \in \mathbb{N}$ fixés, soit $C \in S_{N,\delta}$. Alors:

$$\frac{\log p(C)}{N} = \sum_i \frac{n_i}{N} \log p_i = -1 + X + \sum_\ell Y_\ell \left(\sum_i \frac{n_i}{N} f(e_i)\right)$$

et par conséquent, pour tout $C, C' \in S$ on a:

$$\left| \frac{\log p(C)}{N} - \frac{\log p(C')}{N} \right| = \left|\sum_\ell Y_\ell \left(\sum_i \frac{n_i}{N} f_\ell(e_i) - \sum_i \frac{n_i}{N} f_\ell(e_i)\right)\right|$$

$$\leqslant \sum_{\ell} |Y_\ell| \left(\left| \sum_i \frac{n_i}{N} f_\ell(e_i) - \hat{f}_\ell \right| + \left| \sum_i \frac{n'_i}{N} f_\ell(e_i) - \hat{f}_\ell \right| \right) <$$

$$< 2\delta \sum_{\ell} |Y_\ell|$$

Nous en déduisons la condition (6), si pour tout $\varepsilon > 0$ nous choisissons $\delta = \varepsilon / (2 \sum |Y_\ell|)$.

b) Supposons que p_i satisfasse (6) et (1). Nous posons $\hat{f}_0 = 1$, et nous définissons dans R^k les vecteurs suivants:

$$\vec{f}_0 = (\underbrace{1,1,\ldots 1}_{k \text{ fois}}), \quad \vec{f}_\ell = (f_\ell(e_1), \ldots f_\ell(e_k)), \quad \vec{q} = (\log p_1, \ldots \log p_k) \qquad (13).$$

Soient $\vec{r}$ et $\vec{r}\,'$ deux vecteurs quelconques de R^k tels que:

$$\begin{aligned} &\vec{r} \cdot \vec{f}_\ell = \vec{r}\,' \cdot \vec{f}_\ell = \hat{f}_\ell, \ \forall \ell = 0,1,2,\ldots m, \\ &r_i > 0 \ , \ r'_i > 0, \forall i = 1,2,\ldots k. \end{aligned} \qquad (14)$$

Nous allons prouver que $\vec{r} \cdot \vec{q} = \vec{r}\,' \cdot \vec{q}$. Pour cela, nous procédons comme suit. Pour tout $\varepsilon > 0$, soient $\delta > 0$ et N_0 tels que (6) soit vraie. Nous pouvons prendre $\delta < \varepsilon$. Définissons $\delta' = \delta / \max_\ell \sum_i |f_\ell(e_i)|$. Puisque l'ensemble des rationnels est dense dans R, on peut trouver des entiers non négatifs n_i, n'_i et un entier $N > N_0$ tels que:

$$\forall i = 1, \ldots k : \left| \frac{n_i}{N} - r_i \right| < \delta', \left| \frac{n'_i}{N} - r'_i \right| < \delta', \ \sum n_i = \sum n'_i = N \qquad (15)$$

Soient maintenant dans Ω^N deux suites C, C' dans lesquelles les e_i apparaissent respectivement n_i et n'_i fois. Alors:

$$\begin{aligned} \left| \sum_i \frac{n_i}{N} f_\ell(e_i) - \hat{f}_\ell \right| &= \left| \sum_i \frac{n_i}{N} f_\ell(e_i) - \sum_i r_i f_\ell(e_i) \right| \leqslant \\ &\leqslant \sum_i \left| r_i - \frac{n_i}{N} \right| |f_\ell(e_i)| < \delta \end{aligned}$$

et donc C, C' $\in S_{N,\delta}$. Le choix du δ implique que:

$$\left| \sum_i \frac{n_i}{N} \log p_i - \sum_i \frac{n'_i}{N} \log p_i \right| < \varepsilon \qquad (16)$$

On a maintenant:

$$|\vec{r}\cdot\vec{q} - \vec{r}'\cdot\vec{q}| = |\sum_i r_i \log p_i - \sum_i r'_i \log p_i| \leq$$

$$\leq |\Sigma\, r_i \log p_i - \Sigma \frac{n_i}{N} \log p_i| + |\Sigma \frac{n_i}{N} \log p_i - \Sigma \frac{n'_i}{N} \log p_i| +$$

$$+ |\Sigma \frac{n'_i}{N} \log p_i - \Sigma r'_i \log p_i| \leq \Sigma |\frac{n_i}{N} - r_i|\ |\log p_i| +$$

$$+ |\Sigma \frac{n_i}{N} \log p_i - \Sigma \frac{n'_i}{N} \log p_i| + \Sigma |\frac{n'_i}{N} - r'_i|\ |\log p_i| <$$

$$< 2\delta' \Sigma |\log p_i| + \varepsilon < \varepsilon\left(1 + \frac{2\sum_i |\log p_i|}{\max_\ell \sum_i |f_\ell(e_i)|}\right)$$

On en déduit que pour tout $\varepsilon > 0$ on a $|\vec{r}\cdot\vec{q} - \vec{r}'\cdot\vec{q}| < \varepsilon$, ce qui implique $\vec{r}\cdot\vec{q} = \vec{r}'\cdot\vec{q}$. Nous avons donc démontré que:

$$(\vec{r}\cdot\vec{f}_\ell = \vec{r}'\cdot\vec{f}_\ell = \hat{f}_\ell \;\forall \ell \text{ et } r_i > 0,\ r'_i > 0) \Longrightarrow \vec{q}\cdot\vec{r} = \vec{q}\cdot\vec{r}'$$

Définissons maintenant les ensembles:

$$A = \{\vec{r} \in R^k : \vec{r}\cdot\vec{f}_\ell = 0,\ \ell = 0, 1, \ldots m\}$$

$B = [\vec{f}_o, \vec{f}_1, \ldots \vec{f}_m]$ (sous-espace engendré par les $\vec{f}_\ell$)

Il est évident que A est le complément orthogonal de B, i. e. $A = B^\perp$ et donc $B = A^\perp$. Nous allons maintenant montrer que pour tout $\vec{r} \in A$ on a $\vec{r}\cdot\vec{q} = 0$. Si $\vec{r} = 0$, cela est évident. Si $\vec{r} \neq 0$, considérons le vecteur $\vec{p} = (p_1, \ldots, p_k)$ et définissons pour tout $b \in \mathbb{R}$: $\vec{r}(b) = \vec{p} + b\,\vec{r}$. Puisque $r_i(0) = p_i > 0$, on peut trouver $b \neq 0$ tel que $r_i(b) > 0$. Alors $\vec{r} = (\vec{r}(b) - \vec{p})/b$. Comme nous avons évidemment $\vec{r}(b)\cdot\vec{f}_\ell = \vec{p}\cdot\vec{f}_\ell = \hat{f}_\ell$, nous déduisons de (17) que $\vec{r}\cdot\vec{q} = 0$. Cela est vrai pour tout $\vec{r} \in A$, et donc $\vec{q} \in B$. Par conséquent, il existe des constantes a_ℓ, $\ell = 0, 1, \ldots, m$, telles que $\vec{q} = \sum_{\ell=0}^{m} a_\ell \vec{f}_\ell$ ou, de façon équivalente:

$$\forall\, i : \log p_i = a_o + \sum_{\ell=1}^{m} a_\ell f_\ell(e_i)$$

Cette relation est équivalente à la relation (3), via une identification évidente des termes. Il s'ensuit que p_i maximise l'entropie sous les contraintes (1). $\blacksquare$

CHAPITRE V

L'INFERENCE STATISTIQUE EN THEORIE QUANTIQUE ET LA DISTANCE ENTRE LES ETATS.

1. INTRODUCTION.

Nous avons vu que dans le cas classique, le problème de l'inférence statistique consiste à spécifier une loi de probabilité qui représente le mieux l'information disponible sur une expérience aléatoire. Dans le cadre de la Théorie Quantique, le même problème consiste à spécifier un descripteur d'état qui "représente le mieux nos connaissances" sur un système physique donné (nous aurons plus tard l'occasion de préciser notre langage sur ce point).

Il s'agit là d'un problème crucial. En effet, si le formalisme quantique permet de calculer des prédictions statistiques lorsque l'état est connu, il ne donne guère d'indications concernant la détermination de cet état. D'autre part, sauf dans des cas idéalisés (mesure de première espèce d'un ensemble complet d'observables qui commutent), l'information disponible est en général insuffisante pour fixer de façon unique ce descripteur d'état.

Le problème de l'inférence statistique est beaucoup plus difficile dans le cas quantique que dans sa version classique, et cela pour plusieurs raisons. Tout d'abord, comme nous l'avons déjà dit, la théorie quantique traite avec plusieurs espaces de probabilité à la fois, un pour chaque observable {66,67}. Ensuite, le formalisme quantique est complexe du point de vue formel, ce qui crée des difficultés purement mathématiques. Enfin et surtout, l'interprétation du formalisme pose des questions délicates et pas encore résolues. C'est pourquoi tout problème qui concerne les fondements même de la théorie, comme le problème de l'inférence statistique, doit être traité avec une grande attention. Il est bien connu en effet que l'intuition physique doit être utilisée avec discernement, faute de quoi elle trahit souvent le chercheur en conduisant à des paradoxes*

---

* Dans la référence {68} , nous avons mis en évidence la clé d'un de ces paradoxes, produit par F. Selleri dans {69, 70} .

Comme nous l'avons fait dans le chapitre précédent pour la version classique du problème, on peut distinguer deux cas différents dans la théorie quantique suivant la nature de l'information disponible. Lorsque l'information consiste en la donnée de valeurs moyennes de certaines observables, on traite le problème par une méthode qui transpose celle définie par Jaynes pour le cas classique. Lorsque l'information consiste en la donnée des valeurs moyennes et d'un "état *a priori* " le problème est analogue à celui de Kullback. C'est ce dernier cas qui nous intéressera dans ce chapitre. Au lieu du "gain d'information" de Kullback nous définirons un concept de distance entre les états quantiques, que nous utiliserons d'une manière équivalente.

Ce chapitre comprend deux sections principales. Dans la section 2 nous introduirons le concept qui sera l'instrument de notre étude du problème de l'inférence statistique, à savoir la distance entre deux états quantiques, et nous trouverons son expression dans l'espace d'Hilbert. Ensuite, dans la section 3 nous utiliserons la distance dans deux cas d'inférence statistique. L'un de nos résultats sera la déduction de ce qu'on appelle "postulat de projection" dans un cadre qui sera précisé.

Le chapitre est complété par deux annexes mathématiques se trouvant à la fin de la thèse dont chacune contient la solution d'un problème précis. Dans la première nous donnons une caractérisation simple des mélanges d'états non orthogonaux qui nous sera utile dans la partie mathématique de ce chapitre. Dans la deuxième, nous étudions la distance proposée par V. Cantoni.

2. LA DISTANCE ENTRE DEUX ETATS D'UN SYSTEME.

2.1 *La définition de la distance.*

Nous allons tout d'abord spécifier certaines conditions auxquelles une distance appropriée devrait obéir. D'un point de vue mathématique, la distance $d(p_1,p_2)$ entre deux états doit posséder les propriétés suivantes:

a) $d(p_1, p_2) \geqslant 0, \quad d(p_1, p_2) = d(p_2, p_1)$

b) $d(p_1, p_2) = 0 \Longrightarrow p_1 = p_2$

Il serait aussi souhaitable que la distance satisfasse à l'inégalité triangulaire:

c) $d(p_1, p_3) \leqslant d(p_1,p_2) + d(p_2, p_3)$

i.e. qu'elle soit une métrique. Il faut toutefois noter qu'il existe des "distances" comme l'information de discrimination de Kullback qui ne sont pas des métriques (cf. section IV.4 et {71} .

Plusieurs distances différentes, satisfaisant à toutes les conditions (a), (b) et (c) ont été définies jusqu'à ce jour {72 - 79} . Parmi toutes ces distances, il y en a deux seulement dont la définition a été fondée sur des arguments physiques, à savoir: la distance $d_G(p_1,p_2)$ de S. Gudder {72} et la distance $d(p_1, p_2)$ définie par J.M. Jauch, B. Misra et Gibson {79, 80} et aussi par N.S. Kronfli {81} et nous (cf. référence {82}). Une troisième distance interessante est celle

de V. Cantoni, qui acquiert une signification physique *a posteriori.*

La distance de Cantoni est assez particulière et mérite un examen à part. Nous lui avons réservé l'annexe II, où nous construisons son expression dans le formalisme Hilbertien. La distance $d_G(p_1, p_2)$ de Gudder est introduite ainsi:
Soit Σ l'ensemble des opérateurs statistiques (ou "opérateurs densité"), i.e. des opérateurs positifs à trace unité, dans l'espace d'Hilbert H qui décrit un système physique. Ces opérateurs statistiques représentent, comme il est bien connu, les états du système. L'ensemble Σ est convexe: si $0 < a < 1$ et $W_1,W_2 \in \Sigma$, alors le "mélange" $aW_1 + (1 - a) W_2$ appartient aussi à Σ . Sur la base de cette propriété, S. Gudder qualifia deux états p_1, p_2, représentés par les opérateurs densité W_1 et W_2, comme étant "proches" , s'il existe un mélange contenant princi-

palement W_1, qui soit égal à un mélange contenant principalement W_2 ; en d'autres termes, s'il existe W_3, $W_4 \in \Sigma$ et un petit $a \in [0,1]$ tels que $(1-a)W_1+aW_3= (1-a)W_2+aW_4$. Ensuite, en se fondant sur cette conception de la proximité, il a défini la distance $d_G(p_1,p_2)$ entre deux états par la formule:

$$d_G(p_1,p_2)=\inf\{a\in[0,1]:(1-a)W_1+aW_3=(1-a)W_2+aW_4;W_3,W_4 \in \Sigma\} \qquad (1)$$

Bien évidemment on ne peut pas nier le fait que les états quantiques ont la propriété de former des mélanges. Il semble toutefois très improbable *a priori* qu'une distance qui tient compte seulement de cette propriété secondaire puisse exprimer proprement la distance globale entre deux états. C'est pourquoi nous allons plutôt concentrer notre attention sur une autre distance, proposée pour la première fois par Jauch, Misra et Gibson. La définition de cette seconde distance utilise le formalisme propositionnel de von Neumann(cf. Chapitre I)

Nous rappelons que, selon ce formalisme, à chaque système physique correspond un ensemble $\mathcal{L}$ de "propositions". En outre si le système ne possède pas de règles de super-selection {83, 84}, $\mathcal{L}$ est isomorphe à l'ensemble de sous-espaces fermés de l'espace d'Hilbert qui décrit le système.* Enfin nous retenons du formalisme de Jauch la représentabilité des états par des mesures de probabilité $p(a), a\in\mathcal{L}$ (voir chapitre II, définitions DJ_{14} , DJ_{15}). Nous rappelons aussi le théorème de Gleason (chapitre II, théorème TJ2) selon lequel pour chaque état p il existe un opérateur statistique W

---

* Cette hypothèse est, bien-sûr, trop forte. Nous avons vu en effet qu'il serait irréaliste d'exiger que <u>chaque</u> sous-espace corresponde à une proposition physique. Toutefois, tous nos résultats seraient valables si l'on faisait à sa place l'hypothèse beaucoup mieux fondée que $\mathcal{L}$ est isomorphe à un sous ensemble dense de l'ensemble de sous-espaces fermés.

tel que:

$$\forall\, a \in \mathcal{L}: \quad p(a) = \mathrm{Tr}\,(W E_a) \qquad (2)$$

E_a étant le projecteur sur le sous-espace qui correspond à la proposition $a \in \mathcal{L}$. On a maintenant:

Définition 1. La distance entre deux états p_1, p_2 est donnée par l'expression:

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

La justification de cette définition est basée sur l'argument suivant, lequel, sinon inébranlable, est certainement plus solide que les arguments (lorsqu'ils existent) en faveur d'autres définitions.

Comme nous l'avons déjà remarqué dans l'introduction de cette deuxième partie, la mécanique quantique décrit les microsystèmes exclusivement via leur comportement lors d'une mesure. En particulier, un état quantique non seulement fournit des prédictions statistiques, mais il est aussi complètement déterminé à partir de ces prédictions (il s'agit là d'un des aspects du théorème de Gleason). Par conséquent la distance entre deux états p_1, p_2 peut être mesurée par la différence des prédictions statistiques données par p_1 et p_2 . Pour une proposition $a \in \mathcal{L}$ donnée, la différence des prédictions statistiques des p_1, p_2 sur a est $|p_1(a) - p_2(a)|$. Par conséquent une distance appropriée devrait se fonder sur l'étude de l'ensemble de toutes les valeurs de la fonction:

$$f(a) := |p_1(a) - p_2(a)| \ , \ a \in \mathcal{L}$$

Nous allons démontrer dans la section suivante que l'ensemble des valeurs de cette fonction est un intervalle de la forme $[0,d]$. Maintenant, si d est petit, alors la différence des prédictions $|p_1(a)-p_2(a)|$ est petite pour toute $a \in \mathcal{L}$ et donc les deux états devraient être considérés comme étant proche l'un de l'autre.

Si au contraire d est grand, alors $|p_1(a) - p_2(a)|$ est approximativement égal à d pour un nombre infini de propositions et donc la distance entre p_1 et p_2 devrait être grande. C'est pourquoi il apparaît que le nombre d est une mesure appropriée de cette distance, et cela explique le choix de la définition 1.

Jauch, Misra et Gibson ont proposé cette distance dans le seul but de définir une topologie sur l'ensemble des états, pour pouvoir introduire correctement les limites qui se présentent lors d'une étude rigoureuse du processus de diffusion. Ils ont déterminé la topologie introduite par $d(p_1, p_2)$, mais n'ont pas donné l'expression de cette distance dans le formalisme Hilbertien. Notre but est d'utiliser cette distance pour l'étude du problème de l'inférence statistique en Théorie Quantique. Nous avons besoin pour cela de l'expression Hilbertienne de $d(p_1, p_2)$ pour toute paire d'états, purs ou mélangés. Cette expression, et quelques autres propriétés, sont données dans la sous-section suivante.

2.2 *Forme explicite de $d(p_1, p_2)$ dans le formalisme Hilbertien.*

Nous commençons par un bref rappel de quelques définitions et résultats de la théorie de la trace {85, 86} . On dit qu'un opérateur A dans un espace d'Hilbert complexe et séparable H appartient à la classe des opérateurs à trace si la série $\sum_{n=1}^{\infty} (f_n , |A| f_n)$ converge au moins pour une base $\{f_n\}_{n\in N}$. Dans ce cas, l'expression $\mathrm{Tr}A = \sum_{n=1}^{\infty} (f_n , A f_n)$ est indépendante de la base $\{f_n\}_{n\in N}$. La classe des opérateurs à trace est notée $L_1(H)$ et elle est un espace de Banach par rapport à la norme $\| A \|_1 = \mathrm{Tr}|A|$. Les opérateurs statistiques sont, par définition, les éléments positifs de norme unité dans $L_1(H)$.

Tout opérateur autoadjoint $A \in L_1(H)$ a un spectre discret et ses valeurs propres sont de dégénérescence finie. Nous écrirons donc la décomposition spectrale de A comme $A = \sum_n \lambda_n P_{[f_n]}$, $\{f_n\}_{n\in N}$ étant une base de vecteurs propres de A et λ_n ses valeurs propres correspondants. On a alors: $\mathrm{Tr}A = \sum_n \lambda_n$ et $\| A \|_1 = \sum_n |\lambda_n|$. La norme habituelle (borne) des opérateurs continus est donnée dans ce cas par $\|A\| = \sup\{|\lambda_n| : n \in N\}$

Quelques conventions sur les notations: L (H) est l'espace de Banach de tous les opérateurs continus de H. Pour tout sous ensemble $K \subseteq H$, nous notons par [K] le sous espace algébrique engendré par K- i.e.l'ensemble de toutes les combinaisons linéaires finies d'éléments de K- et par $\overline{K}$ la fermeture topologique de K. Pour tout vecteur normalisé g, nous notons par $P_{[g]}$ le projecteur sur le sous espace unidimentionnel qui contient g. Pour tout opérateur A, R(A) sera son image. Si A est autoadjoint, alors A^+ et A^- seront ses parties positive et négative.

Nous donnons maintenant la forme explicite de d (p_1, p_2) dans le cas le plus général.

<u>*Théorème 1*</u>. *Soit p_1 et p_2 deux états représentés par les opérateurs statistiques W_1 , W_2. Alors:*

$$d(p_1,p_2)=\frac{\|W_1-W_2\|_1}{2}$$

où $d(p_1, p_2)$ est defini par la définition 1.

<u>Démonstration</u>. Puisque W_1, $W_2 \in L_1(H)$, on a aussi $W_1 - W_2 \in L_1(H)$. Soit $W_1 - W_2 = \sum_n \lambda_n P_{[f_n]}$ la décomposition spectrale de $W_1 - W_2$, $\lambda_n$étant ses valeurs propres et$\{f_n\}_{n\in N}$ une base de vecteurs propres. La définition de la distance implique alors:

$$d(p_1, p_2) = \sup\{|Tr(W_1E)-Tr(W_2E)| : E \text{ projecteur}\}$$

$$= \sup_E |Tr((W_1-W_2)E)| = \sup_E |\sum_n \lambda_n (f_n, Ef_n)| \qquad (3)$$

D'autre part,

$$\forall n, \forall E : \quad 0 \leqslant (f_n, Ef_n) \leqslant 1 \quad \Longrightarrow$$

$$\Longrightarrow \forall\, \lambda_n > 0: \quad 0 \leqslant \lambda_n (f_n, Ef_n) \leqslant \lambda_n$$

et $$\forall\, \lambda_n < 0: \quad \lambda_n \leqslant \lambda_n (f_n, Ef_n) \leqslant 0$$

Par conséquent,

$$\sum_{\lambda_n<0} \lambda_n \leqslant \sum_n \lambda_n (f_n, Ef_n) \leqslant \sum_{\lambda_n>0} \lambda_n \tag{4}$$

Puisque $TrW_1 = TrW_2 = 1$, on a $\sum\limits_n \lambda_n = Tr\,(W_1 - W_2) = 0$, et donc

$$\sum_{\lambda_n<0} \lambda_n = -\sum_{\lambda_n>0} \lambda_n$$

En substituant dans la relation (4) on trouve:

$$\forall\ E : \left|\sum_n \lambda_n (f_n, Ef_n)\right| \leqslant \sum_{\lambda_n>0} \lambda_n$$

$$\Longrightarrow \sup_E \left|\sum_n \lambda_n (f_n, Ef_n)\right| \leqslant \sum_{\lambda_n>0} \lambda_n \tag{5}$$

Posons $E_o = \sum\limits_{\lambda_n>0} P_{[f_n]}$. Alors

$$\left|\sum_n \lambda_n (f_n, E_o f_n)\right| = \sum_{\lambda_n>0} \lambda_n \tag{6}$$

En utilisant (3), (5) et (6) on obtient:

$$d\,(p_1, p_2) = \sum_{\lambda_n>0} \lambda_n \tag{7}$$

D'autre part,

$$\| W_1 - W_2 \|_1 = \sum_n |\lambda_n| = 2 \sum_{\lambda_n>0} \lambda_n$$

Le théorème découle par une substitution dans la relation (7). ∎

Notons un fait remarquable: La distance qui nous a semblé la plus "naturelle" parmi toutes les distances définies jusqu'à ce jour, est donnée, à une constante près, par la distance $\| W_1 - W_2 \|_1$ introduite par la norme naturelle de l'espace mathématique $L_1(H)$ associé à l'ensemble des états. Toutefois, en dépit du fait que l'expression $\|W_1 - W_1\|_1/2$ est générale et élégante, il est difficile de la calculer dans la pratique. Heureusement lorsqu'un des deux états est pur, l'expression de la distance devient plus simple comme le montre le résultat suivant:

__Proposition 1__. Soit p_1 un état pur, représenté par un vecteur normalisé f, et p_2 un état quelconque représenté par l'opérateur statistique W. Alors:

a) L'espace $\overline{R(P_{[f]} - W)^+}$ est unidimentionnel, et

b) $d\,(p_1, p_2) = \|P_{[f]} - W\|$.

__Démonstration__ Soit F le projecteur sur $\overline{R(P_f - W^+)}$

Alors

$$P_{[f]} - W = (P_{[f]} - W)^+ - (P_{[f]} - W)^- \Longrightarrow$$

$$\Longrightarrow F(P_{[f]} - W)F = F(P_{[f]} - W)^+ F \Longrightarrow$$

$$\Longrightarrow FP_{[f]}F = (P_{[f]} - W)^+ + FWF \qquad (8)$$

L'opérateur $FP_{[f]}F$ a une image unidimentionnelle, et donc la même chose sera vraie pour le second membre de (8). En appliquant maintenant le corollaire 1 de l'annexe I, on trouve:

$$\overline{R((P_{[f]} - W)^+ + FWF)} = \overline{R((P_{[f]} - W)^+) + R(FWF)}$$

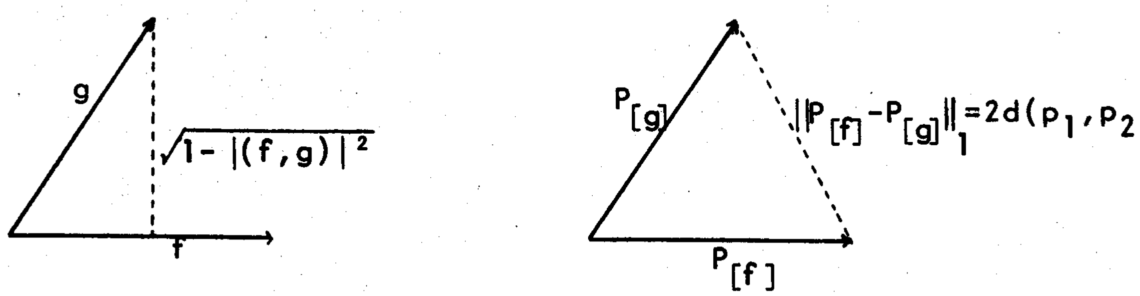
et donc l'image de $(P_{[f]} - W)^+$ aussi est unidimentionnelle. Cela montre (a). L'affirmation (b) est une conséquence de (a) et de la relation (7). ■

Lorsque les états p_1, p_2 sont tous les deux purs, leur distance prend une forme très simple:

Corollaire 1. Si p_1, p_2 sont deux états purs représentés par les vecteurs f et g, alors

$$d(p_1, p_2) = \sqrt{1-|(f,g)|^2}$$

Démonstration Il suffit de calculer les valeurs propres non nulles de $P_{[f]} - P_{[g]}$. Ce calcul a été fait par exemple dans {79} et donne le résultat: $\pm\sqrt{1-|(f,g)|^2}$.
Le corollaire découle aisement en substituant cette expression dans la relation (7). $\blacksquare$



Espace H

Espace $L_1(H)$

Figure 5

La distance entre deux états purs dans l'espace H et dans l'espace des états $L_1(H)$.

Comparons maintenant la distance $d(p_1, p_2)$ avec la distance de Gudder.

Théorème. (Gudder, {72} . La distance d_G entre deux états représentés par W_1 et W_2 est donnée par:

$$d_G(p_1,p_2) = \frac{\|W_1-W_2\|_1}{2+\|W_1-W_2\|_1}$$

En combinant le théorème de Gudder avec le théorème 1 on trouve

$$d_G\ (p_1,\ p_2)\ = \frac{d(p_1,p_2)}{1+d(p_1,p_2)}$$

En d'autres termes, la distance de Gudder est donnée par une fonction croissante de notre distance. Il s'ensuit que non seulement les deux distances produisent la même topologie, mais de plus, elles donneraient le même résultat à tout problème comportant une minimisation ou maximalisation de distance. Cela est plutôt surprenant, étant donné que, comme nous l'avons remarqué la définition de la distance d est fondée sur un aspect plutôt secondaire des états.

Pour terminer notre discussion sur les distances, nous démontrons la proposition concernant les valeurs de $|p_1\ (a) - p_2(a)|$ que nous avons mentionnée dans la sous-section précédente.

Proposition 2. Pour toute paire d'états p_1, p_2, l'image de la fonction définie sur l'ensemble de toutes les propositions par

$$\forall a \in \mathcal{L} : \quad f(a) = |p_1(a) - p_2(a)|$$

est l'intervalle $[0, \frac{\|W_1 - W_1\|_1}{2}]$ *tout entier.*

Preuve. Nous savons déjà, d'après le théorème 1 que

$$\max_{a\in\mathcal{L}} f(a) = d(p_1,p_1) = \frac{\|W_1-W_2\|_1}{2}$$

Soit maintenant $\{f_i\}_{i\in I}$ une base de $\overline{R(W_1-W_2)^+}$ et $\{g_j\}_{j\in J}$ une base de $\overline{R(W_1-W_2)^-}$ Nous noterons par R_o celui de ces deux sous espaces qui a la plus petite dimension, par exemple $R_o = \overline{R(W_1-W_2)^+}$. Nous pouvons alors supposer que $I \subseteq J$. L'idée centrale de cette démonstration consiste à faire "tourner" de façon continue R_o de sa position initiale jusqu'à ce qu'il tombe sur $\overline{R(W_1-W_2)^-}$

Pour tout $i \in I$ et tout $t \in [0,1]$ nous définissons:

$$f_i(t) = \sqrt{1-t^2}\; f_i + t\, g_i \;,\quad R(t) = \overline{[\{f_i(t)\}_{i\in I}]}\;,$$

$$E(t) = \sum_i P_{[f_i(t)]}$$

On trouve trivialement

$$(f_i(t), f_j(t')) = \delta_{ij}\left(\sqrt{(1-t^2)(1-t'^2)} + tt'\right) \tag{9}$$

et aussi

$$\|E(t)-E(t')\|^2 = \sup_{\|e\|=1} \|(E(t)-E(t'))e\|^2 =$$

$$= \sup_{\|e\|=1} \sum_i \|(P_{[f_i(t)]} - P_{[f_i(t')]})e\|^2 =$$

$$= \sup_{\|e\|=1} \sum_i \|(P_{[f_i(t)]} - P_{[f_i(t')]})(P_{[f_i]} + P_{[g_i]})e\|^2 \leq$$

$$\leq \sup_{\|e\|=1} \sum_i \|P_{[f_i(t)]} - P_{[f_i(t')]}\|^2 \|(P_{[f_i]} + P_{[g_i]})e\|^2 \tag{10}$$

Du corollaire 1 et de la proposition 1 on déduit la relation:

$$\|P_{[f_i(t)]} - P_{[f_i(t')]}\|^2 = 1 - |(f_i(t), f_i(t'))|^2$$

En substituant cela dans (10) et en utilisant (9) on trouve

$$\forall\, t,t' \in [0,1] : \|E(t)-E(t')\|^2 \leq 1 - \left(\sqrt{(1-t^2)(1-t'^2)} + tt'\right)^2$$

ce qui montre que la fonction $t \longrightarrow E(t)$ est continue. Par conséquent la fonction $t \longrightarrow h(t) := \mathrm{Tr}((W_1-W_2)\,E(t))$ aussi est continue et donc son image est un intervalle. Pour $t = 0$, $E(0)$ est le projecteur sur $R_o = \overline{R(W_1-W_2)^+}$, et donc $h(0) = \|W_1 - W_2\|_1/2$. Pour $t = 1$, $E(1)$ est le projecteur sur $R_1 \subsetneq \overline{R(W_1-W_2)^-}$, ce qui donne $h(1) < 0$. Si a_t est la proposition correspondant à $E(t)$, alors $f(a_t) = |h(t)|$.
La fonction $f(a_t)$ prend donc toutes les valeurs de 0 à $\|W_1-W_2\|_1/2$. ∎

3. APPLICATION DE LA DISTANCE A L'ETUDE DU PROBLEME D'INFERENCE STATISTIQUE.

3.1 *L'enoncé du problème.*

Etant donné la difficulté des problèmes d'interprétation que soulève le formalisme quantique, nous allons d'abord préciser notre langage concernant l'emploi du terme "état". Nous allons pour cela rester dans le cadre de la théorie quantique conventionnelle.

Les états sont en général considérés comme le résultat d'une préparation du système. Ainsi, selon la définition de J.M. Jauch, "un état est le résultat d'une série de manipulations physiques qui constituent la préparation de l'état" (cf. chapitre II, définition DJ_{14}, et aussi {27} . Il est bien évident que le concept d'état ainsi défini possède un caractère objectif: une préparation définit un état indépendamment de ce que l'observateur sait ou croit savoir. Nous appellerons donc dorénavant cet état "état objectif du système". Dans le formalisme Quantique, cet état est représenté, comme il est bien connu, par un opérateur statistique.

Toutefois les informations que l'on possède sur une préparation ne sont pas toujours suffisantes pour permettre de déterminer l'opérateur statistique qui représente l'état objectif. On essaye alors de choisir, parmi tous les opérateurs statistiques, celui qui représente "le mieux" l'information disponible. La spécification de ce choix constitue l'objet du problème de l'inférence statistique en Mécanique Quantique. L'opérateur statistique qui représente le mieux l'information disponible est souvent conçu comme le descripteur mathématique de "l'état subjectif" ou de "l'état des connaissances de l'observateur".

Tout comme en théorie classique, le problème de l'inférence statistique en mécanique quantique présente plusieurs aspects, en fonction de la nature de l'information disponible. Les spécificitésde la Mécanique Quantique nous obligent toutefois de faire une distinction supplémentaire. On peut en effet obtenir des informations sur l'état des systèmes produits par une préparation donnée de deux manières:

a) soit en effectuant des mesures sur d'autres systèmes produits par la même préparation et en tirant des conclusions concernant le système considéré,

b) soit en effectuant des mesures sur le système considéré lui-même. Dans ce cas l'état objectif n'est plus le même après la mesure.

En Mécanique Classique on ne fait pas cette distinction car on postule en général que l'on peut effectuer des mesures sur un système sans que son état soit modifié.

3.2 *Travail de Jaynes.*

E.T. Jaynes a été l'un des premiers à étudier un aspect du problème de l'inférence statistique en Mécanique Quantique. La question à laquelle il a essayé de répondre est la suivante: étant donné les valeurs moyennes $k_1,...k_m$ de certaines observables $C_1,...C_m$, il y a en général une infinité d'opérateurs statistiques W tels que la moyenne Tr (WC) des C_i dans l'état représenté par W soit k_i. Lequel parmi tous ces opérateurs statistiques représente-t-il le mieux l'information disponible?

Pour répondre à cette question, Jaynes {58} a utilisé le concept d'entropie quantique introduit par von Neumann{8}. Par analogie au cas classique, il a proposé le choix suivant: parmi tous les opérateurs statistiques pour lesquels les C_i ont les moyennes données k_i , il faut choisir celui qui maximise "l'entropie" Tr(-W log W). Ce choix est couramment utilisé aujourd'hui; il importe toutefois de dire que sa justification physique {87, 88} est bien plus faible que celle que nous avons exposée dans le chapitre précédent pour le cas classique.

3.3 *Modification de l'état subjectif dûe à une information supplémentaire.*

Nous allons maintenant employer le concept de distance pour étudier un aspect différent du problème de l'inférence statistique, qui est semblable à celui étudié par Kullback dans la théorie classique des probabilités. Supposons que notre état de connaissance sur un système physique soit W_o . En un moment donné nous acquérons de nouvelles données qui changent notre état de connaissance. Le problème est de déterminer l'état final W_1 en fonction de l'état initial W_o et de la nouvelle information. Dans ce qui suit, nous allons utiliser la "représentation" d'Heisenberg, et donc il n'y aura pas d'évolution dynamique de l'état dans le temps. Les états peuvent changer seulement lorsque notre information sur le système change.

Nous allons considerer deux cas différents, d'après la nature de la nouvelle information, i. e. suivant les cas (a) et (b) de la section 3.1.

3.3.1. Information provenant de mesures sur des systèmes autres que le système considéré.

Lorsque l'information provient des mesures effectuées sur des systèmes autres que le système considéré, nous pouvons enoncer le problème de la manière suivante: on considère des systèmes produits par une préparation donnée, et supposons que notre information sur cette préparation nous ait permis de choisir (d'une manière que nous ne spécifions pas) un opérateur statistique W_o qui représente l'état des connaissances à l'égard de ces systèmes. Afin de tester si oui ou non W_o représente aussi l'état objectif défini par la préparation considérée, nous pouvons mesurer certaines observables sur un grand ensemble d'autres systèmes produits par la même préparation. Cela fournit certaines données statistiques qui peuvent être identiques ou non aux prédictions calculées à partir de W_o . Si elles sont identiques alors on a une indication que W_o est proche de l'opérateur statistique décrivant l'état objectif, ou même qu'il peut être identique à celui-ci.. Par conséquent, l'on maintiendra W_o comme le représentant de l'état de connaissance. Si les données statistiques diffèrent des prédictions de W_o , l'on est obligé de choisir un autre W_1 comme représentant le nouvel état de connaissance. Le problème de l'inférence statistique est dans ce cas de trouver ce W_1 sur la base de W_o et des nouvelles données statistiques.

La solution que nous proposons à ce problème est fondée sur l'argument qualitatif suivant: Soit S l'ensemble de tous les opérateurs statistiques qui donnent des prédictions en accord avec les nouvelles données statistiques. Evidemment W_1 doit être l'un d'entre eux. Mais d'autre part W_1 doit tenir compte dans toute la mesure du possible,de l'information contenue dans W_0 Cette condition est remplie si W_1 , tout en appartenant à S, est choisi de telle façon qu'il donne des prédictions les plus proches possibles de celles données par W_0 ,et cela pour toutes les mesures imaginables. Selon la définition de la distance entre états (déf. 1) cela signifie que d (W_0 , W_1) devrait être le plus petit possible. En d'autres termes, W_1 devrait rendre minimale la distance d (W_0 , W) entre W_0 et un élément $W \in S$.

Certes, à cause de son caractère qualitatif, cet argument n'a pas la force d'une évidence. Il est toutefois tout-à-fait dans l'esprit des arguments utilisés habituellement dans l'inférence statistique classique. Nous élevons donc sa conclusion au rang d'un postulat:

<u>Postulat 1</u>. Soit W_1 , W_2 deux opérateurs statistiques représentant les états subjectifs avant et après l'acquisition d'une information qui consiste en certaines données statistiques. Soit encore S l'ensemble de tous les opérateurs statistiques donnant des prédictions en accord avec ces données statistiques nouvellement obtenues. Alors W_1 est l'élément de S qui satisfait à la relation

$$d\,(W_0 , W_1) = \inf_{W \in S} d(W_0, W) . \tag{11}$$

Le postulat 1 traduit en langage mathématique le problème physique soulevé plus haut. En effet, nous devons maintenant spécifier S et trouver celui qui, parmi ses éléments, satisfait à la relation (11).

La première partie de ce problème i. e. la détermination de S dans le cas d'une information indirecte, est simple. En effet, les données statistiques quantitatives sont en général de trois sortes:

(a) Elles peuvent être des valeurs moyennes, dispersions, etc... de certaines observables.

(b) des probabilités de certaines propositions

(c) des distributions de probabilité de certaines observables.

Toutes ces données peuvent être exprimées comme des valeurs moyennes de certains opérateurs autoadjoints. Dans le cas (a) cela est évident, puisque les dispersions sont exprimables comme des valeurs moyennes des carrés d'opérateurs. Dans le cas (b), il est suffisant de se rappeler que les propositions aussi sont représentées par des opérateurs (projecteurs). En ce qui concerne le cas (c), notons que si $C = \int \lambda dE_\lambda$ est la décomposition spectrale de l'opérateur C, alors la distribution de probabilité de C dans l'état W est donnée par la formule: $f(\lambda) = \mathrm{Tr}\,(WE_\lambda)$. Or Tr (AW) pour tout opérateur autoadjoint A est, par définition, la valeur moyenne de A dans l'état W. Par conséquent, la donnée de la distribution de probabilité de A équivaut à la donnée de toutes les valeurs moyennes des projecteurs E_λ .

On en déduit que, en effet, toute information indirecte peut être exprimée en donnant les valeurs moyennes k_i, $i \in I$ de certaines observables C_i. Par conséquent, l'ensemble S de tous les états compatibles avec une information indirecte est de la forme:

$$S = \{W \in \Sigma : \mathrm{Tr}\,(W C_i) = k_i,\ i \in I\} \qquad (12)$$

Σ étant l'ensemble des opérateurs statistiques et C_i des observables définies pour le système étudié.

La première partie du problème mathématique posé par le postulat, i. e. la détermination de l'ensemble S pour le cas d'une information indirecte, est donc résolue. La deuxième partie -la recherche de l'opérateur W dans S qui satisfait la relation 11 - est beaucoup plus difficile. En fait, nous avons établi seulement certains résultats partiels. Afin de faciliter le traitement mathématique, nous allons désormais supposer que C sont des opérateurs continus. Si tel est le cas, la structure de S est donée pár la proposition suivante:

Proposition 3. L'ensemble S défini par (12) est convexe et fermé (par rapport à la topologie définie par la norme $\|\cdot\|_1$*).*

Preuve. a) Pour tout W', W'' $\in$ S et tout $a, b \in R^+$ tels que $a + b = 1$, on a

$$\forall i \in I : \quad Tr((aW'+bW'')C_i) = aTr(W'C_i) + bTr(W''C_i) = k_i$$

Par conséquent, $aW' + bW'' \in S$, et donc S est convexe.

b) les fonctions

$$f_i : L_1(H) \ni W \longrightarrow Tr(WC_i) \in \mathbf{C}$$

étant continues {86} , il s'ensuit que les ensembles $f_i^{-1}(k_i)$ sont fermés. Or $S = \bigcap_I f_i^{-1}(k_i)$, et donc S est fermé. $\blacksquare$

Le résultat de la proposition permet d'utiliser la théorie de la meilleure approximation d'un élément d'un espace de Banach par des éléments d'un sous ensemble fermé et convexe {89 - 91} . Dans notre problème, l'espace de Banach sera $L_1(H)$ et nous aurons à trouver, selon notre postulat, la meilleure approximation de W_0 par un élément W_1 de S, par rapport à la norme $\|\cdot\|_1$ (cf. théorème 1). La théorie de la meilleure approximation pour des espaces de Banach, est loin d'être suffisamment développée; pourtant, elle fournit des critères généraux caractérisant W_1 , qui sont bien plus simples que la condition (11). Ces critères sont donnés dans l'appendice mathématique de ce chapitre par les théorèmes suivants: pour un

S général (cf. relation (12)) par les théorèmes 3 et 4; pour le cas spécial d'un S défini par la valeur moyenne d'<u>une</u> observable, cf. théorème 5; enfin, ces critères nous permettront, dans des cas spéciaux, de déterminer complètement W_1 (cf. théorème 6).

3.3.2. <u>Information provenant d'une mesure sur le système considéré.</u>

Nous arrivons maintenant au second des deux cas considéré dans la section 3.1, à savoir celui dans lequel on obtient une information en effectuant une mesure sur le système lui-même. En fait, nous allons étudier seulement un cas très particulier, notre but étant de démontrer, dans un cadre précis, le "postulat de projection" de la Mécanique Quantique. Ce cas particulier peut être énoncé ainsi. Supposons que l'état qui représente nos connaissances sur un système physique (i. e. l'état subjectif) soit à un certain moment W_o . On effectue sur ce système une mesure de première espèce, ce qui signifie dans la terminologie habituelle, {8} , que le système n'est pas détruit par la mesure et qu'une seconde mesure éventuelle juste après la première donnerait le même résultat. Le problème est de trouver l'état subjectif W_1 après la mesure.

La différence évidente avec le cas étudié dans la section précédente, est qu'içi la mesure modifie la préparation du système et, par conséquent, l'état objectif du système n'est plus le même après la mesure.

Le problème possède une solution triviale seulement dans le cas d'une mesure complète ,i.e. quand il s'agit d'une mesure simultanée d'un ensemble complet d'observables qui commutent. L'état W_1 est alors représenté par l'unique vecteur propre commun à ces observables qui correspond au résultat trouvé, et il est indépendant de l'état de départ. Cependant, le cas contraire est réalisé le plus souvent: toute mesure d'une observable à spectre continu -en particulier de la position, quantité de mouvement, énergie d'une particule libre, etc...- est nécessairement incomplète. C'est aussi le cas pour les observables à spectre dégénéré, lorsqu'on mesure une seule observable. Il existe alors un nombre infini d'opérateurs statistiques compatibles avec le résultat de la mesure, et le formalisme quantique généralement utilisé {8} ne contient aucun critère pour déterminer W_1 de façon unique.

Le choix de l'état W_1 après la mesure dépend, bien sûr, du montage expérimental utilisé. Nous allons toutefois faire l'hypothèse que l'on ignore complètement le montage, et que notre information est constituée seulement par la spécification de l'observable mesurée et du résultat observé. Cette hypothèse n'est ni trop restrictive ni si abstraite qu'elle pourrait sembler. En effet, certaines informations sur le montage expérimental peuvent être incluses dans le résultat de la mesure. Par exemple, la position de l'appareil dans l'espace implique une certaine localisation de la particule, et cela peut être considéré comme une partie du résultat. D'autre part le formalisme quantique ne se réfère jamais au montage expérimental dans ses principes fondamentaux.

Nous allons maintenant proposer un choix pour l'opérateur statistique représentant nos connaissances après la mesure. Nos arguments seront analogues à ceux utilisés dans la section précédente. Tout d'abord, nous remarquons que le résultat de la mesure d'une grandeur K peut être exprimé par le fait qu'une certaine proposition b, à savoir "la valeur de K appartient au Borélien B ", a été trouvée vraie. Etant donné que la mesure est de première espèce, on sait qu'une seconde mesure donnerait le même résultat. Par conséquent, la probabilité de la proposition b pour l'état final W_1 , est égale à 1. Par l'isomorphisme de von Neumann la proposition b correspond à un projecteur E, et la relation (2) implique que

$$\mathrm{Tr}\,(W_1 E) = 1 \qquad (13).$$

Nous avons trouvé cette expression en tenant compte de la nature de l'observable K et du résultat observé (utilisés pour la détermination de la proposition b et du projecteur E) aussi bien que du fait que la mesure a été de première espèce. D'autre part, on sait aussi que l'état subjectif avant la mesure était W_0. Par conséquent, W_1 doit tenir compte, dans la mesure du possible, de l'information contenue dans W_0. Par un argument identique à celui utilisé dans la sous-section précédente, on déduit que cela sera réalisé si W_1, tout en satisfaisant à la condition (19), est le plus proche possible de W_0 au sens de la distance d. Nous proposons donc le postulat suivant:

Postulat 2. Supposons qu'une mesure de première espèce soit effectuée sur un système quantique. Supposons que l'opérateur statistique représentant l'état subjectif avant la mesure ait été W_0, et que le résultat de la mesure soit représenté par le projecteur E. Alors l'opérateur statistique W_1 qui représente l'état subjectif après la mesure est le plus proche de W_0 parmi tous les opérateurs statistiques satisfaisant à (13). En d'autres termes, on aura:

$$Tr(W_1E) = 1 \quad \text{et} \quad d(W_0,W_1) = \min\{d(W_0,W): Tr(WE) = 1\}. \qquad (14)$$

Si l'on accepte ce postulat sur la base des arguments physiques exposés, alors on peut <u>démontrer</u> la proposition qui est souvent appelée "postulat de projection" {92-94}:

Théorème (Postulat de projection) Supposons qu'une mesure de première espèce soit effectuée sur un système quantique. Si l'état du système avant la mesure était un état pur représenté par un vecteur f, le résultat de la mesure étant représenté par le projecteur E, alors l'état subjectif après la mesure sera lui aussi pur, et il sera représenté par la projection normalisée* $\frac{Ef}{\|Ef\|}$.

La démonstration de ce théorème se trouve en appendice (cf. théorème 6).

---

* Certains auteurs appellent cette proposition "postulat de projection généralisé", réservant le terme "postulat de projection" au cas spécial où la mesure est complète.

4. CONCLUSION

Nous avons vu comment le concept de distance entre les états quantiques peut être utilisé dans l'étude du problème de l'inférence statistique. En particulier, nous avons montré que son emploi permet d'aboutir à une justification, dans un cadre conceptuel précis, du "postulat de projection". Cependant, plusieurs problèmes restent ouverts, conduisant chacun à une nouvelle voie de recherche.

Le premier problème est d'ordre purement mathématique. Nous avons vu en effet que l'inférence statistique conduit, dans les cas que nous avons examinés, à un problème d'approximation dans un espace de Banach: il s'agit de trouver un élément $W_1 \in S$ qui soit le plus proche possible d'un opérateur statistique donné W_o, S étant un ensemble d'opérateurs statistiques défini à partir de l'information disponible. Or notre étude n'aboutit à une solution complète que pour un W_o correspondant à un état pur, et pour un S spécial (cf. théorème 6). Il serait donc intéressant de chercher la solution W_1 pour des états initiaux W_o non purs, et pour des ensembles S plus généraux.

La deuxième question est plus fondamentale et concerne la justesse de notre définition de distance, face au but que nous nous sommes fixés. Il est en effet certain que la solution du problème de l'inférence statistique, comme nous l'avons posé dans la section 3.3, implique une minimisation d'une certaine distance. La distance $d\ (p_1, p_2)$ introduite par la définition 1 est-elle la plus appropriée à ce but? En particulier, est-elle la plus appropriée parmi la classe infinie de distances que nous avons exhibée dans l'annexe 2 (cf. théorème 2)? Seule une étude plus approfondie permettrait de trancher.

APPENDICE AU CHAPITRE V

APPROXIMATIONS DANS L' ESPACE L_1 (H).

Tout au long de cet appendice, S sera un sous-ensemble fermé et convexe de Σ et W_0 un opérateur statistique n'appartenant pas à S. Notre but est de donner des critères qui caractérisent "l'élément de meilleure approximation", i.e. l'opérateur statistique inconnu W_1 qui remplit la condition (11).

Nous ferons usage du résultat fondamental de la théorie de l'approximation des ensembles convexes dans des espaces normés, dû à I. Singer {89,90} :

Théorème 2. (I. Singer): Soit E un espace normé, G un sous-ensemble convexe, $x \in E \setminus \overline{G}$ et $g_0 \in G$. Alors $\|x - g_0\| = \inf\{\|x-g\| : g \in G\}$ si et seulement si il existe un $f \in E^*$ (le dual topologique de E) tel que

$$\|f\| = 1, \qquad Re\,(f(x-g_0)) = \|x-g_0\|,$$

$$Re\,(f(g_0-g)) \geqslant 0, \quad \forall g \in G .$$

Dans le cas qui nous intéresse, le théorème de Singer implique le résultat suivant:

Théorème 3: Si $W_1 \in S$, alors une condition nécessaire et suffisante pour avoir

$$\|W_0-W_1\|_1 = \min\{\|W_0-W\|_1,\ W \in S\} \tag{15}$$

est la suivante: Il existe un opérateur hermitique A tel que

$$\|A\|=1, \qquad Tr(A(W_0-W_1)) = \|W_0-W_1\|_1 \tag{16}$$

et

$$Tr\,(A\,(W_1 - W)) \geqslant 0, \ \forall\, W \in S \tag{17}$$

Preuve Nous appliquons le théorème de Singer au cas

$E = L_1(H)$, $G = \overline{G} = S$, $x = W_o$, $g_o = W_1$.
Il est bien connu {85, 86} que le dual de L_1 (H) est isomorphe à L (H). La dualité est établie via la forme bilinéaire continue

$$L(H) \times L_1(H) \ni (B,V) \longrightarrow Tr\,(BV) \in \mathbb{C}$$

où
$$|Tr\,(BV)| \leq \|B\| \cdot \|V\|_1$$

Par conséquent, si la relation (15) est vraie, alors il existe un élément $B \in L\,(H)$ tel que:

$$\|B\| = 1, \qquad Re\,[Tr\,(B(W_o - W_1))] = \|W_o - W_1\|_1$$

et
$$Re\,[Tr\,(B(W_1 - W))] \geq 0, \ \forall\ W \in S$$

En posant $A = \frac{B+B^*}{2}$,on trouve:

$$Tr\,(A(W_o - W_1)) = Re\,[Tr\,(B(W_o - W_1))] = \|W_o - W_1\|_1$$

$$Tr\,(A(W_1 - W)) = Re[\,Tr\,(B(W_1 - W))] \geq 0$$

$$\|A\| \leq \frac{\|B\| + \|B^*\|}{2} = 1$$

et aussi

$$\|W_o - W_1\|_1 = Tr(A(W_o - W_1)) \leq \|A\| \cdot \|W_o - W_1\|_1 \Longrightarrow \|A\| \geq 1$$

donc $\|A\| = 1$. ∎

Le problème de la recherche de l'élément de meilleure approximation est donc transformé, via le théorème 3, en une paire de problèmes plus simples. Nous devrions d'abord trouver la solution générale des équations (16) et ensuite examiner s'il en existe une qui satisfait la condition (17). Nous établissons maintenant un théorème qui résout le premier de ces problèmes.

Théorème 4. *La solution générale de la paire d'équations (16) par rapport à l'opérateur hermitique A est la suivante:*

$$A = F - F' + B \tag{18}$$

F et F' étant, respectivement, les projecteurs sur $\overline{R(W_0-W_1)^+}$ et $\overline{R(W_0-W_1)^-}$ et B étant un opérateur hermitique dont l'image est orthogonale à $R(W_0 - W_1)$ et tel que $\|B\|<1$.

Démonstration. Nous montrons d'abord que si B a les propriétés décrites plus haut, alors A défini par (18) satisfait aux relations (16).

Si $x \in R(F)$ alors $Bx = 0$ et donc $(x, Ax) = (x,Fx) = \|x\|^2$ ce qui montre que $\|A\| \geqslant 1$. D'autre part pour tout $x \in H$, écrivons $x = x_1 + x_2 + x_3 + x_4$, où $x_1 \in R(F)$, $x_2 \in R(F')$, $x_3 \in \overline{R(B)}$ et $x_4 \perp R(F) \oplus R(F') \oplus \overline{R(B)}$. Alors

$$|(x,Ax)| = |(x_1,Fx_1)+(x_2,F'x_2)+(x_3,Bx_3)| \leqslant$$

$$\leqslant \|F\|.\|x_1\|^2 + \|F'\|.\|x_2\|^2 + \|B\|.\|x_3\|^2 \leqslant \|x\|^2 ,$$

ce qui montre que $\|A\| = 1$. En outre, si l'on pose pour simplifier $\tilde{W} = W_0 - W_1$, alors

$$Tr(A\tilde{W}) = Tr((F-F')\tilde{W})$$

Or $\tilde{W} = (F - F')|\tilde{W}|$ est la décomposition polaire de $\tilde{W}$ et donc

$$Tr((F-F')\tilde{W}) = Tr|\tilde{W}| = \|\tilde{W}\|_1$$

ce qui montreavec la relation précédente, que les conditions (16) sont satisfaites.

Montrons maintenant que toute solution de (16) a la forme (18). Si $\tilde{W} = \sum_n p_n P_{[f_n]}$ est la décomposition spectrale de $\tilde{W}$ alors (16) implique

$$\sum_n p_n (f_n, Af_n) = Tr(\tilde{W}A) = \|\tilde{W}\|_1 = \sum_n |p_n| \tag{19}$$

Puisque $|(f_n, Af_n)| \leqslant \|A\| = 1$, on en déduit que

$$p_n > 0 \Longrightarrow (f_n,Af_n) = 1 \text{ , et } p_n < 0 \Longrightarrow (f_n,Af_n) = -1 \tag{20}$$

Soit $A = \int_{-1}^{1} \lambda \, dE_\lambda$ la décomposition spectrale de A, on trouve alors:

$$p_n > 0 \Longrightarrow 1 = (f_n, Af_n) = \int_{-1}^{+1} \lambda \, d(f_n, E_\lambda f_n) \leqslant \int_{-1}^{+1} d(f_n, E_\lambda f_n) = 1$$

$$\Longrightarrow (f_n, E_\lambda f_n) = 0 \, , \forall \lambda < 1 \Longrightarrow Af_n = f_n$$

et, de même,

$$p_n < 0 \Longrightarrow Af_n = -f_n$$

Par conséquent, en posant par définition $B = A - (F - F')$, on a $Bf_n = 0$ chaque fois que $p_n \neq 0$. On en déduit que $R(B)$ est orthogonal à $\overline{R(F-F')} = \overline{R(\widetilde{W})}$. Puisque $A = B + (F - F')$ et $\|A\| = 1$ cette orthogonalité implique que $\|B\| \leqslant 1$. $\blacksquare$

La proposition suivante est utile pour la démonstration de l'unicité de l'élément de meilleure approximation dans les cas qui nous intéressent.

__Proposition 4__. Si l'équation (15) qui caractérise les élements de meilleure approximation a deux solutions par rapport à W_1, par exemple W_1, W'_1, alors tout A satisfaisant aux conditions (16) et (17) satisfait aussi aux mêmes conditions avec W'_1 à la place de W_1.

__Preuve__. Supposons que W_1, W'_1 soient des solutions de (15) et que A satisfasse aux (16), (17). Alors on a

$$\|W_o - W_1\|_1 = \|W_o - W'_1\|_1$$

et, en vertu de (17),

$$\mathrm{Tr}\,(AW_1) \geqslant \mathrm{Tr}\,(AW'_1)$$

En remplaçant ces expressions dans (16) on trouve:

$$\mathrm{Tr}\,(A(W_o - W'_1)) \geqslant \mathrm{Tr}\,(A(W_o - W_1)) = \|W_o - W_1\|_1 = \|W_o - W'_1\|_1$$

ce qui implique facilement

$$\mathrm{Tr}\,(A(W_o-W'_1)) = \| W_o-W'_1\|_1 \;, \quad \mathrm{Tr}\,(AW'_1) = \mathrm{Tr}(AW_1)\;. \blacksquare$$

Résumons les résultats obtenus jusqu'içi. Le théorème 3 montre que W_1 est un élément de meilleure approximation s'il existe une solution commune A aux équations (16) qui satisfait aussi à la condition (17). Le théorème 4 donne la solution générale de (16). Il reste donc la deuxième partie du problème, i. e. de trouver pour quel W_1 il existe, parmi les solutions de (16), une solution qui satisfait aussi à (17). Cela s'est avéré beaucoup plus difficile. Nous avons seulement établi certains résultats partiels, remplaçant la condition (17) par des conditions plus simples quand on passe du cas général à des cas plus particuliers. Ces nouvelles conditions nous permettront, par exemple, de démontrer le postulat de projection.

Nous avons considéré jusqu'à maintenant un ensemble d'états S très général, qui était seulement convexe et fermé. Nous ferons dorénavant une particularisation, en définissant S comme l'ensemble de tous les états pour lesquels la valeur moyenne d'une observable continue C est un nombre k, i. e.

$$S = \{ W \in \Sigma : \mathrm{Tr}\,(WC) = k \}$$

Nous avons établi dans l'annexe 1 un théorème qui montre que tout opérateur statistique dans S peut être écrit comme un mélange d'états purs qui appartiennent tous à S. Ce résultat nous conduit à simplifier considérablement la condition (17):

Théorème 5 Soit $S = \{ W \in \Sigma: \quad \mathrm{Tr}\,(WC) = k \}$ et $W_1 \in S$. Alors (17) est équivalente à chacune des conditions suivantes:

a) $\mathrm{Tr}\,(W_1C) \geq (f, Af), \forall P_{[f]}\ S$

b) W_1 peut être ecrit comme un mélange d'états purs $P_{[f_i]}$ qui rendent maximale l'expression (f, Af) pour $P_{[f]} \in S$.

<u>Preuve</u> L'implication (17) $\Rightarrow$ (a) est évidente.

(a) $\Rightarrow$ (17) : D'après le théorème 3, annexe I, tout $W \in S$ peut être écrit sous la forme: $W = \sum_j a_j P_{[g_j]}$, où $P_{[g_j]} \in S$. La condition (a) implique:

$$\forall j : \quad Tr(AW_1) \geqslant (g_j, Ag_j) \Longrightarrow Tr(AW_1) \geqslant Tr(AW)$$

i. e. la condition (17).

(a) $\Longrightarrow$(b) : En utilisant de nouveau le théorème 3, annexe 1, on peut écrire: $W_1 = \sum_i \lambda_i P_{[f_i]}$ avec $P_{[f_i]} \in S$. De la condition (a) on déduit

$$\forall P_{[f]} \in S : \quad \sum_i \lambda_i (f_i, Af_i) \geqslant (f, Af)$$

ce qui implique aisément

$$\forall i, \forall P_{[f]} \in S : \quad (f_i Af_i) \geqslant (f, Af).$$

(b)$\Longrightarrow$(a) : évident. ∎

En combinant maintenant les théorèmes 3, 4 et 5 on obtient deux critères caractérisant l'élément de meilleure approximation. Le premier est la condition (a) plus haut, qui doit être remplie par l'élément de meilleure approximation W_1 et l'opérateur A, défini par la relation (18). Il est beaucoup plus simple que la condition (17), puisqu'il fait intervenir une comparaison de W_1 avec des états <u>purs</u> seulement. Le deuxième critère, i. e. la condition (b), n'inclue aucune comparaison; il exprime une relation entre A, qui dépend de W_1, et W_1 lui-même.

Le théorème suivant donne la solution du problème de meilleure approximation lorsque E est un projecteur et W_1 un état pur. Il est utilisé dans la section V.32 pour la démonstration du postulat de projection.

Théorème 6. *Soit E un projecteur quelconque et $S = \{ W \in \Sigma : \operatorname{Tr}(WE) = 1 \}$. Soit encore $W_0 = P_{[f]}$ où $Ef \neq 0$. Alors le seul opérateur statistique W_1 qui satisfait à la condition*

$$\|W_0 - W_1\|_1 = \min\{ \|W_0 - W\|_1 : W \in S \} \quad \text{et} \quad W_1 \in S$$

est le projecteur $W_1 = P_{[g]}$ où $g = \frac{Ef}{\|Ef\|}$.

Preuve Le théorème est une application immédiate des résultats précédents. Posons en effet:

$$W_1 := P_{[g]}, \; B := -(E - P_{[g]}) \; , \; A := F - F' + B$$

F et F' étant les projecteurs sur les sous-espaces unidimensionnels qui correspondent aux valeurs propres $\pm\sqrt{1-(f,g)^2}$ de $P_{[f]} - P_{[g]}$ (cf. corollaire 1). D'après le théorème 4, W_1 satisfait à la relation (16). D'autre part, pour tout $h \in R(E)$ on a:

$$(h,(F-F')h) = (h,E(F-F')Eh) = \left(h, \frac{E(P_{[f]}-P_{[g]})Ef}{\sqrt{1-|(f,g)|^2}}\right) =$$

$$= -\sqrt{1-(f,g)^2}\,(h,P_{[g]}h)$$

et donc

$$(h,Ah) = -1 + (1-\sqrt{1-(f,g)^2})(h,P_{[g]}h)$$

Par conséquent, (h, Ah) prend sa valeur maximale sur $R(E)$ pour $h = g$. D'après le théorème 5, W_1 satisfait à (17), et d'après le théorème 3 il satisfait la relation (15). Enfin, $W_1 = P_{[g]}$ est l'unique solution de (15) à cause de la proposition 4 et du fait que (h, Ah) est maximale sur $R(E)$ seulement pour $h = g$. ∎

ANNEXE I

PROPRIETES DES MELANGES D'ETATS NON ORTHOGONAUX

1. INTRODUCTION

En Théorie Quantique on utilise habituellement des états de deux sortes : les états purs, représentés par des vecteurs normalisés (ou par de sous-espaces unidimentionnels), et les mélanges, représentés par des opérateurs densité. En général les mélanges sont composés d'états purs orthogonaux. Toutefois, des mélanges d'états non orthogonaux apparaissent souvent dans des situations physiques {95} et mathématiques {95} . D'autre part, ils possèdent des propriétés intéressantes, qui concernent surtout l'aspect informationnel de la théorie quantique ; les démonstrations de ces propriétés donnés dans la littérature sont excéssivement longues et laborieuses. La raison de cette difficulté dans les démonstrations est sans doute l'absence d'une caractérisation simple et unitaire de tels mélanges. En dimension finie, une telle caractérisation a été donnée, par exemple, par E. T. Jaynes {58} . Malheureusement, on ne peut pas passer au cas infini par un simple passage à la limite.

L'objet de cette annexe est de donner l'extension au cas infini de cette caractérisation des mélanges d'états non orthogonaux. En particulier, nous allons donner la loi de formation de tels mélanges. Nos résultats nous permettront de donner des démonstrations beaucoup plus simples et plus générales des résultats déjà connus dans le cas fini, et d'établir aussi des résultats nouveaux.

2. PROPRIETES GENERALES.

Tout au long de cette section, nous ferons usage des notations utilisées dans la section V.2.2. Tous les ensembles d'indices sont égaux à N ou à { 1,2,...n} pour un certain n.

Nous appelons mélange toute somme finie au dénombrable $\sum_i a_i P_{[f_i]}$, où $a_i \geqslant 0$, $\sum_i a_i = 1$ et f_i sont des vecteurs normalisés. Il est bien connu qu'une telle somme converge toujours vers un opérateur densité W, par rapport à la norme $||\cdot||_1$ (et aussi par rapport à la norme $||\cdot||$ des opérateurs continus).

2.1. *Propriétés des images des mélanges*

Proposition 1. Si $W = \sum_i a_i P_{[f_i]}$, $\sum_i a_i = 1$,

$a_i > 0$, alors $\overline{R(W)} = \overline{[f_1, f_2,...]}$.

Preuve. Pour tout $g \perp \overline{R(W)}$ on a :

$$0 = (g, Wg) = \sum_i a_i \, |(g, f_i)|^2$$

ce qui implique $g \perp f_i$ pour tout i et tout $g \perp \overline{R(W)}$. Il s'ensuit que $f_i \in \overline{R(W)}$. D'autre part, on a évidemment $\overline{R(W)} \subseteq \overline{[f_1,f_2,...]}$. Donc $\overline{R(W)} = \overline{[f_1,f_2,...]}$

Corrolaire 1. *Si W, W' sont des opérateurs-densité, alors* $\overline{R(W + W')} = \overline{\overline{R(W)} + \overline{R(W')}}$.

Preuve. Il est toujours possible d'écrire W, W' sous forme de mélanges $W = \sum_i a_i P_{[f_i]}$, $W' = \sum_j b_j P_{[g_j]}$,

avec $a_i > 0$, $b_j > 0$ (nous pouvons faire usage, par exemple, de la décomposition spectrale). Selon la proposition 1, il est donc suffisant de démontrer que

$$\overline{[f_1, f_2,...,g_1,g_2,..]} = \overline{\overline{[f_1, f_2,...]} + \overline{[g_1, g_2,..]}}$$

Or cette égalité est évidemment vraie. ∎

2.2. *Passage d'un mélange à un autre*

Soit W un opérateur densité. En écrivant sa décomposition spectrale, $W = \sum_i \lambda_i \, P_{[g_i]}$, on obtient un mélange d'états orthogonaux. Nous allons démontrer que tous les autres mélanges correspondants à W peuvent être obtenus de celui-là via un opérateur partiellement isométrique, pourvu que la dimension de l'espace d'Hilbert H soit suffisamment grande (sinon, on peut ajouter des dimensions auxiliaires à H). Nous rappelons qu'un opérateur U s'appelle partiellement isométrique si sa restriction à un sous-espace H_o est isométrique tandis que sa restriction sur le complément orthogonal de H_o s'annule. Si U est partiellement isométrique, alors UU^+ est le projecteur sur R(U) et U^+U le projecteur sur H_o. Nous pouvons maintenant établir le théorème :

<u>Théorème 1</u>. *Soit* $W = \sum_{i \in I} \lambda_i \, P_{[g_i]}$, g_i étant une base de vecteurs propres de $W, \lambda_i \geqslant 0$ et $\sum_i \lambda_i = 1$. On a alors :

a) Si U est partiellement isométrique et tel que $R(U^+U) \supseteq R(W)$, *alors on peut définir un mélange* $\sum_{i \in I} a_i \, P_{[f_i]}$ *par*

$$\forall i \in I : a_i = \|\sqrt{W} U^+ g_i\|^2, \qquad f_i = \begin{cases} \dfrac{\sqrt{W}\, U^+ g_i}{\sqrt{a_i}}, & \text{si } a_i \neq 0 \\ \text{un vecteur normalisé quelconque,} & \\ \text{si } a_i = 0 & \end{cases} \tag{1}$$

et on aura : $W = \sum_i a_i \, P_{[f_i]}$

b) Réciproquement, si $W = \sum_{j \in J} a_j \, P_{[f_j]}$, $a_j \geqslant 0$, *et si* $\dim(H) \geqslant \operatorname{card} J$, *alors il existe un opérateur partiellement isométrique U tel que*

$$\forall j \in J : \sqrt{a_j}\, f_j = \sqrt{W} \; U^+ g_j$$

<u>Démonstration</u>.

a) Si U est partiellement isométrique et tel que $R(U^+U) \supseteq R(W)$, alors la restriction de U^+U sur $\overline{R(W)}$ est l'opérateur identité. Si maintenant a_i et f_i sont définis comme dans (1), alors

$$\sum_i a_i = \sum_i \| \sqrt{W} \; U^+ g_i \|^2 = \sum_i (g_i , U\, W\, U^+ g_i) =$$

$$= \operatorname{Tr}(UWU^+) = \operatorname{Tr}(WU^+U) = \operatorname{Tr} W = 1$$

Par conséquent, $W' = \sum_i a_i P_{[f_i]}$ est un opérateur densité bien défini et l'on a, pour tout k, l $\in$ I :

$$(g_k , W'g_l) = \sum_i (g_k , f_i) a_i (f_i , g_l) =$$

$$= \sum_i (g_k, \sqrt{W} U^+ g_i)\cdot(\sqrt{W} U^+ g_i , g_l) =$$

$$= \sqrt{\lambda_k \lambda_l} \sum_i (Ug_k, g_i) (g_i, Ug_l) = \lambda_k \delta_{kl} = (g_k, Wg_l)$$

Il s'ensuit que $W = W'$.

b) Soit $W = \sum_{j \in J} a_j P_{[f_j]}$, $a_j \geqslant 0$, $J \subseteq I$. On peut supposer, sans nuire à la généralité, que J = I (il suffit d'ajouter certains termes $a_j = 0$ à cette somme). Définissons maintenant un opérateur linéaire V par $Vg_i = \sqrt{a_i}\ f_i$. Il est facile de voir que V est continu, de norme 1, et il peut donc s'étendre sur tout l'espace H par continuité.
En fait,

$$\forall g \in H : \|Vg\| = \| \sum_i \sqrt{a_i}\ (g,g_i) f_i \| \leqslant \sum_i \sqrt{a_i}\ |(g,g_i)| \leqslant$$

$$\leqslant \sqrt{(\sum_i a_i) (\sum_i |(g,g_i)|^2)} = \| g \|$$

D'autre part, pour tout f, g $\in$ H on a :

$$(f, VV^+g) = \sum_i a_i (f,f_i).(f_i,g) = (f,Wg)$$

et par conséquent, $VV^+ = W$. Il s'ensuit que la décomposition polaire de V^+ s'écrit : $U \sqrt{W}$, pour un certain opérateur U partiellement isométrique. Cela implique que $V = \sqrt{W}\ U^+$ et achève la démonstration. ■

Le théorème 1 a deux corrolaires immédiats donnant des conditions nécessaires et suffisantes pour qu'un état pur participe à un mélange qui correspond à un opérateur densité W donné.

<u>Définition 1</u>. Nous dirons que l'état pur $P_{[g]}$ est candidat pour l'inclusion dans un mélange correspondant à un opérateur densité W, ssi il existe $a > 0$, $a_i \geqslant 0$ et des

états purs $P_{[f_i]}$ tels que $W = a\,P_{[g]} + \sum_i a_i\,P_{[f_i]}$.

<u>Corollaire 2</u>. *Un état pur* $P_{[g]}$ *est candidat à l'inclusion dans un mélange correspondant à* W *ssi* $g \in R(\sqrt{W})$.

<u>Preuve</u>. Conséquence immédiate du théorème 1. $\blacksquare$

<u>Définition 2</u>. Nous dirons qu'un mélange $\sum_i a_i\,P_{[f_i]}$ est irréductible ssi $\forall i : f_i \notin \overline{[f_1, \dots f_{i-1}, f_{i+2}, \dots]}$.

<u>Corollaire 3</u>. *Un état pur* $P_{[g]}$ *est candidat à l'inclusion dans un mélange irréductible correspondant à* W *ssi* $g \in R(W)$.

<u>Preuve</u>.

a) Si g participe à un mélange irréductible correspondant à W, alors il existe des nombres positifs $a, a_1, a_2, \dots$ et des vecteurs normalisés $f_1, f_2, \dots$ tels que $W = a\,P_{[g]} + \sum_i a_i\,P_{[f_i]}$ avec $g \notin \overline{[f_1, f_2, \dots]}$. Soit maintenant un vecteur h tel que $h \in \overline{[g, f_1, f_2, \dots]}$ et $h \perp \overline{[f_1, f_2, \dots]}$.
On voit alors facilement que $Wh = a\,(g,h)\,g$, i.e. $g \in R(W)$.

b) Supposons maintenant que $g \in R(W)$. Il existe donc $f \in H$ tel que $g = Wf$. Posons $a = 1 / \|\sqrt{W} f\|^2$. Alors

$$\forall e \in H : (e, a\,P_{[g]}\,e) = a\,|(g,e)|^2 = a \cdot |(Wf,e)|^2 \leq$$

$$\leq a\,\|\sqrt{W}f\|^2 .\ \|\sqrt{W}e\|^2 = (e, We)$$

Par conséquent, l'opérateur $T = W - a\,P_{[g]}$ est positif.
D'autre part, $Tf = 0$ et donc $f \perp \overline{R(T)}$. Puisque $(f,g) = \|\sqrt{W} f\|^2 \neq 0$, on en déduit que $g \notin \overline{R(T)}$. Soit maintenant $T = \sum_i \lambda_i\,P_{[f_i]}$ la décomposition spectrale de T. Alors le mélange

$$a\,P_{[g]} + \sum_{\lambda_i \neq 0} \lambda_i\,P_{[f_i]}$$

est irréductible et égal à W. $\blacksquare$

<u>3. APPLICATIONS</u>

La caractérisation des mélanges d'états non orthogonaux donnée par le théorème 1 peut-être utilisée chaque fois que de tels mélanges interviennent. Nous en donnons deux applications illustratives.

3.1. *Entropie généralisée des opérateurs densité*

Soit h une fonction continue et concave sur l'intervalle [0,1] , telle que h(0) = h(1) = 0. Nous entendons ici par "concave" une fonction qui possède la propriété :

$$\forall\, x, y \in [0,1] \;:\; 2h\left(\frac{x+y}{2}\right) \geqslant h(x) + h(y)$$

Pour tout opérateur densité W, A. Uhlmann {97} a étudié les propriétés de l'expression Tr(h(W)).
Cette expression est souvent appelée "entropie généralisée de W", puisque pour h(x) = -x logx, Tr(h(W)) est égale à l'entropie de Von Neumann - Shannon. Une propriété triviale de Tr(h(W)) est que

$$Tr(h(W)) \leqslant \sum_i h\,((x_i, Wx_i)) \tag{3}$$

pour toute base $\{x_i\}_{i \in I}$ de H.

L'entropie généralisée a aussi d'autres propriétés, beaucoup moins triviales. Une parmi elles est la suivante, qui fait usage des mélanges d'états non orthogonaux :

$$\left(W = \sum_i a_i\, P_{[f_i]} \;,\; a_i \geqslant 0\right) \Longrightarrow Tr(h(W)) \leqslant \sum_i h(a_i) \tag{4}$$

La démonstration donnée par Uhlmann {97} de la propriété (4) est valable seulement dans un espace H de dimension finie et pour un mélange $\sum_i a_i\, P_{[f_i]}$ qui contient un nombre fini de composantes. Une démonstration pour H de dimension infinie est donnée par Lanford et Robinson {98} mais de nouveau on a considéré seulement des mélanges finis et en plus on s'est restreint au cas h(x) = -xlogx. Enfin, une démonstration dans le cas général a été donnée par A. Wehrl {96} . Malheureusement, la démonstration de Wehrl est erronée*.

Toutes ces démonstrations sont très compliquées. Or le théorème 1 nous permet de donner une démonstration qui allie la généralité à la simplicité :

* La série utilisée dans {96} , p. 231, première ligne, ne converge pas en général.

Théorème 2. *Si* $W = \sum_i a_i P_{[f_i]}$, $a_i \geqslant 0$, $\sum_i a_i = 1$,

alors $Tr(h(W)) \leqslant \sum_i h(a_i)$

Preuve. Nous supposons, sans nuire à la généralité, que $I = \mathbb{N}$. Soit H' un espace d'Hilbert séparable contenant H et de dimension infinie. Soit encore W' l'extension de W sur H, telle que W'= 0 sur le complément de H. Il est évident que $W' = \sum_i a_i P_{[f_i]}$ sur H', et que

$$Tr(h(W')) = Tr(h(W)) \tag{5}$$

Si $\{g_i\}_{i \in \mathbb{N}}$ est une base de vecteurs propres de W', alors le théorème 1 montre qu'il existe un opérateur partiellement isométrique U tel que

$$\forall i \in \mathbb{N} : \sqrt{a_i} \; f_i = \sqrt{W'} \; U^+ g_i \tag{6}$$

Un calcul simple montre que $R(U^+U) \supseteq R(W')$, et donc $Tr(h(W')) = Tr\,(h(UW'U^+))$. Les relations (3) et (6) impliquent maintenant :

$$Tr\,(h(W')) = Tr\,(h(U\,W'U^+)) \leqslant \sum_i h(\,(g_i, UW'U^+g_i)\,) = \sum_i h(a_i)$$

ce qui, combiné à (5), démontre le théorème.

3.2. *Représentation de certains ensembles convexes d'opérateurs densité.*

Des ensembles convexes d'opérateurs densité interviennent lorsqu'on applique des méthodes de la théorie d'information en Mécanique Quantique. Ils ont en général la forme

$$S = \{W : Tr\,(WC_i) = k_i \;,\; i \in I\} \tag{7}$$

où C_i sont des opérateurs hermitiques continus représentant des observables et k_i des nombres réels (Cf. section V.3.2). Dans le cas spécial d'un S donné par la relation (7) et où une seule observable C intervient, on peut caractériser les éléments de S de façon simple :

Théorème 3. *Soit C un opérateur hermitique et continu et W un opérateur densité. Posons Tr(WC) = k. Alors W* *peut être écrit comme un mélange* $W = \sum_i a_i P_{[f_i]}$, $a_i > 0$, *tel que* $(f_i, Cf_i) = k$ *pour tout* i.

<u>Démonstration</u>. Soit $\mathcal{A}$ l'ensemble de toutes les familles finies ou dénombrables de paires $\{(a_1, f_1), (a_2, f_2), ..\}$ telles que les a_i soit des nombres positifs et les f_i de vecteurs normalisés tels que $(f_i, Cf_i) = k$, et $W - \sum_i a_i P_{[f_i]} \geqslant 0$. Nous ordonnons $\mathcal{A}$ par l'inclusion ensembliste. Nous avons l'intention de démontrer que les conditions d'applicabilité du lemme de Zorn sont satisfaites.

a) $\mathcal{A}$ est non vide : en fait, si $W = \sum_n \lambda_n P_{[g_n]}$ est la décomposition spectrale de W, alors

$$\mathrm{Tr}(WC) = \sum_n \lambda_n (g_n, Cg_n) = k$$

S'il existe un g_n tel que $(g_n, Cg_n) = k$, alors $\{(\lambda_n, g_n)\}$ est une famille à un élément qui appartient à $\mathcal{A}$ et donc $\mathcal{A}$ est non vide. Sinon, il existe des g_n et g_m tels que $(g_n, Cg_n) > k$, $(g_m, Cg_m) < k$. Posons

$$g(t) = tg_m + \sqrt{1-t^2}\; g_n \;, \; t \in [0,1]$$

L'application $t \longrightarrow (g(t), Cg(t))$ est continue et ses valeurs pour $t = 0$, $t = 1$ sont respectivement supérieure et inférieure à k. Il existe donc une valeur de t telle que $(g(t), Cg(t)) = k$. Puisque $g(t) \in R(W)$, le corollaire 3 montre que $\mathcal{A}$ n'est pas vide.

b) Tout sous-ensemble totalement ordonné de $\mathcal{A}$ est borné : en effet, soit $\mathcal{B} = \{M_j\}_{j \in J}$ un sous-ensemble totalement ordonné de $\mathcal{A}$, et posons $M = \bigcup_{j \in J} M_j$. Pour tout sous-ensemble fini K de M, par exemple $\{(a_1, f_1), (a_2, f_2), ...(a_n, f_n)\}$, il existe un M_j tel que $K \subset M_j$. Par conséquent, on a $W \geqslant \sum_{i=1}^{n} a_i P_{[f_i]}$ et cela montre que $\sum_{i=1}^{n} a_i \leqslant 1$. Il s'ensuit que pour tout $\varepsilon > 0$ il existe tout au plus un nombre fini de paires (a, f) telles que $a > \varepsilon$. On en déduit que M est dénombrable :

$$M = \{(a_1, f_1), (a_2, f_2), ...\}$$

et puisque

$$\forall n: \sum_{i=1}^{n} a_i \leqslant 1, \quad W - \sum_{i=1}^{n} a_i P_{[f_i]} \geqslant 0$$

on aura donc

$$\sum_i a_i \leqslant 1, \; W - \sum_i a_i P_{[f_i]} \geqslant 0$$

Il s'ensuit que $M \in \mathcal{A}$, i.e. $\mathcal{B}$ est borné.

Le lemme de Zorn peut donc être appliqué, et par conséquent $\mathcal{A}$ contient un élément maximal, par exemple $M_o = \{ (a_1, f_1), (a_2, f_2), \ldots \}$. On a évidemment $W \geqslant \sum_i a_i P_{[f_i]}$.
Supposons que > soit vrai. Alors l'opérateur

$$W' := \frac{W - \sum_i a_i P_{[f_i]}}{1 - \sum_i a_i}$$

serait un opérateur densité, et on aurait $Tr(W'C) = k$.
Or la partie (a) de cette démonstration implique l'existence d'une paire (a, f) telle que $a > 0$, $(f, Cf) = k$ et $W' \geqslant a P_{[f]}$, ou, en d'autres termes :

$$W \geqslant \sum_i a_i P_{[f_i]} + a \left(1 - \sum_i a_i\right) P_{[f]}$$

Par conséquent, la famille

$$\{ (a(1-\sum_i a_i), f), (a_1, f_1), (a_2, f_2), \ldots \}$$

appartiendrait à $\mathcal{A}$, contrairement à l'hypothèse faite que M_o est maximal. On en déduit que $W = \sum_i a_i P_{[f_i]}$. $\blacksquare$

Une autre façon de formuler le théorème 3 est de dire que tout élément de l'ensemble S défini par la relation (7), peut être écrit comme un mélange d'états purs appartenant à S. Ce fait a une interprétation géométrique simple : les états purs sont les éléments extrêmes de l'ensemble convexe S, les mélanges formant "l'intérieur" de S. Comme nous l'avons vu dans l'appendice, ch. V, le théorème trouve des applications au problème de l'inférence statistique en Théorie Quantique.

ANNEXE II

LA "PROBABILITE DE TRANSITION" ET LA DISTANCE DE V. CANTONI

1. INTRODUCTION

Parmi les distances entre états quantiques proposées jusqu'à ce jour, une des plus intéressantes est celle de V. Cantoni {74-77} . Son intérêt réside surtout dans le fait qu'elle est reliée à une "probabilité de transition généralisée" et applicable à toute théorie physique.

La définition de cette distance repose sur le système axiomatique de G. Mackey {42} , dont nous donnons brièvement les éléments principaux : à tout système physique, on associe un ensemble $\mathcal{O}$ d'observables et un ensemble $\mathcal{S}$ d'états. Pour toute observable $A \in \mathcal{O}$, pour tout état $\alpha \in \mathcal{S}$ et pour tout Borélien E, on note par $p(A, \alpha, E)$ la probabilité qu'une mesure de A, effectuée sur un système dans l'état α, donne un résultat appartenant à E. Mackey impose sur $\mathcal{O}$, $\mathcal{S}$ et $p(A, \alpha, E)$ trois axiomes qui sont trivialement valables dans toutes les théories physiques :

<u>Axiome 1</u>. *Pour tout* $A \in \mathcal{O}$ *et pour tout* $\alpha \in \mathcal{S}$, $\alpha_A(E) := p(A, \alpha, E)$ *est une mesure de probabilité sur les boréliens de R.*

<u>Axiome 2</u>. *Si* $\alpha_A(E) = \alpha'_A(E)$ *pour tout A et E, alors* $\alpha = \alpha'$. *De même, si* $\alpha_A(E) = \alpha_{A'}(E)$ *pour tout* α, *E, alors* $A = A'$.

<u>Axiome 3</u>. *Soit* $A \in \mathcal{O}$ *une observable et f une fonction réelle mesurable. Alors il existe une observable* $B \in \mathcal{O}$ *telle que*

$$\forall \alpha \in \mathcal{S}, \forall E: \quad \alpha_B(E) = \alpha_A(f^{-1}(E)).$$

On symbolise B par f(A).

Le système axiomatique de Mackey contient aussi d'autres axiomes, moins évidents, mais les trois axiomes ci-dessus suffisent pour l'étude de la "probabilité de transition généralisée" et de la distance de V. Cantoni. La définition de ces concepts peut maintenant être donnée ainsi : pour toute paire d'états α, β et toute observable A, définissons l'expression T_A (α,β) par

$$T_A^{1/2}(\alpha,\beta) := \int_R \sqrt{\frac{d\alpha_A}{d\sigma}\,\frac{d\beta_A}{d\sigma}}\; d\sigma \qquad (1)$$

σ étant une mesure quelconque par rapport à laquelle α et β sont absolument continus. Il est facile de voir que T_A ne dépend pas de σ. Enfin, définissons

$$T(\alpha,\beta) := \inf_{A\in\mathcal{O}} T_A(\alpha,\beta) \qquad (2)$$

La fonction $T(\alpha,\beta)$ est la "probabilité de transition généralisée" proposée par V. Cantoni. En ce qui concerne la distance entre les états α et β, elle est définie par

$$d_c(\alpha,\beta) = \sqrt{2\,(1 - T^{1/2}(\alpha,\beta))} \qquad (3)$$

Les définitions (2) et (3) ont deux défauts évidents : tout d'abord, elles n'ont aucune justification physique*. En effet, $d_c(\alpha,\beta)$ est appelée "distance" par Cantoni uniquement parce qu'elle est une métrique, i.e. satisfait aux conditions 1, 2 et 3 de la section V.2.1. Quant à $T(\alpha,\beta)$, elle est appelée "probabilité de transition généralisée", parce que dans le cas spécial de la Théorie Quantique et pour deux états α,β purs, elle est égale à la probabilité de transition habituelle $|(\alpha,\beta)|^2$. Enfin, un autre handicap de ces définitions est la complexité des relations (1) et (2), complexité qui n'a pas permis jusqu'à ce jour de calculer $T(\alpha,\beta)$ dans des cas plus généraux que le cas de deux états purs. S. Gudder dans son article intitulé "Some unsolved problems in quantum logics", a posé le problème de savoir si oui ou non, l'expression de $T(\alpha,\beta)$ pour deux états quantiques W_1,W_2 est donnée par la formule:

$$T(W_1,W_2)=1-\frac{\|(W_1-W_2)\|_1^2}{4}$$

---

* S. Gudder a essayé de spécifier une signification pour T dans {76} .

(Cf{ 99 }).

Dans cette annexe nous donnerons des expressions plus simples pour T et pour d_G , tant pour la théorie quantique que pour la mécanique statistique classique. En particulier, nous montrerons que la conjecture de S. Gudder n'est pas bonne . Nous montrerons aussi que T(α, β) est égale à la "probabilité de transition" quantique chaque fois que ce concept a un sens.

2. LE THEOREME PRINCIPAL

Suivant les définitions de Mackey (Cf {42} , p. 69), nous dirons qu'une observable A possède un spectre discret et fini, s'il existe des nombres réels $\lambda_1, \lambda_2, \ldots, \lambda_n$, tels que $\alpha_A(\{ \lambda_1, \lambda_2, \ldots, \lambda_n \}) = 1$ pour tout $\alpha \in \mathcal{S}$. Notre théorème principal affirme que, dans (2), il suffit de prendre l' infimum pour toutes les observables à spectre discret et fini. Nous montrerons d'abord deux lemmes simples :

Lemme 1. Soit (X, $\mathcal{B}$, μ) un espace de mesure et $L^2(X,\mu)$ l'espace d'Hilbert de toutes les fonctions réelles de carré sommable sur X. Alors pour tout élément positif $f \in L^2(X,\mu)$ et pour tout $\varepsilon > 0$, il existe une partition D_i, i=1,2,...n de X telle que, pour toute partition B_j,j=1,2,...m plus fine que D_i, on a :

$$\left\| f - \sum_{\mu(B_j)\neq 0} \sqrt{\int_{B_j} f^2 \, d\mu} \, \frac{\chi_{B_j}}{\sqrt{\mu(B_j)}} \right\| < \varepsilon$$

χ_{B_j} étant la fonction caractéristique de B_j.

Preuve. Comme les fonctions en escalier sont denses dans $L^2(X,\mu)$, pour tout $\varepsilon > 0$ et tout $f \in L^2(X,\mu)$ il existe une partition D_i, i = 1, 2, ... n, et des nombres a_i tels que

$$\| f - \sum_i a_i \chi_{D_i} \| \equiv \sqrt{\int_X | f - \sum_i a_i \chi_{D_i} |^2 \, d\mu} < \varepsilon \tag{4}$$

Nous rappelons que, d'après la définition habituelle des fonctions en escalier, ({100} p. 231), $\mu(D_i) = \infty$ implique $a_i = 0$. Si l'on a en plus $f \geqslant 0$, alors on peut prendre $a_i \geqslant 0$. Pour toute partition B_j plus fine que D_i, définissons les nombres b_j par $b_j = a_i$ ssi $B_j \subseteq D_i$. La relation (4) implique alors

$$||f - \sum_j b_j \chi_{B_j}|| < \varepsilon \tag{5}$$

En posant $f_1 = \sum_{\mu(B_j)\neq 0} \sqrt{\int_{B_j} f^2 \, d\mu} \, \frac{\chi_{B_j}}{\sqrt{\mu(B_j)}}$, on trouve

$$||f - f_1|| \leqslant ||f - \sum_j b_j \chi_{B_j}|| + || \sum_j b_j \chi_{B_j} - f_1|| \tag{6}$$

et aussi

$$|| \sum_j b_j \chi_{B_j} - f_1||^2 = \sum_{\mu(B_j)\neq 0,\infty} \left| b_j - \frac{\sqrt{\int_{B_j} f^2 \, d\mu}}{\sqrt{\mu(B_j)}} \right|^2 \mu(B_j) \leqslant$$

$$\leqslant \sum_j \Big| || f . \chi_{B_j} || - || b_j \chi_{B_j} || \Big|^2 \leqslant$$

$$\leqslant \sum_j ||f . \chi_{B_j} - b_j \chi_{B_j}||^2 = || f - \sum_j b_j \chi_{B_j} ||^2 \tag{7}$$

En combinant (5),(6) et (7) on en déduit que $||f - f_1|| \leqslant 2\varepsilon$. ∎

<u>Lemme 2</u>. Si $f, g \in L^2(X, \mu)$ et $f \geqslant 0$, $g \geqslant 0$, alors $\forall \varepsilon > 0$ il existe une partition B_j, $j = 1, 2, \ldots m$, telle que:

$$\left| \int_X fg \, d\mu - \sum_j \sqrt{(\int_{B_j} f^2 \, d\mu) \, (\int_{B_j} g^2 \, d\mu)} \right| < \varepsilon$$

Preuve. Selon le lemme 1, pour tout $\varepsilon > 0$ on peut trouver des partitions B_j, B'_j, telles que $\|f-f_1\| \leqslant \varepsilon$, $\|g-g_1\| \leqslant \varepsilon$, où f_1 a été définie dans la preuve précédente et g_1 est définie de façon analogue. De plus, nous pouvons supposer que les deux partitions sont identiques (sinon,on peut utiliser la partition $B_j \cap B'_{j'}$). On a maintenant :

$$\left| \int_X fg\, d\mu - \int_X f_1 g_1\, d\mu \right| \leqslant \int |f-f_1|\, g\, d\mu + \int f_1 |g-g_1|\, d\mu \leqslant$$

$$\leqslant \varepsilon\, (\|f\| + \|g\|) \qquad (8)$$

D'autre part,

$$\int_X f_1\, g_1\, d\mu = \sum_{\mu(B_j) \neq \infty} \sqrt{\left(\int_{B_j} f^2\, d\mu\right)\left(\int_{B_j} g^2\, d\mu\right)} \qquad (9)$$

tandis que le lemme 1 montre de nouveau que pour cette partition B_j on a

$$\sum_{\mu(B_j)=\infty} \int_{B_j} f^2\, d\mu < \varepsilon^2, \quad \sum_{\mu(B_j)=\infty} \int_{B_j} g^2\, d\mu < \varepsilon^2$$

ce qui implique

$$\sum_{\mu(B_j)=\infty} \sqrt{\left(\int f^2\, d\mu\right)\left(\int g^2\, d\mu\right)} < \varepsilon^2 \qquad (10)$$

La démonstration du lemme 2 est achevée en combinant (8),(9) et (10). ∎

Venons en maintenant à l'établissement du théorème. Nous allons noter par $\mathcal{O}_f$ l'ensemble des observables à spectre discret et fini.

Théorème 1. *Dans toute théorie physique satisfaisant aux axiomes 1, 2 et 3 de l'axiomatique de Mackey, $T(\alpha, \beta)$ est donnée aussi par la formule.*

$$T(\alpha, \beta) = \inf_{A \in \mathcal{O}_f} T_A(\alpha, \beta) \qquad (11)$$

Preuve. Pour toute paire d'états α,β , et pour toute observable A, soit σ une mesure par rapport à laquelle α_A et β_A sont absolument continues (par exemple, $\sigma=\alpha_A+\beta_A$). Puisque α_A et β_A sont des mesures de probabilité sur R, $\sqrt{\frac{d\alpha_A}{d\sigma}}$ et $\sqrt{\frac{d\beta_A}{d\sigma}}$ appartiennent à $L^2(R,\sigma)$. Le lemme 2 et la relation (1) impliquent donc que pour tout $\varepsilon > 0$ il existe une partition B_k, k = 1, 2, ... n, de R telle que

$$\left| T_A^{1/2} - \sum_k \sqrt{\alpha_A(B_k)\ \beta_A(B_k)} \right| < \varepsilon \qquad (12)$$

Définissons maintenant la fonction $h(x)= \sum_{k=1}^{n} k\ \chi_{B_k}(x)$

et l'observable A' = h(A) (voir l'axiome 3). Il est facile de voir que:

(a) le spectre de A' est discret et fini, constitué des points k = 1, 2, ... n,

(b) on a : $\alpha_{A'}(\{k\}) =\alpha_A(B_k)$ et $T_{A'}^{1/2}(\alpha ,\beta)= \sum_k\sqrt{\alpha_A(B_k)\beta_A(B_k)}$.

En remplaçant cette dernière expression dans (12) on trouve

$$|T_A^{1/2} - T_{A'}^{1/2}| < \varepsilon \qquad (13)$$

Les relations (2) et (13) impliquent trivialement la relation (11). $\blacksquare$

3. APPLICATION A LA MECANIQUE STATISTIQUE CLASSIQUE

En Mécanique Statistique Classique, les systèmes physiques sont décrits par leur espace des phases Ω . Sur les Boréliens de Ω on définit la mesure invariante de Liouville μ.

Les états sont des densités de probabilité sur Ω , i.e. des fonctions non-négatives ρ telles que $\int_\Omega \rho d\mu = 1$ (voir, par exemple {42} p. 48), tandis que les observables sont des fonctions réelles mesurables f sur Ω (i.e. des variables aléatoires). La mesure de probabilité $p(f,\rho,E) \equiv \rho_f(E)$ est donnée par

$$\rho_f(E) = \int_{f^{-1}(E)} \rho \, d\mu \tag{14}$$

La valeur de $T(\rho_1,\rho_2)$ peut être facilement déduite du Théorème 1 :

<u>Proposition 1</u> Dans le cas de la Mécanique Statistique Classique, la valeur de $T(\rho_1,\rho_2)$ est donnée par la formule :

$$T(\rho_1,\rho_2) = \left(\int_\Omega \sqrt{\rho_1 \rho_2} \, d\mu \right)^2$$

<u>Preuve</u> Soit f une variable aléatoire représentant une observable à spectre discret et fini. Cela signifie simplement que f prend seulement un nombre fini de valeurs y_1, y_2, ..., y_n, et que par conséquent $f^{-1}(y_i)$, i = 1, 2, ... n, est une partition de Ω . Les relations (1) et (14) donnent maintenant

$$T_f^{1/2}(\rho_1, \rho_2) = \sum_i \sqrt{\int_{f^{-1}(y_i)} \rho_1 \, d\mu \cdot \int_{f^{-1}(y_i)} \rho_2 \, d\mu} \tag{15}$$

Puisque f peut être choisie arbitrairement, il en est de même pour la partition $f^{-1}(y_i)$. Le théorème 1 et la relation (15) impliquent donc

$$T^{1/2} = \inf_{\{B_j\}_{j\in J}} \sum_j \sqrt{\int_{B_j} \rho_1 \, d\mu \cdot \int_{B_j} \rho_2 \, d\mu} \tag{16}$$

l'infimum étant pris sur toutes les partitions. D'autre part, l'inégalité de Schwartz donne

$$\int_\Omega \sqrt{\rho_1 \rho_2} \, d\mu \leq \sum_j \sqrt{\int_{B_j} \rho_1 \, d\mu \cdot \int_{B_j} \rho_2 \, d\mu} \tag{17}$$

La proposition découle de (16) et (17) et du lemme 2. ∎

4. LE CAS QUANTIQUE

En Théorie Quantique les états sont des opérateurs densité et les observables des opérateurs auto-adjoints. La mesure p(A,W,E) est donnée par

$$p(A,W,E) = Tr(F_E^A\ W) \qquad (18)$$

F_E^A étant la mesure spectrale qui correspond à A.

Nous donnons maintenant dans le cas quantique une expression de $T(\alpha\ ,\beta)$, beaucoup plus simple que la définition générale (1) + (2).

Proposition 2. Dans le cas Quantique, la "probabilité de transition généralisée de Cantoni, entre deux états représentés par deux opérateurs densité W_1 et W_2, est donnée par

$$T^{1/2}(W_1, W_2) = \inf_{\{f_i\}_{i \in n}} \sum_{i \in N} \sqrt{(f_i, W_1 f_i)\ (f_i, W_2 f_i)} \qquad (19)$$

l'infimum étant pris sur toutes les bases de l'espace d'Hilbert H.

Preuve. Le théorème 1 montre que $\forall\ \varepsilon > 0$ il existe un opérateur auto-adjoint A à spectre discret et fini $\lambda_1, \lambda_2, \ldots, \lambda_n$, tel que $0 \leqslant T_A^{1/2} - T^{1/2} \leqslant \varepsilon$, ou, de manière équivalente :

$$0 \leqslant \sum_k \sqrt{Tr(W_1\ F^A_{\{\lambda_k\}}) . Tr(W_2 F^A_{\{\lambda_k\}})} - T^{1/2} \leqslant \varepsilon \qquad (20)$$

$F^A_{\{\lambda_k\}}$ étant les projecteurs sur les sous-espaces propres de A.

Soit $\{x_{k_i}\}_{k_i \in I_k}$ une base du sous-espace sur lequel projette $F^A_{\{\lambda_k\}}$. Alors $\bigcup_k \{x_{k_i}\}_{k_i \in I_k}$ est une base de H et l'on a :

$$\sum_k \sqrt{Tr(W_1 F^A_{\{\lambda_k\}})\ Tr(W_2\ F^A_{\{\lambda_k\}})} = \sum_k \sqrt{\Sigma_{I_k} (x_{k_i}, W_1\ x_{k_i})\ \Sigma_{I_k} (x_{k_i}, W_2\ x_{k_i})} \geqslant$$

$$\geqslant \sum_k \Sigma_{I_k} \sqrt{(x_{k_i}, W_1\ x_{k_i})(x_{k_i}, W_2\ x_{k_i})} \qquad (21)$$

Soit maintenant B un opérateur autoadjoint à spectre discret et non dégénéré, et ayant les x_{k_i} comme vecteurs propres.

Alors (1) et (2) impliquent

$$\sum_k \sum_{I_k} \sqrt{(x_{k_i}, W_1 x_{k_i})(x_{k_i}, W_2 x_{k_i})} = T_B^{1/2}(W_1, W_2) \geqslant T^{1/2}(W_1, W_2) \quad (22)$$

Le résultat désiré peut être maintenant déduit des relations (20), (21) et (22). ∎

<u>Corollaire 1</u>. La distance $d_c(W_1, W_2)$ est donnée par

$$d_c(W_1, W_2) = \sup_{\{f_i\}_{i \in \mathbb{N}}} \sqrt{\sum_{i \in \mathbb{N}} \left[\sqrt{(f_i, W_1 f_i)} - \sqrt{(f_i, W_2 f_i)} \right]^2}$$

le supremum étant pris sur toutes les bases de l'espace d'Hilbert H.

<u>Preuve</u>. Conséquence évidente des (3) et (19). ∎

Le résultat de la proposition précédente est considérablement simplifié lorsque l'un de deux états est pur.

<u>Proposition 3</u>. La "probabilité de transition généralisée" de Cantoni entre un état pur $P_{[g]}$ et un état W est égale à (g, Wg).

<u>Preuve</u>. Pour toute paire d'états W_1, W_2 et toute base $\{f_n\}_{n \in \mathbb{N}}$, on trouve :

$$\left(\sum_n \sqrt{(f_n, W_1 f_n)(f_n, W_2 f_n)} \right)^2 = \sum_{n,m} \sqrt{(f_n, W_1 f_n)(f_m, W_1 f_m)(f_n, W_2 f_n)(f_m, W_2 f_m)}$$

$$\geqslant \sum_{n,m} |(f_n, W_1 f_m)| \, |(f_n, W_2 f_m)| \geqslant \left| \sum_{n,m} (W_1 f_n, f_m)(f_m, W_2 f_n) \right| = \mathrm{Tr}(W_1 W_2)$$

En combinant ce résultat avec la proposition 2 on en déduit que

$$\forall W_1, W_2 : T(W_1, W_2) \geqslant \mathrm{Tr}(W_1 W_2) \quad (23)$$

En posant dans (23), $W_1 = W$ et $W_2 = P_{[g]}$ on trouve :

$$T(W,P_{[g]}) \geqslant Tr(W\ P_{[g]}) = (g,\ Wg) \tag{24}$$

D'autre part, on peut facilement vérifier que

$$(g,Wg) = T_{P_{[g]}}(W,P_{[g]}) \geqslant T\ (W,\ P_{[g]})$$

Ce qui achève la démonstration. ■

5. COMMENTAIRES

5.1. *Commentaires sur la "probabilité de transition généralisée."*

Comparons le résultat de la proposition 3 à celui donné par la Mécanique Quantique. La "probabilité de transition" possède une signification précise en Théorie Quantique dans le cas suivant : une mesure d'une observable à spectre discret et non dégénéré est effectuée sur un système se trouvant à l'état W. Si les valeurs propres et vecteurs propres de l'observable sont respectivement λ_i et g_i, $i=1,2,...,$ alors la probabilité de trouver la valeur λ_i est $Tr(WP_{[g]}) = (g_i,\ Wg_i)$. Si en plus la mesure est de première espèce (voir section V.3.2), alors l'état après la mesure sera g_i. On en déduit que la probabilité de transition de l'état W à l'état g_i, calculée selon le formalisme Quantique, est égale à $(g_i,\ Wg_i)$. On voit donc que la fonction $T(\alpha,\beta)$ définie par Cantoni donne la probabilité de transition quantique, chaque fois que ce concept a un sens.

Peut-on affirmer pour autant que la "probabilité de transition généralisée" mérite son nom? Il est vrai en effet qu'elle se trouve partiellement justifiée à postériori puisque $T(\alpha,\beta)$ donne le bon résultat quand l'un des deux états est pur. Toutefois, il existe dans la littérature d'autres propositions pour définir une probabilité de transition généralisée, qui donnent le même résultat lorsque l'un des états est pur, et qui diffèrent parfois de T dans le cas général. A titre d'exemple, nous citons P.C.Zabey {101} , qui a proposé l'expression $Tr(W_1W_2)$ et A. Uhlmann {102} , qui a proposé

$(\mathrm{Tr}\,|\sqrt{W_1}\ \ \sqrt{W_2}\,|\,)^2$. On voit facilement que si l'un des deux états est pur, ces expressions sont égales à $T(W_1, W_2)$. Mais si les deux états sont des mélanges alors,

$$\mathrm{Tr}(W_1W_2) = \mathrm{Tr}(\sqrt{W_2}\ \ W_1\ \ \sqrt{W_2}) = \mathrm{Tr}\,|\sqrt{W_1}\quad \sqrt{W_2}\,|^2 < (\mathrm{Tr}|\sqrt{W_1}\ \ \sqrt{W_2}|)^2$$

et par conséquent, $T(W_1,W_2)$ est différente de l'une au moins de ces deux expressions. Le fait que $T(W_1, W_2)$ donne le bon résultat dans un cas particulier ne signifie donc pas que son appellation soit justifiée.

Enfin, on voit aisément, d'après la proposition 3, que la réponce à la question posée par S.Gudder (section 1) est négative. En effet, on a en général:

$$T\,(W,P_{[g]}) = (g,\ Wg) \neq 1 - \frac{||\,P_g - W_1\,||_1^{\ 2}}{4}$$

5.2. Commentaires sur la distance de Cantoni.

D'après le corollaire 1, on voit que la distance de Cantoni est obtenue en comparant les W_1 et W_2 sur une base et en prenant le supremum pour toutes les bases. On peut toutefois se demander pourquoi cette comparaison devrait se faire à travers les racines. En fait, on peut montrer que la distance de Cantoni et la nôtre appartiennent à toute une classe de distances indexées par un réel positif q, la distance de Cantoni correspondant au cas q = 2 et la notre à q = 1 :

<u>Théorème</u> 2: *Pour tout $q \geqslant 1$, l'expression*

$$d_q(W_1,W_2) := \sup_{\{f_n\}_{n\in N}} \left(\sum_{n\in N} \left|(f_n,W_1f_n)^{\frac{1}{q}} - (f_n,W_2f_n)^{\frac{1}{q}}\right|^{\,q} \right)^{\frac{1}{q}}$$

(le supremum étant pris sur toutes les bases de H), définit une distance (i.e. métrique) sur l'ensemble des opérateurs densité. En plus, on a les égalités :

$d_c\,(W_1,W_2) = d_2\,(W_1,W_2)$ *et* $d(W_1,W_2) = \frac{1}{2}\,d_1(W_1,W_2)$.

<u>Preuve</u>. Puisque $\mathrm{Tr}(W_1) = \mathrm{Tr}(W_2) = 1$, les suites

$$\{(f_n,W_1f_n)^{\frac{1}{q}}\}_{n\in N}\ ,\quad \{(f_n,W_2f_n)^{\frac{1}{q}}\}_{n\in N}$$

appartiennent à l'espace de Banach $L^q(N)$ (espace des suites $\{a_n\}_{n\in N}$ telles que $\sum_n |a_n|^q < \infty$). Par conséquent, pour toute base $B = \{f_n\}_{n\in N}$, l'expression

$$d_q(W_1,W_2\ ;\ B) = \Big(\sum_{n\in N}\Big|(f_n,\ W_1f_n)^{\frac{1}{q}} - (f_n,\ W_2f_n)^{\frac{1}{q}}\Big|^q\Big)^{\frac{1}{q}}$$

est finie et obéit à l'inégalité triangulaire. On a donc :

$$d_q(W_1,W_3\ ;\ B) \leqslant d_q\ (W_1,\ W_2\ ;\ B\ ) + d_q\ (W_2,\ W_3\ ;\ B)$$

ce qui implique trivialement

$$d_q(W_1,W_3) \leqslant d_q(W_1,W_2) + d_q\ (W_2,\ W_3)$$

i.e. d_q est une distance. L'égalité $d_2 = d_c$ découle du corollaire 1. D'autre part, pour $q = 1$ on a :

$$d_1(W_1,W_2) = \sup_B\ \sum_n |(f_n,(W_1-W_2)f_n)| \leqslant$$

$$\leqslant \sup_B\ \sum_n\ (f_n,\ |W_1 - W_2|\ f_n) = \mathrm{Tr}|W_1-W_2|=2d(W_1,W_2)$$

En prenant une base de vecteurs propres de $W_1 - W_2$ on voit alors que $d_1(W_1,W_2) = 2d(W_1,W_2)$. $\blacksquare$

Vu le théorème 2 , nous croyons que le choix le plus naturel parmi toutes les distances correspond à $q = 1$, choix qui est soutenu aussi par les considérations de la section V.2. et qui conduit à notre distance $d(W_1,\ W_2)$.

CONCLUSION

Les problèmes fondamentaux posés par le formalisme quantique restent toujours ouverts; nous espérons néanmoins que les quelques contributions qui sont contenues dans ce travail auront aidé à clarifier certains aspects de ces problèmes. Nous croyons aussi que certaines de ces contributions - et en particulier le concept de distance entre états quantiques, ainsi que l'établissement des liens entre le principe d'entropie maximale et le principe d'équiprobabilités de Laplace - peuvent s'avérer fructueuses dans un proche avenir.

BIBLIOGRAPHIE

1 -N.HADJISAVVAS,F.THIEFFINE et M.MUGUR-SCHACHTER, Found.Phys. 10,751 (1980)

2 -N.HADJISAVVAS,F.THIEFFINE et M.MUGUR-SCHACHTER,*Critical remark on Jauch's program*,Let.Nuovo Cimento 30, 530 (1981)

3 -F.THIEFFINE,N.HADJISAVVAS et M.MUGUR-SCHACHTER, *Supplement to a critique of Piron's system of questions and propositions*, Found. Phys. 11,645 (1981)

4 -N.HADJISAVVAS,*The Maximum Entropy Principle as a consequence of the Principle of Laplace*,J. Stat. Phys.(à paraître)

5 -N.HADJISAVVAS,*Distance between states and statistical inference in Quantum theory*, Ann. Inst.H.Poinc.,T. 34,n° 4.

6 -N.HADJISAVVAS, *Properties of mixtures of non-orthogonal states*, Lett.in Math.Phys. 5,327 (1981)

7 -N.HADJISAVVAS, *On Cantoni's generalised transition probability*, Comm.Math.Phys. (à paraître en 1981)

8 -J.von NEUMANN, *Mathematical foundations of quantum mechanics*, Princeton University Press (Princeton,1955)

9 -G.BIRKOFF et J.von NEUMANN,Ann.Math. 37,823 (1936)

10-H.REICHENBACH,Ann.Inst.Henri Poincaré 13,109 (1952-1953)

11-J.LUKASIEWICZ, dans *Polish Logic*, Oxford University Press (London,1967)

12-J.M.JAUCH,Helv.Phys.Acta 35,711 (1960)

13-J.M.JAUCH,Helv.Phys.Acta 37,293 (1964)

14-J.M.JAUCH et B.MISRA,Helv.Phys.Acta 34,699 (1961)

15-C.PIRON,Helv.Phys.Acta 34,504 (1961)

16-G.EMCH et C.PIRON,Helv.Phys.Acta 35,542 (1962)

17-L.H.LOOMIS,Mem.Am.Math.Soc. 18,36 (1955)

18-C.PIRON, *Axiomatique Quantique* ,Thèse (Université de Genève,1964)

19-J.M.JAUCH et C.PIRON,Helv. Phys. Acta 42,842 (1969)

20-C.PIRON,Found.Phys. 2,287 (1972)

21-C.PIRON,Conférences à Lyon, 4-8 Juin 1973

22-C.PIRON,*Foundations of Quantum Physics*,W.A.Benjamin Inc., (Reading,Mass.1976)

23-C.PIRON,Ann.Fond.Louis de Broglie 3,131 (1978)

24-K.POPPER,Nature 219,682 (1968)

25-M.JAMMER, *The Philosophy of Quantum Mechanics,* Wiley(New York 1974)

26-J.PARK et H.MARGENAU,Int.J.Theor.Phys. 1,211 (1968)

27-J.M.JAUCH,*Foundations of Quantum Mechanics*,Addison-Wesley (Reading,Mass.1968)

28-J.M.JAUCH,dans: *Proceedings of the International School of Physics "Enrico Fermi"*,Course 49,Academic Press (New York,1971)

29-J.CYRANSKI,Information and Control 41,275 (1979)

30-G.BIRKHOFF, *Lattice Theory,* Amer.Math.Soc.Colloq.Publ.Vol.25 revised edition (Providence,R.I. 1948)

31-K.R.PARTHASARATHY,dans: *Les probabilités sur les structures algébriques,* Colloques Internationaux du CNRS,n°186 (1970)

32-A.M.GLEASON,J.Rat.Mech.and Anal.6,385 (1965)

33-K.GÖDEL,Monatshefte für Math.und Phys.3_,173 (1931)

34-J.F.PABION, *Logique Mathématique,* Hermann (Paris,1976)

35-J.SHONFIELD , *Mathematical Logic,* Addison-Wesley (Reading, Mass. 1967)

36-A.SHIMONY,Found.Phys. 1,325 (1971)

37-R.J.GREECHIE etS.P.GUDDER,dans: *Contemporary Research in the Foundations and Philosophy of Quantum Mechanics,* C.A.Hooker ed. D.Reidel (Dordrecht,Holland 1974)

38-V.S.VARADARAJAN, *Geometry of Quantum Theory,* D.Van Nostrand (Princeton,1968)

39-S.GUDDER,dans: *The Uncertainty Principle and Foundations of Quantum Mechanics,* W.C.Price and S.S.Chissich eds.,Wiley (New York, 1977)

40-W.GUZ,Ann.Inst.henri Poincaré,28,1 (1978)

41-W.GUZ,Ann.Inst.Henri Poincaré 33,63 (1980)

42-G.MACKEY, *The Mathematical Foundations of Quantum Mechanics*, Benjamin (New York,1963)

43-K.BUGAJSKA et S.BUGAJSKI,Ann.Inst.Henri Poincaré 19,332 (1973)

44-J.CYRANSKI dans:*Théorie de l'Information*,Colloques Internationaux du CNRS n° 276 (Paris,1978)

45-R.CRISTESCU, *Ordererd vector spaces and linear operators*, Abacus Press (England,1976)

46-R.J.GREECHIE,J.of Comb.Theory,10,119 (1971)

47-R.J.GREECHIE,Car.J.Scien.Math. 1,15 (1969)

48-A.EINSTEIN,B.PODOLSKI et N.ROSEN,Phys.Rev. 47,777(1935)

49-W.OCHS, Z.Naturforsch. 27A, 893 (1972)

50-B.MIELNIK,dans: *Quantum Mechanics,Determinism,Causality and Particles*, M.Flato et al,eds.,D.Reidel (Dordrecht, Holland 1976)

51-J.M.JAUCH et C.PIRON,dans:*Quanta* ,P.Feund *et al* eds.,University of Chicago Press (Chicago,1970)

52-T.FINE, *Theories of probability: an examination of foundations*, Academic Press (New York,1973)

53-N.HADJISAVVAS et M.MUGUR-SCHÄCHTER,Lett.Nuov.Cim.23,439 (1978)

54-L.SCHMETTERER, *Inroduction to Mathematical Statistics*,Springer Verlag (Berlin 1974)

55-M.G.KENDALL et A.STUART,*The advanced theoru of statistics*,Griffin (London,1969)

56-J.CYRANSKI,Found.Phys. 8,493 (1978)

57-K.FRIEDMAN et A.SHIMONY,J.Stat.Phys. 3,381 (1971)

58-E.T.JAYNES,Phys.Rev. 106,620 (1957); 108,171(1957)

59-E.T.JAYNES,I.E.E.E.Trans.Sys.Scie.Cyb.,SSC-4,228 (1968)

60-E.T.JAYNES,dans: *Colloquiumletters in pure and applied science,* n°4 (Février 1978)

61-E.T.JAYNES,dans: *Statistical Physics,* vol.3,W.K.Ford,ed., Benjamin Inc. (New York,1963)

62-S.KULLBACK, *Information Theory and Statistics,* Wiley (New York 1959)

63-S.KULLBACK et M.A.KHAIRAT,Ann.Math.Stat.37,279 (1966)

64-M.MUGUR-SCHÄCHTER,Ann.Inst.Henri Poincaré 32,33 (1980)

65-M.MUGUR-SCHÄCHTER et N.HADJISAVVAS,Kybernetes (à paraître)

66-M.MUGUR-SCHÄCHTER, *Les arbres de probabilité en théorie quantique,* Séminaire à la Fondation Louis de Broglie,16 avril 1980 (non publié)

67-G.MAROULIS, *Molecular properties and basis set quality: an approach based on information theory* , Thèse, Université de Louvain-la-Neuve,1981

68-N.HADJISAVVAS, Epist. Lett. 24, 14 (1979)

69-F.SELLERI, Communication au Collège de France (Table ronde à l'occasion de l'année Einstein, Juin 1979)

70-A.GARUCCIO et F.SELLERI, Epist.Lett.24, 1 (1979)

71-I.SCISZAR,Stud. Sci. Math. Hung. 2, 299 (1967); 2,329 (1967)

72-S.GUDDER, Comm.Math.Phys. 29, 249 (1973)

73-S.GUDDER,J.P.MARCHAND et W.WYSS,J.Math.Phys. 20, 1963 (1979)

74-V.CANTONI,Comm. Math.Phys. 44, 125 (1975)

75-V.CANTONI,Comm.Math.Phys. 56,189 (1977)

76-S.GUDDER,Comm.Math.Phys. 63;265 (1978)

77-W.K.WOOTERS,Phys.Rev.D. 23,357 (1981)

78-W.WOOTERS, *The acquisition of information from quantum measurements,* PHD thesis, University of Texas, 1980

79-J.M.JAUCH,B.MISRA et A.G.GIBSON, Helv.Phys.Acta 41,513 (1968)

80-B.MISRA,dans: *Physical reality and Mathematical description,* C.P.Enz et J.Mehra, eds.,D.Reidel (Dordrecht,Holland,1974)

81-N.S.KRONFLI,Int.J.Theor.Phys. 3,191 (1970)

82-N.HADJISAVVAS,Ann.Fond.Louis de Broglie, 3,155 (1978)

83-G.C.WICK,A.S.WIGHTMAN et E.WIGNER,Phys.Rev. 88,101 (1952)

84-N.BOGOLUBOV,A.LOGUNOV et I.TODOROV, *Introduction to axiomatic Quantum Field theory*, Benjamin Inc.(New York,1975)

85-R.SCHATTEN, *Theory of cross-norm spaces*, (Princeton,1950)

86-O.BRATELLI et D.ROBINSON, *Operator algebras and quantum statistical mechanics*, vol.1 (1980)

87-W.OCHS,Rep.Math.Phys.8,109 (1975)

88-P.STASZEWSKI,Rep.Math.Phys.13,67 (1978)

89-I.SINGER,*Best approximation in normed linear spaces and by elements of linear subspaces*,Springer Verlag (1970)

90-I.SINGER,*The theory of best approximation and functional analysis* S.I.A.M. (Philadelphia,1974)

91-F.R.DEUTSCH et P.H.MAZERICK,S.I.A.M.Rev. 9,516 (1967)

92-K.BUGAJSKA et S.BUGAJSKI,Bull.Acad.Pol.Sci. 21,873 (1976)

93-E.G.BELTRAMENTI et G.CASSINELLI,Riv.Nuov.Cim. 6,321 (1976)

94-A.MESSIAH, *Mécanique Quantique*, Dunod (Paris,1965)

95-B.D'ESPAGNAT,*Conceptual Foundations of Quantum Theory*,Benjamin Inc. (New York,1971)

96-A.WEHRL,Rev.Mod.Phys.50,221 (1978)

97-A.UHLMANN,Rep. Math.Phys.1,147 (1970)

98-D.LANFORD et D.ROBINSON,J.Math.Phys. 9,1120 (1968)

99-S.GUDDER,dans: *Mathematical foundations of quantum theory*,A.R. Marlow,ed., Academic Press (New York,1978)

100-S.LANG,*Real Analysis*,Addison-Wesley (Reading,Mass.1973)

101-P.ZABEY,*Probabilités de transition en axiomatique quantique*,Thèse, Université de Genève,(1972)

102-A.UHLMANN,Rep.Math.Phys. 9,973 (1976)