SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

PAUL-ANDRÉ MEYER

Éléments de probabilités quantiques (exposés I à V)

Séminaire de probabilités (Strasbourg), tome 20 (1986), p. 186-312

<http://www.numdam.org/item?id=SPS_1986__20__186_0>

ELEMENTS DE PROBABILITES QUANTIQUES

par P.A. Meyer

Les exposés qui suivent ont été faits à Strasbourg pendant l'année universitaire 1984/85, devant des auditeurs en nombre variable, parmi lesquels D.Bakry, M.Emery, O.Gebuhrer, J-L.Journé, M.Ledoux, R.Léandre sont parvenus à tenir aussi longtemps que le conférencier. Le texte que l'on trouvera ci-après a été considérablement amélioré grâce à leurs remarques. Nous espérons que le travail sera poursuivi en 85/86, et que les exposés pourront être réunis en un volume séparé.

L'idée de tenir un séminaire sur ce sujet est née à Bangalore en 1982, en écoutant les exposés de R.L. Hudson et K.R. Parthasarathy sur le calcul stochastique non commutatif, renouvelée à Bénarès en 1984 au cours de discussions avec L. Accardi. Tous nos remerciements vont à ceux-ci, ainsi qu'à J. Lewis et R.F. Streater, pour leur aide et leurs encouragements répétés. De nombreuses occasions de discussions avec des physiciens théoriciens ont été fournies par des rencontres régulières à Strasbourg (RCP 25 du CNRS) et à Bielefeld (BiBoS), ou irrégulières (Heidelberg, Ascona...). Nous remercions aussi le programme d'échange Franco-Britannique pour une fructueuse mission de 15 jours en Angleterre.

TABLE DES MATIERES ET MODE D'EMPLOI

Pour une lecture suivie, on fera bien de se borner à lire les sections indispensables pour notre but (qui est le calcul stochastique non commutatif), marquées d'un i dans la table des matières. On peut aussi parcourir les sections marquées d'un c, qui ont une valeur "culturelle" (elles contiennent le plus souvent les motivations physiques, dans la mesure où le rédacteur lui même les a comprises). Les sections marquées d'un m sont en marge du texte, et on fera bien de les omettre.

I. Les notions fondamentales
- I. Axiomatique de von Neumann (tribus et lois) — i
- II. Axiomatique de von Neumann (variables aléatoires) — i
- III. Noyaux — m
- IV. Appendice (langage des C^*-algèbres, convergence en loi) — c

II. Quelques exemples discrets
- I. Le spin — m sauf n^{os} 1-3 i
- II. La tribu engendrée par deux événements — c
- III. Systèmes finis de spins — (appendice m) i

III. Couples canoniques
I. Le théorème de Stone-von Neumann c
II. Calculs sur les couples canoniques c et i
III. Fonctions de Wigner m

IV. Probabilités sur l'espace de Fock
I. L'espace de Fock (définition algébrique) i
II. Interprétations probabilistes i
III. Compléments m

V. Calcul stochastique non commutatif
I. Intégrales stochastiques de processus d'opérateurs i
II. Développement du calcul stochastique, et é.d.s. i
III. Opérateurs définis par des noyaux c

ELEMENTS DE PROBABILITES QUANTIQUES. I

Les notions fondamentales

par P.A. Meyer

La théorie moderne des probabilités est toujours écrite par les mathématiciens dans le système axiomatique de Kolmogorov. La plupart des physiciens pensent que ce système est insuffisant pour décrire certains aspects du monde réel, et utilisent une autre axiomatique, due pour l'essentiel à von Neumann (et d'ailleurs contemporaine de celle de Kolmogorov , sinon légérement antérieure). Bien que la mécanique quantique soit une théorie essentiellement probabiliste, les probabilités utilisées par les physiciens sont longtemps restées assez rudimentaires, à l'inverse de leur analyse fonctionnelle, et les probabilistes << classiques >> ont pu se permettre de les ignorer. Mais il n'en est plus du tout de même aujourd'hui. En particulier, on assiste depuis quelques années à une floraison de travaux sur le calcul stochastique non commutatif, qui présentent une version très attirante du calcul d'Ito pour des processus d'opérateurs.

Ce séminaire a pour but de préparer quelques auditeurs probabilistes, tous dépourvus de connaissances techniques en physique , à lire ces travaux récents. Nous avons essayé d'utiliser un langage et des notations aussi proches que possible de nos habitudes (par exemple "loi", " variable aléatoire " plutôt qu'"état", "observable"). De même, l'organisation des exposés est guidée par le développement probabiliste, et non par celui des idées physiques. Une première rédaction de ce texte a un peu circulé : celle que l'on trouvera ici est notablement plus courte sur les questions générales touchant aux fondements de la mécanique quantique, qui finalement ne nous servaient à rien.

I . Ω , $\underline{\underline{A}}$, P .

Ce paragraphe contient les définitions fondamentales de l'axiomatique de von Neumann, avec des commentaires.

1. L'ensemble Ω des probabilistes est ici remplacé par un espace de Hilbert complexe . Le produit scalaire sur Ω , noté $< , >$, sera supposé linéaire par rapport à la seconde variable, antilinéaire par rapport à la première : lorsque Ω est de la forme $L^2_{\mathbb{C}}(E,\underline{\underline{E}},\mu)$, où μ est une mesure positive, le produit scalaire sera donc

$$< f,g > \; = \int_E \overline{f}(x)g(x)\mu(dx) \; .$$

Nous supposerons toujours que Ω est <u>séparable</u>.

Les physiciens utilisent souvent la notation de Dirac, qui s'interprète ainsi : le choix d'une base orthonormale (e_i) de Ω étant sous-entendu, on peut identifier $x = \Sigma_i\, x^i e_i \in \Omega$ à la matrice colonne $|x> = (x^i)$, et aussi à la matrice ligne $<x| = x^* = (\overline{x}_i)$ où $x_i = \Sigma_j\, x^j \delta_{ji}$. Le produit scalaire $<x|y>$ ("bracket") est alors une notation contractée pour $<x||y> = x^* y$, le produit de la matrice ligne $<x|$ ("bra") par la matrice colonne $|y>$ ("ket"). L'antilinéarité par rapport à la <u>première</u> variable provient alors des conventions usuelles du calcul matriciel.

2. La tribu $\underline{\underline{A}}$ des probabilités classiques va être représentée ici par l'ensemble des <u>sous-espaces fermés</u> de Ω ; plus précisément, ceux-ci correspondent aux événements, l'indicatrice de l'événement A étant remplacée par le projecteur (orthogonal) I_A correspondant.

L'ensemble de ces événements est muni d'une structure qui rappelle la structure familière des tribus : l'inclusion $\subset$ correspond à l'inclusion usuelle des sous-espaces (pour les projecteurs, $A \subset B \iff I_A I_B = I_B I_A = I_A$) ; l'événement minimal (impossible) correspond au sous-espace $\{0\}$, au projecteur 0 ; l'événement maximal (certain) à Ω et au projecteur I . L'opération qui correspond à $\cap$, ici notée $\wedge$, se traduit par l'intersection des sous-espaces, et celle qui correspond à $\cup$, notée $\vee$, s'obtient en considérant le sous-espace fermé engendré par la réunion. Enfin, à l'opération ^{c} (complémentaire) correspond l'opération $A \longmapsto A^\perp$ ou $I_A \longmapsto I - I_A$ (orthogonal). On remarquera qu'à bien des égards, la situation est plus simple qu'en probabilités classiques : il n'y a pas d'ensembles non-mesurables, les opérations $\vee, \wedge$ ont un sens sans restriction de dénombrabilité.

COMMENTAIRES. a) La grande différence avec la logique usuelle est l'absence de <u>distributivité</u>. Par exemple, on a bien pour tout événement (sous-espace fermé) A

$$A \wedge A^\perp = \{0\} \quad , \quad A \vee A^\perp = \Omega$$

mais cela ne permet pas de décomposer un événement quelconque B suivant A et $A^\perp$, i.e. d'écrire que $B = (B \wedge A) \vee (B \wedge A^\perp)$: les événements B qui possèdent cette propriété sont exactement ceux pour lesquels I_B et I_A <u>commutent</u>. Par exemple, dans la fameuse expérience idéale des deux fentes d'Young, qui depuis Feynman figure au début de tous les cours de mécanique quantique, A ($A^\perp$) pourra être l'événement " la particule passe par la première (seconde) fente " et B l'événement " la particule arrive en une certaine région de l'écran " ; l'expérience contredit un usage naïf de la phrase " la particule arrivant sur l'écran a dû passer soit par la première fente, soit par la seconde ".

b) Comme en probabilités classiques, on ne peut se contenter de travailler sur l'énorme tribu de tous les événements possibles, on doit distinguer

des sous-tribus. Or il se produit ici un phénomène nouveau : partant d'un ensemble de projecteurs, on ne peut se borner à le stabiliser pour les opérations de la "logique" : on voudrait aussi pouvoir considérer comme des v.a. les combinaisons linéaires finies de projecteurs (au moins à coefficients réels), et considérer les événements naturellement liés à de telles v.a.. Nous étudierons plus loin explicitement le cas de deux projecteurs ne commutant pas, et nous verrons que cela introduit inévitablement une infinité non dénombrable de projecteurs...

La notion d'algèbre d'événements perd donc de son importance, au profit de la notion d'algèbre d'opérateurs bornés, stable pour l'opération * de passage à l'adjoint. Les C^*-algèbres jouent le rôle des algèbres de fonctions continues, souvent utilisées en théorie de la mesure, tandis que les algèbres de von Neumann jouent le rôle des tribus. Pour l'instant, nous travaillerons uniquement sur l'ensemble de tous les événements possibles, ou sur l'algèbre de tous les opérateurs bornés, ce qui nous évitera de recourir à ces notions un peu trop raffinées.

c) Une conséquence de la définition des événements, sans aucun parallèle en probabilités classiques, est la suivante : tout sous-espace fermé étant somme directe de sous-espaces orthogonaux de dimension 1, tout événement se trouve être une " réunion dénombrable d'événements élémentaires disjoints ", un peu comme si tout borélien était dénombrable !

Par analogie avec les notations classiques, nous désignerons par $\{\omega\}$, dans cet exposé, le sous-espace $\mathbb{C}\omega$ de dimension 1 engendré par un vecteur non nul ω . Cette notation, qui prête à confusion avec le singleton $\{\omega\}$ de la théorie des ensembles, ne sera pas utilisée après les premiers paragraphes.

3. Tout naturellement, une mesure positive μ (une loi) sera une fonction d'événement, positive, dénombrablement additive (telle que $\mu(\Omega)=1$). Le mot état est utilisé par les physiciens comme synonyme de loi . Puisque tout sous-espace est somme directe de sous-espaces de dimension 1, une mesure est déterminée par sa valeur sur ceux-ci (un peu à la manière des points d'un espace probabilisé classique dénombrable) .

L'analogue de la mesure classique qui, sur un espace dénombrable, associe à tout point la masse 1, est ici la mesure qui à tout sous-espace associe sa dimension. Nous ne faisons que la mentionner en passant, car nous ne travaillerons que sur des mesures bornées.

Exemple fondamental. Soit $\omega\in\Omega$. Pour tout sous-espace fermé A posons $\mu(A) = \langle \omega, I_A\omega \rangle = \|I_A\omega\|^2$; μ est manifestement une mesure positive de masse totale $\mu(\Omega) = \|\omega\|^2$, donc une loi si ω est de norme 1 . On notera que cette mesure ne change pas si l'on remplace ω par $c\omega$, avec $|c|=1$.

Dans la situation où nous nous sommes placés , où l'on n'impose aucune restriction aux projecteurs de la tribu ou aux vecteurs de l'espace, on peut montrer que l'on obtient ainsi les points extrémaux de l'ensemble des lois de probabilité (comme les ε_x en probabilités classiques). Ces lois sont dites pures[1], et nous utiliserons pour les désigner la notation ε_ω, qui est suggestive et peu dangereuse. On a $\varepsilon_\omega = \varepsilon_{c\omega}$ si $|c|=1$.

Il y a une différence essentielle avec la situation classique : les lois pures ne donnent pas une réponse déterministe (0 ou 1) à toutes les questions : si ω est un vecteur normalisé, A un sous espace fermé, $\varepsilon_\omega(A)$ vaut 1 si $\omega\in A$, 0 si $\omega\in A^\perp$, mais pour ω en position générale prend des valeurs quelconques de $]0,1[$. Par exemple, si A est le sous-espace $\{\phi\}$ engendré par un vecteur normalisé ϕ, on a $\varepsilon_\omega(A) = |\langle\omega,\phi\rangle|$.

On peut évidemment, à partir des lois pures, définir de nouvelles lois par mélange $\int\varepsilon_\omega m(d\omega)$, où m est une loi de probabilité au sens classique sur la boule unité de Ω (compacte métrisable pour sa topologie faible) portée par la sphère unité Ω_1 (borélienne). Si l'on désigne par μ ce mélange, on a pour tout vecteur normalisé ϕ

$$\mu(\{\phi\}) = \int\varepsilon_\omega(\{\phi\})m(d\omega) = \int|\langle\omega,\phi\rangle|^2 m(d\omega) = q(\phi)$$

où $q(.)=\int|\langle\omega,.\rangle|^2 m(d\omega)$ est une forme quadratique positive sur Ω . On a alors pour tout sous-espace fermé A

$$\mu(A) = \Sigma_i\, q(e_i) \text{ , où } (e_i) \text{ est une base orthonormale de } A \text{ ,}$$

expression qui ne dépend pas de la base choisie, et se désigne par la notation $\mathrm{Tr}(q|_A)$, la trace de $q|_A$ (q restreinte à A) par rapport à la forme quadratique fondamentale donnée par le produit scalaire. En particulier, si l'on prend $A=\Omega$, on voit que $\mathrm{Tr}(q)=1$.

La forme quadratique q est continue, et il existe un unique opérateur borné W tel que l'on ait

$$q(x) = \langle x, Wx\rangle \quad (\text{ par polarisation, } \langle x,Wy\rangle = \int\langle x,\omega\rangle\langle\omega,y\rangle m(d\omega))$$

Cet opérateur est autoadjoint positif, de trace 1 (on l'appelle parfois opérateur de densité, opérateur statistique). Il a l'avantage de fournir une représentation explicite et unique de la loi μ , alors que la mesure m définissant le mélange n'était pas du tout unique. Par exemple, si ω est un vecteur normalisé, on a $q(x)=|\langle x,\omega\rangle|^2$ et l'opérateur W est le projecteur sur $\{\omega\}$, $Wx=\langle\omega,x\rangle\omega$ (aussi noté $|\omega\rangle\langle\omega|$ chez Dirac).

---

1. La notion d'état pur est définie par la propriété d'extrémalité, et les états purs ne sont associés à des vecteurs unitaires, en général, que si l'on ne restreint ni les événements, ni les vecteurs (par des règles de supersélection). Je remercie vivement R.F. Streater d'avoir rectifié des erreurs de la première version sur ce point, et sur quelques autres.

4. Voici un exemple pathologique, illustrant le fait (déjà mentionné au n°2) que les projecteurs ne sont pas le point de départ naturel en probabilités non commutatives.

On prend $\Omega=\mathbb{C}^2$. Alors il y a un seul espace de dimension 0, un seul espace de dimension 2, et toute une famille d'espaces de dimension 1 sur lesquels les opérations $\wedge$ et $\vee$ sont triviales. La seule condition imposée à une mesure de probabilité au sens précédent est que $\mu(A)+\mu(A^{\perp}) = 1$ pour tout sous-espace de dimension 1. Il existe donc un ensemble énorme de lois de probabilité qui ne sont pas des mélanges. Mais par ailleurs, on peut montrer que seuls les mélanges se prolongent bien des projecteurs aux opérateurs quelconques par additivité.

En fait, on peut montrer que cette situation est particulière à la dimension 2 ! C'est un résultat non trivial, dû à Gleason, dont la démonstration a été récemment simplifiée (1984) par R. Cooke, M. Keane et W. Moran :

THEOREME. Dès que $\dim(\Omega)\geq 3$, toute loi de probabilité sur Ω est un mélange.

Désormais, nous oublierons l'exemple pathologique ci-dessus, et nous définirons les lois de probabilité comme des mélanges. Si W est l'opérateur statistique associé à la loi, on peut alors définir l'espérance de tout opérateur borné A par la formule $E_{\mu}[A] = \mathrm{Tr}(AW) = \mathrm{Tr}(WA)$ (voir l'appendice).

II. VARIABLES ALEATOIRES, ETC.

1. Soit $(E,\underline{E})$ un espace mesurable classique, et soit X une v.a. réelle.

La v.a. X est uniquement définie par la famille des sous-ensembles mesurables $\{X\leq t\}=E_t$, famille croissante, continue à droite, d'intersection vide et de réunion E. Si de plus μ est une loi sur E, la connaissance des probabilités $\mu(E_t)$ nous donne la fonction de répartition de X, et du même coup toutes les informations probabilistes utiles.

En probabilités quantiques, il est tout naturel de définir une v.a. X comme une famille croissante et continue à droite de sous-espaces fermés de Ω , qu'il n'y a aucun inconvénient à noter $\{X\leq t\}$, d'intersection réduite à $\{0\}$ et de réunion dense dans Ω . Nous désignerons par J_t le projecteur sur $\{X\leq t\}$. En analyse hilbertienne, une telle famille de sous-espaces ou de projecteurs est appelée une famille spectrale. Les physiciens emploient le mot d'observable de préférence à celui de v.a.. (On dit aussi : résolution de l'identité)

Si μ est une loi (i.e. un état) la fonction $t\longmapsto\mu(\{X\leq t\})$ est la fonction de répartition d'une loi de probabilité au sens classique sur $\mathbb{R}$, que nous appellerons la loi de X sous μ (ou dans l'état μ). Comme d'habitude, on dira que X est intégrable, appartient à L^p, etc. si la loi de X admet un moment d'ordre 1, d'ordre p ... En particulier, si $\mu=\varepsilon_{\omega}$ (état pur

la fonction de répartition de X est $< \omega, J_t\omega >$.

Soit x un vecteur de Ω . La courbe $x_t = J_t x$ à valeurs dans Ω est un peu analogue à une martingale de carré intégrable, et la fonction croissante $< x, J_t x > = \|x_t\|^2$ analogue au crochet de cette martingale, c'est pourquoi nous la noterons $<x,x>_t$. Comme en théorie des martingales, on peut "polariser" le crochet en posant $<x,y>_t = < x_t, y_t >$. Comme en théorie des martingales encore, on peut définir de manière unique, par isométrie, pour toute fonction borélienne bornée f sur $\mathbb{R}$ (réelle ou complexe) un opérateur analogue à l'intégrale stochastique, $J_f = \int f(s) dJ_s$, tel que

$$< x, J_f y > = \int f(s) d<x,y>_s$$

l'intégrale au second membre étant une intégrale de Stieltjes. Tout cela est presque évident pour qui connaît l'intégrale d'Ito ! Notons les propriétés suivantes (où * désigne le passage à l'adjoint)

$$J_1 = I \ , \ J_{f+g} = J_f + J_g \ , \ J_{fg} = J_f J_g \ , \ J_{\overline{f}} = J_f^* \ , \ J_t = \int I_{\{s \leq t\}} dJ_s$$

En particulier, si f est l'indicatrice I_A d'un ensemble borélien, J_f (que nous noterons aussi J_A) est un projecteur. Soit alors h une fonction borélienne réelle ; nous pouvons définir une nouvelle v.a. Y en convenant que le projecteur $I_{\{Y \leq t\}}$ est égal à $J_{\{h \leq t\}}$, et la loi de Y sous μ est l'image par h de la loi de X sous μ , comme en probabilités classiques. Il n'y a aucun inconvénient à noter $Y = h(X)$ comme d'habitude.

Finalement, notre famille spectrale $t \longmapsto J_t$ a été prolongée en une <u>mesure à valeurs projecteurs</u>, ou <u>mesure spectrale</u>, $A \longmapsto J_A$, A parcourant la tribu borélienne de $\mathbb{R}$. De plus, nous avons <u>intégré</u> ci-dessus les fonctions boréliennes bornées par rapport à la mesure spectrale. Tout cela est très facile.

2. Soit $(E, \underline{\underline{E}})$ un espace mesurable. Nous appellerons <u>v.a.</u> X <u>à valeurs dans</u> E (les physiciens ne disent malheureusement pas <u>observable à valeurs dans</u> E) une <u>mesure spectrale</u> $A \longmapsto J_A^X$, où A parcourt la tribu $\underline{\underline{E}}$. Autrement dit, une famille de projecteurs possédant les propriétés suivantes :

$J_\emptyset^X = 0$, $J_E^X = I$, $J_{A \cap B}^X = J_A^X J_B^X$, $J_{A \cup B}^X = J_A^X + J_B^X$ si $A \cap B = \emptyset$, $J_{A_n}^X \downarrow 0$ si $A_n \downarrow \emptyset$.

On peut définir les J_f^X pour f $\underline{\underline{E}}$-mesurable bornée, avec des propriétés analogues à celles que l'on a écrites plus haut ; pour un bon espace mesurable $(E, \underline{\underline{E}})$, cela n'exige aucune théorie nouvelle : il suffit de plonger $(E, \underline{\underline{E}})$ dans $\mathbb{R}$! Si μ est une loi quantique, l'application $A \longmapsto \mu(J_A^X)$ est une loi de probabilité sur E , que l'on appelle la <u>loi de la v.a.</u> X .

<u>Exemple fondamental</u>. Soit $(E, \underline{\underline{E}}, P)$ un espace probabilisé, et soit $\Omega = L^2(\mu)$. On définit une v.a. X sur Ω, à valeurs dans E, en posant

$$J_A^X(f) = I_A f \quad \text{pour tout} \quad f \in L^2$$

La loi de cette v.a. sous la loi quantique ε_f est la loi de probabilité

$A \longmapsto < f, J_A f > = \int_A |f|^2(x)P(dx)$ sur E . En particulier, si f=1, la loi de cette v.a. est exactement P .

En probabilités classiques, étant données deux v.a. réelles X et Y, il est possible de définir une v.a. Z=(X,Y) à valeurs dans $\mathbb{R}^2$, qui admet X et Y comme marges. En probabilités quantiques, remarquons que si U et V sont deux parties boréliennes de $\mathbb{R}^2$, on a $J_U^Z J_V^Z = J_{U\cap V}^Z = J_V^Z J_U^Z$: deux projecteurs quelconques d'une mesure spectrale commutent. Prenant U=A×ℝ, V=ℝ×B, nous voyons que J_A^X et J_B^Y doivent commuter, quelles que soient les parties A,B (boréliennes) de ℝ . On dit alors que les v.a. X et Y elles mêmes commutent. Ainsi

> En probabilités quantiques, deux v.a. X et Y ne peuvent être considérées comme les marges d'une même v.a. que si elles commutent.

Cette condition nécessaire est aussi suffisante : la démonstration de ce fait est très voisine de celle du théorème des bimesures en probabilités classiques (cf. par ex. Dellacherie-Meyer, Probabilités et potentiels A, III.74). Plus généralement, les v.a. d'une famille quelconque (X_i) admettent une loi jointe si elles commutent, et peuvent alors être considérées comme formant un processus stochastique au sens classique.

Soient X et Y deux v.a. réelles qui commutent : rien n'est plus facile que de définir leur somme, en les considérant comme marges d'une même v.a. Z à valeurs dans $\mathbb{R}^2$: le sous-espace $\{X+Y \leqq t\}$ s'écrit $\{Z \in \{x+y \leqq t\}\}$. En revanche, l'addition de deux v.a. réelles qui ne commutent pas est une opération délicate, sans signification intuitive immédiate, et tout de même fondamentale. Elle va nous amener à associer à toute v.a. un opérateur linéaire unique, l'addition des v.a. correspondant à l'addition des opérateurs.

3. Nous commençons par le cas où la v.a. X est bornée. Cela signifie qu'il existe deux nombres $a \leqq b$ tels que $J_t=0$ pour $t<a$, $J_t=I$ pour $t \geqq b$. Pour tout polynôme f , la fonction f(t) étant bornée sur [a,b], nous pouvons définir l'opérateur $\int f(t)dJ_t^X = J_f$. En particulier, nous poserons

$$\mathfrak{X} = \int t dJ_t^X \quad .$$

qui est borné autoadjoint. Pour tout polynôme f , $f(\mathfrak{X})$ est un opérateur borné défini de manière évidente, et c'est précisément $\int f(t)dJ_t^X$: l'habitude est d'identifier X et $\mathfrak{X}$, d'écrire f(X) pour $f(\mathfrak{X})$, de noter plus généralement f(X) pour J_f lorsque f est borélienne. L'opérateur $\mathfrak{X}$ est positif ($\forall \omega$, $<\omega, \mathfrak{X}\omega > \geqq 0$) si et seulement si la v.a. X est positive ($J_t^X = 0$ pour $t<0$).

Inversement, si Y est un opérateur borné autoadjoint (i.e. $<\omega, Y\phi>=<Y\omega, \phi>$ pour $\omega, \phi \in \Omega$), on peut montrer sans grandes difficultés, par passage à la limite à partir du cas des polynômes, qu'il existe un calcul symbolique borélien sur Y, i.e., que l'on peut définir raisonnablement f(Y) pour f

borélienne (même complexe) sur $\mathbb{R}$, de telle sorte que pour $f=1$ $f(Y)=I$, pour $f(t)=t$ $f(Y)=Y$, $(f+g)(Y)=f(Y)+g(Y)$, $(fg)(Y)=f(Y)g(Y)$, $\overline{f}(Y)=(f(Y))^*$, et que l'on ait aussi une propriété de continuité que nous n'expliciterons pas (au sens de la topologie forte des opérateurs). Il en résulte que les opérateurs $J_t^Y=I_{]-\infty,t]}(Y)$ forment une famille spectrale, et que l'on peut associer une v.a. à tout opérateur a.a. borné. Plus généralement, on peut associer une v.a. complexe à tout opérateur borné normal (i.e. commutant à son adjoint). Ces résultats ne sont pas difficiles.

Cela permet de définir diverses opérations sur l'ensemble des v.a. bornées : la somme $X+Y$ d'abord : la somme de deux opérateurs a.a. bornés est évidemment un opérateur a.a. borné, mais on n'a aucune relation simple permettant d'obtenir une information sur la loi de la somme $X+Y$ lorsque X et Y ne commutent pas (exceptée l'additivité des espérances). Par exemple, on peut parfaitement avoir dans un état donné $Y=0$ p.s., et des lois différentes pour X et $X+Y$!

Deux autres opérations importantes préservant le caractère a.a. sont les applications

$$(X,Y) \longmapsto i[X,Y] = i(XY-YX) \quad , \quad (X,Y) \longmapsto XY+YX \text{ (parfois noté } \{X,Y\} \text{).}$$

Elles n'ont pas de signification probabiliste simple.

5. Nous restons dans le cas des v.a. bornées pour établir les fameuses relations d'incertitude. Soient X,Y deux observables (v.a.) bornées, U et V les v.a. $XY+YX$ et $i(XY-YX)$. Nous écrivons l'inégalité de Schwarz sous la forme

$$\begin{aligned} \langle X\omega, X\omega\rangle\langle Y\omega, Y\omega\rangle &\geq |\langle X\omega, Y\omega\rangle|^2 = (\mathrm{Re}\langle\ \rangle)^2+(\mathrm{Im}\langle\ \rangle)^2 \\ &= \tfrac{1}{4}\langle \omega, U\omega\rangle^2 + \tfrac{1}{4}\langle \omega, V\omega\rangle^2 . \end{aligned}$$

Ce qui s'écrit encore, sous la loi ε_ω

$$(1) \qquad E_\omega[X^2]E_\omega[Y^2] - \tfrac{1}{4}E_\omega[U]^2 \geq \tfrac{1}{4}\langle \omega, V\omega\rangle^2 = \tfrac{1}{4}E_\omega[V]^2 .$$

Les relations d'incertitude s'obtiennent en oubliant le terme négatif au second membre d'une part, et d'autre part en remplaçant X par $X-E_\omega[X]$, Y par $Y-E_\omega[Y]$, ce qui ne modifie pas le commutateur, donc le second membre. On aboutit ainsi à une borne inférieure pour le produit des variances de deux v.a. ne commutant pas.

Mais il est intéressant aussi d'interpréter la formule (1) complète, et de l'étendre à des lois non nécessairement pures. Introduisons la forme quadratique

$$\phi_\omega(r,s) = E_\omega[(rX+sY)^2] = r^2a(\omega) + 2rsb(\omega)+s^2c(\omega) \quad \text{où} \quad a(\omega)=E_\omega[X^2], \text{ etc.}$$

La relation (1) nous dit que, dans tout état pur ε_ω , le discriminant $\Delta(\omega) = ac-b^2$ satisfait à l'inégalité

$$(2) \qquad \sqrt{\Delta(\omega)} \geq \tfrac{1}{2}|E_\omega[V]| \quad .$$

Or la fonction $\sqrt{ac-b^2}$ sur le cône des matrices $\left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right) \geqq 0$ est concave. Pour le voir , il suffit de vérifier que si A,B sont des matrices strictement positives, la fonction $t \longmapsto \sqrt{\det(A+tB)}$ est concave dans tout intervalle où A+tB est positive. Par une transformation $A \longmapsto U^*AU$, $B \longmapsto U^*BU$ on peut se ramener au cas où B=I. Les valeurs propres de A étant réelles, det(A+tI) a ses racines réelles, et l'hyperbole $y^2=\det(A+xI)$ est placée comme il faut.

Ayant fait cela, on peut intégrer (2) par rapport à une mesure m(dω) pour étendre aux mélanges la relation (2), puis (1), et enfin la relation d'incertitude elle même.

Note. Pour les états purs, la forme des relations d'incertitude que nous avons donnée remonte à Schrödinger , Sitz. Preuss. Akad. Wiss. 1930 (référence empruntée à un preprint de S. Golin, BiBoS Bielefeld n°17). Pour les états non purs, je l'ai apprise dans Cushen-Hudson, a quantum mechanical central limit theorem, J. Appl. Prob. 8, 1971. Il en existe des formes plus raffinées, faisant intervenir une fonctionnelle du type "entropie" [1].

La démonstration pour les états non purs est peu satisfaisante. En voici une autre (pour les détails, voir l'appendice). Soit HS l'espace de Hilbert des opérateurs de Hilbert-Schmidt, avec son produit scalaire $\langle U,V\rangle=\mathrm{Tr}(U^*V)$. L'opérateur de densité W définissant la loi de probabilité ($E[X]=\mathrm{Tr}(WX)$ pour toute v.a. bornée X) peut se mettre sous la forme K^*K, où K est un élément de HS de norme 1 (puisque W est positif, on peut prendre $K=K^*=\sqrt{W}$). Associons maintenant à notre v.a. X un opérateur $\mathfrak{X}$ sur HS par $\mathfrak{X}U=UX$ (à droite, l'opération est la composition des opérateurs). Nous avons

$$< U,\mathfrak{X}V >_{HS} = \mathrm{Tr}(U^*VX) = \mathrm{Tr}(XU^*V) = \mathrm{Tr}((UX)^*V) = < \mathfrak{X}U,V >_{HS}$$

donc $\mathfrak{X}$ est a.a.. Si X est un projecteur ($X^2=X$) , $\mathfrak{X}$ est un projecteur, et si $X=\int t dE_t$ (projecteurs spectraux dans Ω) on a $\mathfrak{X} = \int t d\mathfrak{E}_t$ avec les projecteurs correspondants. Enfin, la loi de X sous W nous est donnée par

$$P\{X\leqq t\} = \mathrm{Tr}(WE_t) = \mathrm{Tr}(K^*KE_t) = < K,\mathfrak{E}_t K>_{HS} = P\{\mathfrak{X}\leqq t\}$$

la dernière probabilité étant calculée sous ε_K dans l'espace de Hilbert HS. Ce changement d'espace de Hilbert ramène donc les mélanges aux lois "pures", ou plus exactement aux lois associées à des vecteurs normalisés, et notre première démonstration s'applique.

6. Passons aux v.a. X non bornées. Soit J_t le projecteur sur $\{X\leqq t\}$. Nous définissons un opérateur $\mathfrak{X}$ de la manière suivante

- Son domaine $\mathcal{D}$ est exactement l'ensemble des $\omega\in\Omega$ pour lesquels l'intégrale $\int t^2 d\langle\omega,J_t\omega\rangle$ est convergente, c.à d. tels que $E_\omega[X^2]<\infty$.
- Pour $\omega\in\mathcal{D}$, $\mathfrak{X}\omega$ est défini par

$$< \phi, \mathfrak{X}\omega > = \int t d\langle\phi,J_t\omega\rangle \quad \text{pour tout } \phi\in\Omega$$

1. I. Bialynicki-Birula, J. Mycielski, Comm. Math. Phys. 44, 1975.

L'intégrale du côté droit est absolument convergente, et définit une forme linéaire continue en la variable ϕ, d'après l'inégalité (qui pour le probabiliste est très proche de l'inégalité de Kunita-Watanabe)

$$\int |f(t)||g(t)||d<\phi, J_t\omega>| \leq (\int |f(t)|^2 d<\phi, J_t\phi>)^{1/2} (\int |g(t)|^2 d<\omega, J_t\omega>|)^{1/2}$$

ici avec $f(t)=t$, $g(t)=1$.

On montre sans peine que le domaine $\mathcal{D}$ est dense, que l'opérateur $\mathfrak{x}$ est symétrique sur son domaine, i.e. $< \phi, \mathfrak{x}\omega > = < \mathfrak{x}\phi, \omega >$ pour tout couple d'éléments de $\mathcal{D}$. Mais il est beaucoup mieux que cela : il est autoadjoint au sens très précis suivant :

($\omega \in \mathcal{D}$ et $\mathfrak{x}\omega = \theta$) <=> (la forme linéaire $\phi \longmapsto < \mathfrak{x}\phi, \omega >$ sur $\mathcal{D}$ est continue, et $< \mathfrak{x}\phi, \omega > = < \phi, \theta >$)

Inversement, un théorème fondamental dit que, pour tout opérateur a.a. ξ il existe une v.a. X (une famille spectrale !) et une seule telle que l'opérateur $\mathfrak{x}$ associé soit égal à ξ . Comme dans le cas borné, on identifiera désormais v.a. et opérateur associé.

COMMENTAIRE. L'objet probabiliste intéressant est la famille spectrale, mais en fait, c'est l'opérateur qui a une signification mécanique ou physique, et la famille spectrale est construite à partir de celui-ci.

Il est donc fondamental de savoir vérifier qu'un opérateur donné est autoadjoint. Ou plus exactement (un opérateur étant presque toujours défini par extension à partir d'un domaine assez restreint sur lequel on sait bien calculer), qu'un opérateur symétrique défini sur un domaine dense admet une fermeture autoadjointe. Il y a une gigantesque littérature consacrée à des théorèmes de ce genre, à tous les degrés de généralité.

Nous n'aurons besoin dans la suite que de deux théorèmes, qui nous fourniront toutes les v.a. dont nous aurons besoin :

1) Soit $(U_s)_{s\in\mathbb{R}}$ un groupe à un paramètre, fortement continu, d'opérateurs unitaires de Ω , et soit iX son générateur : le domaine de X est l'ensemble des $\omega \in \Omega$ tels que $\frac{d}{ds} U_s\omega|_{s=0}$ existe au sens de la norme de Ω, et $X\omega = \frac{1}{i} \frac{d}{ds} U_s\omega|_{s=0}$. Alors X est a.a. .

Inversement, si X est un opérateur a.a., $X=\int t dJ_t$ (où les J_t sont les projecteurs spectraux), nous définirons dans un instant les opérateurs $\int f(t) dJ_t$ pour f borélienne complexe. Alors les opérateurs

$$U_s = \int e^{ist} dJ_t$$

constituent un groupe unitaire fortement continu, de générateur iX. La connaissance de ce groupe détermine complètement la loi μ de la v.a. X sous la loi ε_ω :

$$E_\omega[e^{isX}] = < \omega, U_s\omega > .$$

L'ensemble de ces deux résultats constitue le théorème de Stone.

Dans de très nombreuses situations physiques, l'opérateur X est non seulement a.a., mais a.a. positif (i.e., la v.a. est positive, la famille spectrale J_t satisfait à $J_t=0$ pour $t<0$), et l'observable correspondante a la signification physique de l'énergie (X est appelé l'hamiltonien[1]). on remplace alors les exponentielles complexes par des exponentielles réelles, la transformée de Fourier par une transformation de Laplace :

2) Soit $(P_s)_{s\in\mathbb{R}_+}$ un semi-groupe fortement continu de contractions positives de Ω , et soit $-X$ son générateur : le domaine de X est formé des $\omega\in\Omega$ tels que $\frac{d}{ds}P_s\omega|_{s=0}$ existe au sens de la norme, et $X\omega = -\frac{d}{ds}P_s\omega|_{s=0}$. Alors X est a.a. positif.

Inversement, pour tout opérateur a.a. positif, on définit un semi-groupe (P_s) du type ci-dessus en posant

$$P_s = \int e^{-ist} dJ_t \qquad (J_t \text{ famille spectrale de X })$$

et la connaissance du semi-groupe détermine la loi de X par

$$E_\omega[e^{-sX}] = < \omega, P_s\omega > \quad (s>0).$$

En pratique, Ω sera très souvent un espace $L^2(\mu)$, et (P_s) sera un semi-groupe de transition sousmarkovien du type couramment utilisé en théorie des processus de Markov.

7. Nous ne distinguons plus désormais l'opérateur X de sa famille spectrale (J_t^X). Pour toute fonction borélienne complexe f , définissons un opérateur noté $f(X) = \int f(t) dJ_t^X$, comme nous l'avons fait plus haut lorsque f était bornée.

- Le domaine de f(X) est exactement formé des $\omega\in\Omega$ tels que
$$\int |f(t)|^2 \, d<\omega, J_t\omega > < \infty ,$$
- Pour un tel ω , $f(X)\omega=\Theta$ est caractérisé par
$$< \phi, \Theta > = \int f(t) d<\phi, J_t\omega > \text{ pour tout } \phi\in\Omega .$$

Un tel opérateur est fermé, et son adjoint est $\overline{f}(X)$. Nous n'insisterons pas sur cette notion, qui est facile à utiliser. Lorsque l'on tronque à n la fonction f , on obtient des opérateurs bornés $f_n(X)=\int f(s) I_{\{|f|<n\}}(s) dJ_s^X$, et pour tout ω appartenant au domaine $f_n(X)\omega$ converge en norme vers $f(X)\omega$.

Nous allons appliquer cela aux relations d'incertitude, pour des opérateurs a.a. X et Y non nécessairement bornés, et en nous plaçant pour simplifier sous une loi pure ε_ω . Rappelons qu'il s'agit de minorer le produit $E_\omega[X^2]E_\omega[Y^2]$ (cf.(1)) : on peut donc supposer que les deux facteurs sont finis, ce qui revient à dire que ω appartient aux domaines de X et de Y. Dans ces conditions, on a comme dans le cas borné

$$E_\omega[X^2]E_\omega[Y^2] \geq |<X\omega, Y\omega>|^2 = (\mathrm{Re}< >)^2 + (\mathrm{Im}< >)^2$$

1. Plus exactement, en physique on écrit $U_t = e^{itX/\hbar}$, où X est l'hamiltonien.

On ne peut déduire cela directement de l'inégalité de Schwarz, mais on peut écrire l'inégalité déjà établie pour les opérateurs tronqués $X_n = \int_{-n}^{n} t dJ_t^X$, $Y_n = \ldots$ et passer à la limite. Négligeant le premier terme on obtient la relation d'incertitude

$$E_\omega[X^2]E_\omega[Y^2] \geq \frac{1}{4}|\langle X\omega, Y\omega\rangle - \langle Y\omega, X\omega\rangle|^2$$

qui est formellement $\frac{1}{4}|\langle\omega, i[X,Y]\omega\rangle|^2$ - mais on n'a pas eu besoin de définir ce commutateur de manière précise. Cette remarque, due à Kraus et Schröter, Int. J. of Th. Phys. 8, 1973, figure dans le livre de Beltrametti et Cassinelli, The Logic of Quantum Mechanics, Encycl. Math. Appl. n°15 .

III. NOYAUX .

1. Puisque les lois de probabilité sont remplacées ici par des opérateurs de densité, a.a. positifs à trace égale à 1, l'espace classique des mesures bornées est remplacé tout naturellement par l'espace des opérateurs à trace sur Ω . Nous en rappelons en appendice la définition précise, et quelques propriétés importantes. Pour souligner l'analogie classique, nous le désignerons ici par $\mathfrak{M}_\Phi(\Omega)$ (l'espace des opérateurs a.a. à trace, le cône positif, l'ensemble des lois de probabilité, sont notés $\mathfrak{M}$, $\mathfrak{M}^+$, $\mathfrak{M}_1^+ = \mathcal{P}$).

En probabilités classiques, étant donnés deux espaces mesurables $(E,\underline{\underline{E}})$, $(F,\underline{\underline{F}})$, un noyau sousmarkovien de E dans F est une application $x \overset{N}{\longmapsto} N_x$ associant à tout point de E une sous-probabilité N_x sur F (si N_x est une probabilité pour tout x, N est un noyau markovien) , de telle sorte que $x \longmapsto N_x(A)$ soit $\underline{\underline{E}}$-mesurable pour tout ensemble $A \in \underline{\underline{F}}$. L'application N se prolonge alors par intégration

- en une application $\lambda \longmapsto \lambda N$ ($\lambda N(dy) = \int \lambda(dx) N(x,dy)$) de $\mathfrak{M}(E)$ dans $\mathfrak{M}(F)$,
- en une application $f \longmapsto Nf$ ($Nf(x) = \int N(x,dy) f(y)$) de l'espace des v.a. bornées sur F dans l'espace des v.a. bornées sur E .

Une application mesurable de E dans F peut s'interpréter comme un noyau markovien N de E dans F, tel que pour tout $x \in E$ $N(x,.)$ soit une masse unité $\varepsilon_{n(x)}$ (autrement dit, un noyau qui transforme les lois pures en lois pures). On sait le rôle fondamental que jouent, en probabilités classiques, les semi-groupes (sous)markoviens, leurs générateurs, les évolutions markoviennes associées.

Toutes ces notions ont été étendues aux probabilités quantiques, avec plus ou moins de simplicité et d'utilité - car il ne suffit pas de poser des définitions, il faut encore qu'elles servent à quelque chose en physique ! Un livre très agréable à lire sur ces sujets, pour un mathématicien du moins, est celui de E.B. Davies , Quantum Theory of Open Systems, Academic Press 1976.

La lecture de ce paragraphe n'est pas indispensable pour la suite.

Nous nous bornerons ici à un peu de vocabulaire. Nous remplaçons notre premier espace mesurable classique $(E,\underline{\underline{E}})$ par l'espace de Hilbert Ω, muni de la << tribu >> de tous ses projecteurs. Quant au second espace, il est clair que nous obtenons deux notions de noyaux, selon que nous le prenons du type classique ou du type quantique. C'est la seconde notion qui est la plus importante.

1) La première notion est celle d'une application N associant, à tout $A\in\underline{\underline{F}}$, une v.a. $N(I_A)$ au sens quantique comprise entre 0 et 1, c'est à dire un opérateur a.a. positif compris entre 0 et I ($0 \leqq \langle\omega, N(I_A)\omega\rangle \leqq \langle\omega,\omega\rangle$ pour tout $\omega\in\Omega$). Nous exigeons une propriété d'additivité complète :

(3) $\quad A=\cup_n A_n$, A_n disjoints $\Rightarrow N(I_A) = \Sigma_n N(I_{A_n})$

série convergente pour la topologie forte des opérateurs. Dans ces conditions, on peut aussi définir un opérateur a.a. positif $N(f)$ compris entre 0 et 1 pour toute fonction mesurable f comprise entre 0 et 1, par intégration, etc. Nous ne donnerons pas de détails ici. Si les $N(I_A)$ étaient des projecteurs, l'additivité complète (3) entraînerait que tous ces opérateurs $N(f)$ commutent, mais il n'en est plus de même dans la situation où nous sommes maintenant.

Les v.a. quantiques comprises entre 0 et 1 sont souvent appelées effets, et Davies appelle observables les noyaux du type que nous venons de définir. Nous n'utiliserons aucun de ces deux mots dans la suite.

2) Soient Ω et Φ deux espaces de Hilbert, $\mathfrak{M}(\Omega)$, $\mathfrak{M}(\Phi)$ les espaces de mesures bornées correspondants. Il semble naturel d'appeler noyau sousmarkovien de Ω dans Φ une application linéaire positive N de $\mathfrak{M}(\Omega)$ dans $\mathfrak{M}(\Phi)$, qui diminue la trace (elle diminue alors la norme-trace, donc est continue). La situation est meilleure ici qu'en probabilités classiques, où l'on ne sait pas caractériser simplement les applications entre espaces de mesures provenant de noyaux. L'extension aux mesures complexes est immédiate. Nous verrons dans l'appendice que le dual de $\mathfrak{M}_\phi(\Omega)$ est $\mathfrak{L}(\Omega)$, l'espace de tous les opérateurs bornés, le dual de $\mathfrak{M}(\Omega)$ est $\mathfrak{L}_a(\Omega)$, l'espace des opérateurs a.a. bornés. Par transposition, N définit donc une application linéaire, continue, positive, de $\mathfrak{L}_a(\Phi)$ dans $\mathfrak{L}_a(\Omega)$ (et de même sans les $_a$), et cela correspond exactement à l'application $f \longmapsto Nf$ dans le cas classique.

Le mot opération est souvent utilisé pour désigner ce que nous appelons ici un noyau sousmarkovien. Nous ne l'emploierons pas.

Exemple : si A est un opérateur continu de Ω dans Ω, A^* son adjoint, on peut poser, pour tout opérateur borné T sur Ω

$$N(T) = ATA^*$$

Il est très facile de vérifier que cet opérateur est positif si T est positif, et nous verrons en appendice que $N(\cdot)$ diminue la trace si A^*A est une

contraction. Plus généralement, si les A_n sont des opérateurs bornés tels que $\Sigma_n A_n^* A_n$ soit une contraction , on définit un noyau sousmarkovien sur Ω en posant

$$N(T) = \Sigma_n A_n T A_n^* .$$

Un important théorème, dû à Stinespring , affirme que l'on obtient ainsi les noyaux sur Ω qui possèdent une propriété très raisonnable : la positivité complète (identique à la positivité dans un contexte commutatif, mais ici strictement plus forte). Cela signifie que si l'on fait la somme Ω_n de n copies de Ω , de sorte qu'un opérateur borné sur Ω_n se lit comme une matrice (n,n) d'opérateurs bornés sur Ω , l'extention naturelle de N en une application de $\mathfrak{M}(\Omega_n)$ dans lui-même est encore positive, quel que soit n. La positivité complète doit être considérée comme faisant partie de la définition des noyaux en probabilités quantiques. Voir K. Kraus, States, Effects and Operations, LN in Phys. 190, p. 42. Cf. surtout Evans et Lewis, Dilations of Irreversible Evolutions in Algebraic Quantum Theory, Dublin Inst for Adv. Studies 1977. Nous espérons revenir sur ce sujet.

IV . APPENDICE

Cet appendice tente de répondre à une question épineuse. Que faut il savoir au juste en analyse fonctionnelle pour aborder les travaux de probabilités quantiques ? La réponse évidente est : tout , beaucoup d'auteurs ne faisant (comme ailleurs en mathématiques) aucun effort pour être compris des non-spécialistes, et adoptant tout de suite le langage des algèbres de von Neumann... Or une expérience de quelques mois m'a montré que, dans la plupart des cas, on a besoin surtout d'un peu de vocabulaire, et de quelques résultats généraux, toujours les mêmes, et plutôt faciles et agréables à apprendre. Nous tâcherons peu à peu d'en faire la liste.

Cet appendice contient le vocabulaire des C^*-algèbres, dont on se servira avec plaisir dans les exposés ultérieurs, et quelques compléments sur les opérateurs bornés (surtout les opérateurs à trace). Pour les démonstrations, on pourra se reporter à Bratteli-Robinson, Operator Algebras and Quantum Statistical Mechanics, Springer, ou à Pedersen, C^*-Algebras and Their Automorphism Groups, Acad. Press. B-R est bien réduit à l'essentiel, les démonstrations parfois un peu lourdes. Pedersen est parfait, son seul défaut est d'être trop complet ! L'appendice ne contient rien sur les opérateurs non bornés : cf Reed-Simon, Methods of Modern Math. Physics, I-VIII.

1. Vocabulaire des C^*-algèbres. Les éléments de la théorie de Gelfand étant très facilement accessibles , nous supposerons connue la notion d'algèbre de Banach $\mathcal{A}$ sur $\mathbb{C}$ (à unité $\mathbf{1}$, norme notée $\| \|$) munie d'une involution notée * . Rappelons la définition du spectre d'un élément a de $\mathcal{A}$ (nous le notons Sp(a) : c'est l'ensemble des $\lambda \in \mathbb{C}$ tels que $a-\lambda\mathbf{1}$ ne soit pas

inversible) : c'est un compact non vide du plan complexe, et le rayon spectral $r(a)$ est par définition $\sup_{\lambda \in Sp(a)} |\lambda|$. On a

$$(1) \qquad r(a) = \inf_n \|a^n\|^{1/n} = \lim_n \|a^n\|^{1/n} \leqq \|a\| \ .$$

Soit A un opérateur borné dans un espace de Hilbert. On a

$$\|A\|^2 = \sup_{\|x\| \leqq 1} < Ax, Ax > = \sup_x < A^*Ax, x > \leqq \|A^*A\| \ .$$

Une algèbre de Banach à unité dans laquelle la norme possède cette propriété sera appelée C^*-algèbre, et nous n'en considèrerons pas d'autre. Une C^*-algèbre concrète est une sous*algèbre fermée de l'algèbre des opérateurs bornés $\mathcal{L}(H)$ d'un Hilbert H . Une algèbre de von Neumann est une C^*-algèbre concrète qui est fermée dans $\mathcal{L}(H)$ pour la topologie forte (ou faible, cela revient au même) des opérateurs.

Dans toute C^*-algèbre, on a en fait (très facilement)

$$(2) \qquad \|a^*a\| = \|a\|^2 \ , \quad \|a\| = \|a^*\| \ , \quad \|1\| = 1 \ .$$

Vocabulaire. Un élément a d'une C^*-algèbre est

- autoadjoint si $a=a^*$ - unitaire si $a^*a=aa^*=I$
- normal si $a^*a=aa^*$ - un projecteur si $a=a^*$, $a^2=a$.

Pour tout élément normal n on a $r(n)=\|n\|$. Pour tout b, b^*b est a.a., donc normal, donc $\|b\|^2=\|b^*b\|=r(b^*b)$. Or le rayon spectral s'exprime au moyen de la structure d'algèbre seule : celle-ci détermine donc entièrement la norme.

Un élément unitaire (resp. a.a.) a son spectre contenu dans le cercle unité (resp. l'axe réel).

Probablement, le plus important des résultats élémentaires de la théorie est celui qui permet le calcul symbolique continu : soit a un élément a.a., et soit P un polynôme ; on définit de manière évidente $P(a) \in \mathcal{G}$. Alors

THEOREME. Posons $\|P\|_a = \sup_{\lambda \in Sp(a)} |P(\lambda)|$. Alors $\|P(a)\| = \|P\|_a$, et l'application $P \mapsto P(a)$ se prolonge en un isomorphisme de l'algèbre des fonctions complexes continues sur $Sp(a)$ (avec la norme uniforme) sur l'algèbre fermée engendrée dans $\mathcal{G}$ par a et I . On a $f(a)^* = \overline{f}(a)$ et $Sp(f(a)) = f(Sp(a))$.

On a le même résultat pour a normal , à condition de considérer l'algèbre fermée engendrée par a, a^* et I , et des éléments $P(a,a^*)$, où $P(.,.)$ est un polynôme, et $\|P\|_a = \sup_{\lambda \in Sp(a)} |P(\lambda, \overline{\lambda})|$. Ce théorème figure dans à peu près tous les traités sur les algèbres de Banach.

> Le lecteur se demandera sans doute à quelle condition doit satisfaire une C^*-algèbre pour que le calcul symbolique borélien soit possible sur ses a.a. sans sortir de l'algèbre. Pour une C^*-algèbre concrète d'opérateurs sur un Hilbert séparable, cela caractérise les algèbres de von Neumann (Pedersen, C^*-algebras and their automorphism groups, Academic Press, th. 2.8.8.) de même que la propriété des classes monotones (th. 2.4.3).

On dit qu'un élément a.a. b est positif si son spectre est contenu dans $\mathbb{R}_+$: le calcul symbolique continu montre que b admet une racine carrée positive $a=\sqrt{b}$, et l'on montre que a est le seul élément a.a. positif tel que $a^2=b$. Un résultat trivial pour les C^*-algèbres concrètes, mais non trivial en général, est le très important

THEOREME. Pour tout $a\in\mathcal{A}$, a^*a est positif.

Inversement, tout élément positif b s'écrit a^*a, avec $a=\sqrt{b}$.

Un résultat bien utile est le suivant : tout élément a de norme ≤ 2 est somme de 4 unitaires. Il suffit de poser $q=\frac{1}{4}(a^*+a)$, $p=\frac{i}{4}(a^*-a)$, a.a. ; $u^{\pm}=q\pm i\sqrt{1-q^2}$, $v^{\pm}=ip\pm\sqrt{1-p^2}$; ces opérateurs sont unitaires, et leur somme est $2q+2ip=a$. Pour un joli raffinement, cf. Pedersen, th. 1.1.12.

Il faut se méfier de la notion d'ordre pour les opérateurs qui ne commutent pas. Par exemple, la relation $0\leq a\leq b$ entraîne $\sqrt{a}\leq\sqrt{b}$, mais non $a^2\leq b^2$! Les livres contiennent sur ce sujet des résultats techniques utiles.

On appelle loi de probabilité (ou état) sur la C^*-algèbre $\mathcal{A}$ une forme linéaire μ sur $\mathcal{A}$, positive ($\mu(a^*a)\geq 0$) et telle que $\mu(\mathbf{1})=1$. On a tout naturellement une inégalité de Schwarz

$$|\mu(a^*b)| \leq \mu(a^*a)^{1/2}\ \mu(b^*b)^{1/2} \tag{3}$$

et en faisant $b=\mathbf{1}$ et en remarquant que $a^*a\leq \|a\|^2\mathbf{1}$ on voit que μ est continue, de norme 1 . Tout naturellement, on appellera mesures (réelles) les formes linéaires continues λ telles que $\lambda(a)=\lambda(a^*)$, et l'on peut montrer que toute mesure est combinaison linéaire de deux lois. L'ensemble des lois de probabilité est convexe compact pour sa topologie faible : il contient donc beaucoup de points extrémaux, appelés lois pures. Par exemple, si $\mathcal{A}\subset\mathcal{L}(H)$ est une C^*-algèbre concrète, et x est un vecteur normalisé de H, $\mu(a)=\langle x,ax\rangle$ est une loi de probabilité, mais non nécessairement pure si $\mathcal{A}\neq\mathcal{L}(H)$.

On dit que la loi μ est fidèle, si $\mu(a^*a)=0 \Rightarrow a=0$ (cela veut dire aussi que si b est autoadjoint positif et $\mu(b)=0$, alors $b=0$). On dit que μ est une trace si $\mu(ab)=\mu(ba)$, ce qui équivaut (d'après le théorème de décomposition en quatre unitaires ci-dessus) à l'invariance unitaire de μ : $\mu(uau^*)=\mu(a)$ pour tout u unitaire et tout a. Il n'existe pas de loi tracia-le sur $\mathcal{L}(H)$ si H est de dimension infinie, mais les C^*-algèbres que nous rencontrerons par la suite admettront des traces - c'est une excellente raison de considérer d'autres algèbres que $\mathcal{L}(H)$!

Le théorème de Hahn-Banach permet d'établir l'existence de beaucoup de lois de probabilité, et le théorème de Krein-Milman celle de beaucoup de lois pures. En particulier, toute C^*-algèbre séparable admet une loi fidèle.

Soit μ une loi sur $\mathcal{A}$. Nous posons $\langle a,b\rangle_\mu=\mu(a^*b)$ et désignons par H_μ l'espace hilbertien séparé-complété de $\mathcal{A}$ pour ce produit scalaire.

A tout $a \in \mathcal{A}$ est associé un vecteur $\tilde{a}$ de H_μ (si μ est fidèle, le $\sim$ est inutile) et un opérateur borné T_a sur H_μ défini par

$$T_a(\tilde{b}) = (ab)^\sim \tag{4}$$

On définit ainsi une représentation de $\mathcal{A}$ dans $\mathcal{L}(H)$ (et l'on peut montrer que $\|T_a\|$ est exactement la norme de a dans $\mathcal{A}/I$, I étant le noyau : en particulier si μ est fidèle, $\|T_a\| = [\![a]\!]$). L'ensemble des $T_a\tilde{1}$ est dense dans H_μ : on dit que $\tilde{1}$ est un vecteur cyclique pour la représentation (et inversement, toute représentation admettant un vecteur cyclique s'obtient de cette manière). Si μ est fidèle, la relation $T_a\tilde{1}=0$ entraîne a=0, et le vecteur $\tilde{1}$ est dit séparant pour la représentation.

Cela suffit à fixer notre vocabulaire pour la suite. Tout ce qui vient d'être exposé (et qui est tout juste une introduction aux C^*-algèbres) constitue vraiment une belle et agréable théorie, polie par de nombreux rédacteurs successifs, et parvenue à sa forme définitive. Par ailleurs, l'axiomatisation << C^*-algébrique >> des probabilités quantiques semble être très généralement considérée comme la meilleure manière d'aborder les problèmes difficiles de mécanique statistique ou de théorie des champs.

2. Opérateurs à trace, etc.

On a vu que les opérateurs positifs de trace 1 jouent le rôle des lois de probabilité sur $\mathcal{L}(H)$. Il est donc important de faire un catalogue de leurs propriétés. Nous donnons quelques démonstrations en langage télégraphique.

a) Soit $a \in \mathcal{L}(H)$. Etant donnée une base o.n. (e_n), on pose

$$\|a\|_2 = (\Sigma_n \|ae_n\|^2)^{1/2} \leqq +\infty \tag{5}$$

A priori cela dépend de la base. Mais soit (e'_n) une seconde base. On a

$$\|a\|_2^2 = \Sigma_{nm} |\langle ae_n, e'_m\rangle|^2 = \Sigma_{nm} |\langle e_n, a^* e'_m\rangle|^2 = \|a^*\|_2^2$$

D'abord $e_n = e'_n$ nous donne $\|a\|_2 = \|a^*\|_2$, puis on voit que $\|a\|_2$ ne dépend pas de la base o.n.. On l'appelle la norme de Hilbert-Schmidt de a , et les opérateurs de norme HS finie forment l'espace (noté HS ou $\mathcal{L}^2$) des opérateurs de Hilbert-Schmidt. Par opposition, l'espace $\mathcal{L}(H)$ et sa norme seront notés parfois $\mathcal{L}^\infty$, $\| \ \|_\infty$. Tout vecteur unitaire pouvant entrer dans une base o.n., on a $\|a\|_\infty \leqq \|a\|_2$. Pour tout opérateur borné b on a d'après (5)

$$\|ba\|_2^2 = \Sigma_n \|bae_n\|^2 \leqq \|b\|_\infty^2 \|a\|_2^2 \quad \text{et} \quad \|ab\|_2^2 \leqq \|b\|_\infty^2 \|a\|_2^2 \tag{6}$$

par passage à l'adjoint. Ceci est très important ! En particulier, $\|ab\|_2 \leqq \|a\|_2 \|b\|_2$.

Il est clair sur (5) que la norme $\| \ \|_2$ est associée à un produit scalaire hermitien $\langle a,b\rangle_{HS} = \Sigma_{nm} \langle ae_n, be_m\rangle$. On montre sans peine que HS est complet. HS est aussi (avec l'involution * et la norme $\| \ \|_2$) une

algèbre de Banach d'un type bien particulier (algèbre de Hilbert), cette structure jouant un rôle en théorie de l'intégration non commutative.

Soit a un opérateur de HS, et soit (a_{nm}) sa matrice dans la base (e_n). Soit a(k) l'opérateur de rang fini obtenu en remplaçant a_{nm} par 0 si $n \geqq k$ ou $m \geqq k$: alors a(k) tend vers a en norme HS, donc en norme $\| \ \|_\infty$ (en particulier, a est un opérateur compact, mais nous n'aurons pas besoin de cela, je pense). Il y a encore beaucoup à dire sur les HS, mais nous n'aurons à utiliser que ce qui précède, qui est très facile.

b) $\mathcal{L}(H)$ est analogue à ℓ^∞, HS à ℓ^2, nous allons définir l'analogue de ℓ^1.

~ D'abord, soit a un opérateur borné <u>positif</u>, et soit $b=a^{1/2}$. On a dans toute base (e_n)

(7) $\quad \|b\|_2^2 = \Sigma_n < be_n, be_n > = \Sigma_n < e_n, ae_n > \leqq +\infty$

Ceci ne dépend pas de la base, et se note tr(a). Ainsi, pour un opérateur positif, tr(a) est toujours défini, unitairement invariant ($tr(uau^*)=tr(a)$ pour u unitaire : cela revient à changer (e_n)), fini ssi $a^{1/2} \in HS$.

~ Plus généralement, soit a un opérateur <u>produit de deux</u> HS : a=bc . On a $\Sigma_n \ |<e_n, ae_n>| = \Sigma_n |<b^* e_n, ce_n>| \leqq \|b\|_2 \|c\|_2 < \infty$, et

(8) $\qquad \Sigma_n < e_n, ae_n > = <b^*, c>_{HS}$ [1]

montrant que ceci est indépendant de la base (côté droit) et de la décomposition a=bc (côté gauche). L'idée est que les produits de deux HS sont les opérateurs <u>intégrables</u>, tr(a) définie par (8) étant l'<u>intégrale</u>. Alors (7) exprime la propriété familière que l'intégrale a un sens ($\leqq +\infty$) pour les fonctions positives. On va préciser cela.

~ Soit a un opérateur borné quelconque : a^*a est positif, on note |a| sa racine carrée (choix arbitraire : pourquoi pas aa^* ?). Un résultat élémentaire (et facile) affirme que l'on peut écrire a=u|a|, $|a|=u^*a$, où u est un opérateur de norme $\leqq 1$, unitaire si a est inversible, uniquement caractérisé par quelques propriétés simples (<u>décomposition polaire</u> de a).

THEOREME. (a <u>est un produit de deux</u> HS) <=> ($tr(|a|)<\infty$).

On écrira alors $a \in \mathcal{L}^1$, on dira que a est un <u>opérateur à trace</u> (ou <u>nucléaire</u>) , on posera $\|a\|_1 = tr(|a|)$.

<u>Dém</u>. Supposons $tr|a|<\infty$, soit $b=|a|^{1/2} \in HS$. Alors a=u|a|=(ub)b est un produit de deux HS. Inversement, soit a=hk un produit de deux HS. On a $|a|=u^*a = (u^*h)k$ et par (8)

(9) $\qquad \|a\|_1 \leqq \|h^*u\|_2 \|k\|_2 \leqq \|h\|_2 \|k\|_2 \quad (\ \|u\|_\infty \leqq 1 \ ).$

Noter aussi que si $b=|a|^{1/2}$, $|tr(a)| = |<b^*u, b>_{HS}| \leqq \|b^*u\|_2 \|b\|_2 \leqq \|b\|_2^2 = \|a\|_1$.

On en tire plusieurs conséquences intéressantes .

---

1. Noter que pour deux HS b,c , $<b,c>_{HS}$ s'écrit $tr(b^*c)$.

~ Si $a\in\mathfrak{L}^1$, $h\in\mathfrak{L}^\infty$, on a $ah\in\mathfrak{L}^1$, $\|ah\|_1\leq\|a\|_1\|h\|_\infty$. En effet, si $|a|^{1/2}=b$ on a (déc. polaire) $a=ubb$, $ah=(ub)(bh)$, $\|ah\|_1\leq\|ub\|_2\|bh\|_2\leq\|b\|_2\|b\|_2\|h\|_\infty$.

~ Si $a\in\mathfrak{L}^1$ on a $a^*\in\mathfrak{L}^1$ avec même norme. Ecrivons $a=u|a|$ (déc. polaire), d'où $a^*=|a|u^*$; le précédent donne $\|a^*\|_1\leq\|a\|_1$ et l'égalité. Par passage à l'adjoint, on peut récrire le précédent avec h à gauche.

~ Si $a\in\mathfrak{L}^1$, $h\in\mathfrak{L}^\infty$, on a $tr(ah)=tr(ha)$. En effet, si h est unitaire, $b=ha$, cela se ramène à $tr(h^{-1}bh)=tr(b)$, invariance unitaire de la trace. On en déduit le cas général, tout opérateur borné étant combinaison linéaire de 4 unitaires.

Voici maintenant le seul résultat non trivial de ce n° :

THEOREME. Soient $a,b\in\mathfrak{L}^1$. Alors $a+b\in\mathfrak{L}^1$ et $\|a+b\|_1\leq\|a\|_1+\|b\|_1$.

Dém. On écrit les trois décompositions polaires $a=u|a|$, $b=v|b|$, $a+b=w|a+b|$. Alors $|a+b|=w^*(a+b)=w^*u^*|a|+w^*v^*|b|$. On en déduit que pour une base o.n. (e_n)

$$\Sigma_1^k \langle e_n,|a+b|e_n\rangle \leq \Sigma_1^k \langle e_n,w^*u^*|a|e_n\rangle+\langle e_n,w^*v^*|b|e_n\rangle \leq \|a\|_1+\|b\|_1 .$$

Donc $tr(|a+b|)$ est bien finie, etc.

~ Si a est a.a., de décomposition spectrale (E_λ), il est facile de voir que $|a|$ ne peut avoir une trace finie que si le spectre est discret (λ_n), et $\|a\|_1=\Sigma_n|\lambda_n|\dim(E_{\lambda_n})$. On en déduit que $a\in\mathfrak{L}^1 \Rightarrow a^+,a^-\in\mathfrak{L}^1$, $\|a\|_1=Tr(a^+)+Tr(a^-)$ Ceci correspond à la décomposition de Jordan des mesures (réelles).

Avant de passer au n° suivant, remarquons que l'espace des opérateurs à trace a une double interprétation en théorie de la mesure classique : soit comme espace des fonctions intégrables (espace $\mathfrak{L}^1$) soit comme espace des mesures bornées (espace $\mathfrak{M}$) . La même situation se rencontre sur $\mathbb{N}$ pour ℓ^1 : comme espace de mesures, c'est le dual de c_0 (suites tendant à 0 à l'infini, adhérence en norme des suites finies). Comme espace de fonctions intégrables, son dual est ℓ^∞ .

3. Propriétés de dualité. Nous désignerons par l'indice $_r$ (pour réel) les sous-espaces de $\mathfrak{L}^1,\mathfrak{L}^2,\mathfrak{L}^\infty$ formés d'opérateurs a.a..

Nous désignerons par $\mathfrak{F}$ l'espace des opérateurs de rang fini, engendré par les opérateurs de rang 1 , c'est à dire

$$E_{xy} : E_{xy}(z) = \langle y,z\rangle x$$

On a $E_{xy}^*=E_{yx}$, $E_{xy}^*E_{yx}=\|y\|^2E_{xx}$, $\|E_{xy}\|_1=\|x\|\|y\|$. L'adhérence en norme de $\mathfrak{F}$ est l'espace des opérateurs compacts, mais nous n'avons pas besoin de le savoir.

THEOREME. Le dual de $(\mathfrak{F},\|.\|_\infty)$ est $(\mathfrak{M}=\mathfrak{L}^1,\|.\|_1)$ (en particulier, $\mathfrak{L}^1$ est complet !), et le dual de $(\mathfrak{L}^1,\|.\|_1)$ est $(\mathfrak{L}^\infty,\|.\|_\infty)$. On a le même énoncé pour les sous-espaces réels (=a.a.) correspondants.

Dans les deux cas, la forme de dualité est $(a,h)\mapsto tr(ah)$.

Dém. 1) Nous remarquons que pour un opérateur positif a, $tr(a)=\sup_{h\in\mathcal{F}_1} tr(ah)$. Cela se voit sur (7), en prenant pour h une matrice diagonale finie qui tend vers l'identité ($\mathcal{F}_1=\{h\in\mathcal{F}, \|h\|_\infty\leqq 1\}$). Si a n'est pas positif, on écrit $a=u|a|$, $|a|=u^*a$, $tr(|a|) = \sup_{h\in\mathcal{F}_1} tr(u^*ah) = \sup_h tr((hu^*)a)$ et comme $hu^*\in\mathcal{F}_1$ on a

$$\|a\|_1 \leqq \sup_{b\in\mathcal{F}_1} tr(ab) \text{ (d'où l'égalité) .}$$

Si a est a.a., on peut prendre b a.a., mais cela se voit par un calcul direct sur la décomposition spectrale.

2) Soit φ une forme linéaire continue sur $\mathcal{F}$ pour $\|.\|_\infty$. Alors φ est continue pour $\|.\|_2$ qui est plus forte, donc est de la forme $\varphi(h)=tr(ah)$ avec $a\in HS$. D'après 1) ci-dessus, on voit que $a\in\mathcal{M}$. Le reste de cette partie est facile.

3) Si $\|x\|=\|y\|=1$, on a en utilisant une base o.n. dont le premier élément est y , $hE_{xy}(e_n)=h(x)$ si $n=1$, 0 si $n\neq 1$, donc $tr(hE_{xy}) = \langle y,hx\rangle$. Donc

$$\|h\|_\infty \leqq \sup_{\|a\|_1\leqq 1} tr(ah) \text{ (d'où l'égalité) .}$$

Si h est a.a. on peut se borner à travailler sur les E_{xx} , donc sur des opérateurs a eux aussi autoadjoints.

4) Soit φ une forme linéaire continue sur $\mathcal{L}^1$. Alors $(y,x)\longmapsto\varphi(E_{xy})$ est une forme hermitienne continue, qui peut donc s'écrire $\langle y,hx\rangle$ pour un $h\in\mathcal{L}$; la norme de h est égale à celle de φ d'après le petit calcul ci-dessus. Les formes $\varphi(.)$ et $tr(h.)$ sont égales sur les E_{xy} , et il suffit de vérifier que ceux-ci forment un ensemble dense. Les détails sont faciles. □

Essayons d'adapter à la situation présente notre vieux langage probabiliste, plus suggestif que celui de l'analyse fonctionnelle. Nous dirons que des lois ρ_i convergent <u>vaguement</u> (suivant un filtre sur l'ensemble d'indices I) vers l'opérateur ρ si $tr(\rho_i a)\longrightarrow tr(\rho a)$ pour tout opérateur de rang fini a : il suffit de le vérifier pour les E_{xy} , et nous voyons que cela revient à la topologie faible des opérateurs. L'opérateur limite est positif et en prenant une base o.n. quelconque, on voit que $tr(\rho)\leqq 1$. Comme $\mathcal{L}^1$ est un dual, et comme il suffit pour identifier un opérateur de connaître ses éléments de matrice dans une base o.n., nous avons le résultat simple

L'ensemble des mesures positives de masse $\leqq 1$ est compact métrisable pour la topologie vague (= faible).

Nous dirons que les ρ_i convergent <u>étroitement</u> vers ρ si l'on a de plus $tr(\rho)=1$. Dans le cas classique, cela équivaut à la convergence de l'intégrale de toute fonction continue bornée. Ici, nous aurons

PROPOSITION. <u>Si</u> $\rho_i\longrightarrow\rho$ <u>étroitement</u> , $tr(\rho_i a)\longrightarrow tr(\rho a)$ <u>pour tout</u> a <u>borné</u>.

Inversement, d'ailleurs, la convergence de la trace correspond à $a=I$.

On voit donc que la topologie étroite n'est autre que la topologie faible $\sigma(\mathfrak{L}^1,\mathfrak{L}^\infty)$, du moins sur les ensembles bornés de mesures positives. Démonstration. Il suffit de traiter le cas d'un opérateur a.a. borné. Tout opérateur a.a. borné est limite en norme d'opérateurs a.a. bornés à spectre discret (approximation de Lebesgue usuelle + calcul symbolique borélien), donc on peut supposer que a admet une base o.n. qui le diagonalise. Dans cette base, on a $tr(\rho_i a)= \Sigma_n \lambda_n \rho_{nn}(i)$ (λ_n valeur propre de a suivant e_n , $\rho_{nn}(i)$ élément de matrice de ρ_i) et de même pour ρ . On est alors ramené à un problème classique de convergence étroite sur $\mathbb{N}$, d'ailleurs évident. □

Pour les suites, on a un résultat plus fort. Davies (Comm. M. Phys. 15, 1969, lemme 4.3, p.291) démontre que toute suite (ρ_n) étroitement convergente satisfait à une " condition de Prokhorov " du type suivant : pour tout $\varepsilon>0$, il existe un projecteur p de rang fini tel que l'on ait pour tout n $\|\rho_n - p\rho_n p\|_1 < \varepsilon$. On en déduit qu'en fait la convergence des ρ_n vers leur limite a lieu en norme trace. Ceci illustre la ressemblance entre la théorie des probabilités quantiques dans l'axiomatique de von Neumann et les probabilités classiques discrètes (cf. Dunford-Schwartz, Linear Operators I, Cor. 14 p. 296). Si au lieu de travailler sur la tribu de tous les projecteurs, ou l'algèbre de tous les opérateurs, on se place sur une algèbre de von Neumann, ce résultat cesse d'être vrai (il ne l'est déjà plus en probabilités commutatives non discrètes !). Ce sujet est étudié par Dell'Antonio, Comm. Pure Appl. M. 20, 1967.

4. Topologies faibles. Nous terminons cet exposé par quelques résultats très simples concernant, non plus $\mathfrak{L}^1=\mathfrak{M}$, mais $\mathfrak{L}(H)=\mathfrak{L}^\infty$. Ces résultats sont indispensables pour la théorie des algèbres de von Neumann.

Tout le monde connaît la topologie faible des opérateurs ($(a_i \to a) \Leftrightarrow (\langle y,a_i x\rangle \to \langle y,ax\rangle$ pour tout couple (x,y)), autrement dit les semi-normes fondamentales sont les $p_{xy}(a)= |\langle y,ax\rangle|$), et la topologie forte des opérateurs ($(a_i \to a) \Leftrightarrow (a_i x \to ax$ en norme pour tout x), semi-normes fondamentales $q_x(a)=\|ax\|$). On a la proposition suivante, qui est utile parce qu'elle permet d'appliquer le th. de Hahn-Banach :

PROPOSITION. Le dual de $\mathfrak{L}$ est le même pour les deux topologies.

Ce dual est explicitement connu : il est très simple de voir qu'il est formé des combinaisons linéaires de formes $a \mapsto \langle y,ax\rangle$, ou encore (de manière un peu plus pédante) de formes $a \mapsto tr(ah)$, où h est de rang fini (Bourbaki, EVT II, § 6, prop.3 : mais en fait c'est redémontré ci-dessous).

Dém. Soit f une forme continue sur $\mathfrak{L}$ pour la top. forte. Il existe des x_i en nombre fini tels que ($\sup_i \|ax_i\| \leqq 1$) $\Rightarrow$ ($|f(a)| \leqq 1$). Soit p le

projecteur sur le sous-espace engendré par les x_i : la relation $ap=0$ entraîne $ax_i=0$ pour tout i donc $f(a)=0$, et il en résulte que $f(a)=f(ap)$ pour tout a . Ecrivons p sous la forme $p(.)=\Sigma_k \langle y_k,.\rangle y_k$ et posons pour tout x $h^k_x(.) = \langle y_k,.\rangle x$, opérateur de rang 1 : on a $ap = \Sigma_k\, h^k_{a(y_k)}$, donc $f(a) = \Sigma_k\, f(h^k_{a(y_k)})$. Mais $x \longmapsto f(h^k_x)$ est une forme linéaire continue sur H, donc de la forme $\langle z_k,.\rangle$, et il reste $f(a)=\Sigma_k \langle z_k, a(y_k)\rangle$, le résultat désiré. □

La troisième topologie faible utile sur $\mathfrak{L}^\infty$ est la topologie $\sigma(\mathfrak{L}^\infty,\mathfrak{L}^1)$ associée à la dualité avec les opérateurs à trace. Elle porte en théorie des algèbres de von Neumann le nom malheureux de topologie ultrafaible (on utilise plutôt σ-faible dans les livres récents) . Remarquons que tout opérateur à trace s'écrit comme produit bc de deux HS, et que $tr(bca)$ s'écrit dans une base o.n. (e_n)

$$tr(bca) = tr(cab) = \Sigma_n \langle c^* e_n, ab(e_n)\rangle = \Sigma_n \langle x_n, ay_n\rangle$$

avec $\Sigma_n \|x_n\|^2 < \infty$, $\Sigma_n \|y_n\|^2 < \infty$: il est facile d'en déduire que la topologie "σ-faible" est définie par les semi-normes fondamentales[1]

$$r_{(x_n),(y_n)}(a) = \Sigma_n\, |\langle x_n, ay_n\rangle| \quad (\ \Sigma_n \|x_n\|^2 < \infty\ ,\ \Sigma_n \|y_n\|^2 < \infty\ )$$

Mais cela n'est pas important. Ce qui compte, c'est que

- Une forme linéaire continue pour cette topologie provient d'un opérateur à trace.
- Sur les boules fermées de $\mathfrak{L}(H)$, cette topologie est compacte, donc coïncide avec la topologie faible (séparée et moins forte), donc est aussi métrisable.

1. Le sens de σ dans σ-faible est un peu le même que dans "σ-additive".

ELEMENTS DE PROBABILITES QUANTIQUES. II

Quelques exemples discrets

Cet exposé contient surtout une série d'exemples : celui du spin et de quelques systèmes de spins . Cette série se poursuit dans l'exposé III, l'exemple du couple canonique ayant été détaché en raison de son importance (et de sa longueur).

I. LE "SPIN"

1. En probabilités classiques, l'espace mesurable non trivial le plus simple a deux points, et une v.a. à valeurs dans cet espace décrit une question à laquelle on ne peut répondre que par oui ou non. En formant des produits de tels espaces élémentaires, on construit des espaces arbitrairement compliqués (ainsi, l'espace $\{0,1\}^{\mathbb{N}}$ est non dénombrable, et porte des lois diffuses).

L'analogue quantique de cet espace mesurable a fait son apparition en mécanique à propos du spin, d'où le titre de ce paragraphe. Mais il s'agit en fait d'une question de pure probabilité , et notre lecteur n'apprendra pas ce qu'est le spin ! Nous nous bornerons à indiquer la traduction du langage probabiliste au langage des physiciens.

Considérons l'espace de Hilbert $H = L^2_{\mathbb{C}}(\mu)$, où μ est la loi de probabilité sur l'ensemble $\{-1,1\}$ qui attribue à chacun de ces deux points la masse 1/2. Ce sera pour nous le modèle naturel d'un espace de Hilbert de dimension (complexe) 2 . L'interprétation probabiliste fournit des bases orthonormales distinguées : soit ν la fonction qui prend les valeurs

$$\nu(-1)=0 \ , \quad \nu(1)=1$$

nous pouvons prendre comme base orthonormale, soit

$$|+> = \sqrt{2}\,\nu \qquad |-> = \sqrt{2}\,(1-\nu)$$

soit

$$1 \ , \quad X = 2\nu-1 \quad (\text{ application identique de } \{-1,1\} \text{ dans } \mathbb{R})$$

Dans ce paragraphe, nous nous servirons plus de la première, et plus de la seconde lorsque nous traiterons des systèmes plus compliqués.

Tout vecteur de H s'écrit $\omega = u|+> + v|->$; supposons ω normalisé. On a $|u|^2+|v|^2=1$, de sorte qu'il existe un angle unique θ entre 0 et π tel que

$$|u| = \cos(\theta/2) \ , \quad |v| = \sin(\theta/2) \ .$$

Nous pouvons alors écrire $u=|u|e^{ia}$, $v=|v|e^{ib}$. Comme la loi ε_ω ne change pas si l'on remplace ω par $c\omega$ avec $|c|=1$, nous pouvons normaliser cette représentation en supposant $b=-a$. Nous atteignons donc toutes les lois pures en nous limitant à des vecteurs normés de la forme

$$(1) \qquad \omega_{\theta,\phi} = \cos(\theta/2)e^{-i\phi/2}|+> + \sin(\theta/2)e^{i\phi/2}|-> $$

avec $0\leqq\theta\leqq\pi$, $0\leqq\phi\leqq 2\pi$ par exemple. En faisant un peu attention aux cas $\theta=0,\pi$ où cette représentation est singulière, on voit que les lois pures sont paramétrées par les points de la sphère S_2 , θ étant la latitude (mesurée à partir du pôle Nord) et ϕ la longitude (à partir de Ox).

Puisque notre espace H est un $L^2(\mu)$, où μ est une mesure sur $\mathbb{R}$, il porte une v.a. quantique (observable) naturelle, qui est l'opérateur de multiplication par l'application identique . Comme μ est portée par l'ensemble $\{-1,1\}$, les valeurs de cette observable sont $-1,1$. Nous la désignerons conventionnellement par σ_z, et son interprétation physique est celle d'une mesure du spin suivant Oz (ce spin pouvant être trouvé soit "vers le haut", soit "vers le bas"). La matrice de cette observable, dans la base $|+>$, $|->$, est $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Plus généralement, nous pouvons écrire toutes les observables (opérateurs a.a.) dans la base $|\pm>$, sous la forme suivante, où les coefficients x,y,z,t sont réels

$$(2) \qquad A = \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} = tI + x\sigma_x + y\sigma_y + z\sigma_z$$

où les matrices σ sont les matrices de Pauli

$$(3) \qquad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

Si l'on donne à t,x,y,z des valeurs complexes, on écrit ainsi tous les opérateurs sur H.

La loi de la v.a. σ_z dans l'état pur associé à $\omega_{\theta,\phi}$ correspond, en probabilités classiques, au remplacement de μ par $|\omega_{\theta,\phi}|^2\mu$; ainsi

$$P_{\theta,\phi}\{\sigma_z=+1\} = \cos^2(\theta/2) \quad , \quad P_{\theta,\phi}\{\sigma_z=-1\} = \sin^2(\theta/2) .$$

Une loi de probabilité (mélange) sur H est une matrice du type (2), mais de plus positive et de trace 1, ce qui nous donne $t=1/2$. Il est donc plus commode de l'écrire sous la forme

$$(4) \qquad W = \frac{1}{2}\begin{pmatrix} 1+z & \bar{r} \\ r & 1-z \end{pmatrix} \quad \text{avec} \quad r=x+iy$$

La positivité s'écrit $<\begin{pmatrix} u \\ v \end{pmatrix}, W\begin{pmatrix} u \\ v \end{pmatrix}> \geqq 0$, soit $r\bar{r}\leqq 1-z^2$, et enfin $x^2+y^2+z^2\leqq 1$. Les lois sont donc paramétrées par les points de la boule unité de $\mathbb{R}^3$. Cherchons à retrouver ainsi le paramétrage des points extrémaux (lois pures) par la sphère unité : la matrice du projecteur sur le vecteur de norme 1

$\omega = u|+> + v|->$ étant $\begin{pmatrix} \bar{u}u & \bar{v}u \\ \bar{u}v & \bar{v}v \end{pmatrix}$, si u,v sont donnés par (1) on a

(5) $\quad W_{\theta,\phi} = \frac{1}{2}(I+\sigma_{\theta,\phi})$, $\sigma_{\theta,\phi} = \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}$

(6) $\quad x = \sin\theta\cos\phi$, $y = \sin\theta\sin\phi$, $z = \cos\phi$ dans (4)

Les matrices (5) sont aussi celles de <u>tous les projecteurs orthogonaux</u> sur des sous-espaces de dimension 1 (i.e., autres que 0 et I) : $W_{\theta,\phi}$ admet les vecteurs propres $\omega_{\theta,\phi}$ et $\omega_{\pi+\theta,\phi}$ (orthogonaux dans $\mathbb{C}^2$, opposés dans $\mathbb{R}^3$) avec les valeurs propres 1,0 respectivement.

Dans la situation physique du spin (pour une particule de spin 1/2), $\sigma_{\theta,\phi}$ représente une mesure de spin dans la direction du vecteur (6) . Sous la loi $\varepsilon_{0,\phi}$ (ϕ arbitraire), pour laquelle une mesure du spin suivant Oz donnerait toujours la réponse + , donc pour laquelle, du point de vue classique, il ne resterait <u>plus rien d'aléatoire</u>, l'observable $\sigma_{\theta,\phi}$ fournit une réponse + avec probabilité $\cos^2(\theta/2)$, une réponse - avec probabilité $\sin^2(\theta/2)$, d'où une moyenne de $\cos\theta$. Plus généralement, si l'on a filtré les particules de manière que leur spin soit toujours dans la direction du vecteur unitaire a, une mesure du spin dans la direction du vecteur unitaire b donnera la valeur moyenne a.b (plus exactement, $\frac{\hbar}{2}$a.b), résultat comparable à ceux que l'on obtient en optique pour la lumière polarisée, mais cette valeur moyenne provenant de l'accumulation de "tirages au sort" individuels fournissant chacun une réponse discrète $\pm \frac{\hbar}{2}$.

Nous regroupons ici quelques formules concernant les matrices de Pauli, qui seront utiles plus tard

(7) $\quad \sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z$ (deux autres formules par permutation circulaire)
$\quad\quad \sigma_x\sigma_x = \sigma_y\sigma_y = \sigma_z\sigma_z = I$.

Les matrices $\mathbf{i}=-i\sigma_x$, $\mathbf{j}=-i\sigma_y$, $\mathbf{k}=-i\sigma_z$ sont les matrices traditionnellement utilisées pour représenter les quaternions : on a $\mathbf{ij} = \mathbf{k} = -\mathbf{ji}$ (et deux autres formules par permutation circulaire), $\mathbf{i}^2=\mathbf{j}^2=\mathbf{k}^2= -1$.

2. Il est très facile de donner des exemples de <u>groupes unitaires</u> représentant des évolutions hamiltoniennes. Avec les notations des physiciens

$$U_t = e^{itH/\hbar}$$

où H est un opérateur a.a., nécessairement borné ici, qui représente l'observable énergie. Excluant le cas où H est proportionnel à I, H admet deux vecteurs propres orthogonaux et normés, avec deux valeurs propres qui représentent deux "niveaux d'énergie" . Comme remplacer H par H+cI avec c réel ne change $e^{itH/\hbar}$ que par un facteur de module 1, et l'énergie n'est jamais définie en physique qu'à une constante additive près, on peut normaliser la situation en supposant que les deux niveaux d'énergie sont -E et

+E avec E>0. On ne perd aucune généralité essentielle en supposant que les deux vecteurs propres sont respectivement $|- >$ et $|+ >$ (cela revient à un changement de base). Alors on a $H=E\sigma_z$ et

$$(8) \qquad e^{itH/\hbar} = \begin{pmatrix} e^{itE/\hbar} & 0 \\ 0 & e^{-itE/\hbar} \end{pmatrix}$$

Si l'état initial était $\omega_{\theta,\phi}$, l'état à l'instant t $U_t\omega_{\theta,\phi}$ est décrit par

$$\theta_t=\theta \quad , \quad \phi_t = \phi - tE/\hbar$$

Comme θ reste constant, on voit que dans une telle évolution, la probabilité des réponses +,- pour l'observable σ_z reste inchangée. En revanche, les lois des observables σ_u varient périodiquement au cours du temps. La signification physique de (8) est que le vecteur représentant le spin tourne autour de Oz d'un mouvement uniforme (précession de Larmor).

3. Avant de donner des exemples d'évolutions irréversibles, nous introduisons certaines matrices remarquables, qui reparaîtront sous des formes diverses dans tous les exposés ultérieurs.

Nous nous plaçons dans la base $|- >$, $|+ >$: l'état $|- >$ sera appelé <u>état vide</u>, et nous le noterons ϕ , et l'état $|+ >$ <u>état occupé</u> (nous sommes donc en train de changer d'interprétation physique : il ne s'agit plus de mesurer un spin !). Nous appelons <u>opérateur de création</u> b^+ et <u>opérateur d'annihilation</u> b^- les opérateurs de matrices

$$(9) \qquad \begin{aligned} b^+ &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{transforme l'état vide en état occupé,} \\ b^- &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{transforme l'état occupé en état vide.} \end{aligned}$$

Ces opérateurs sont adjoints l'un de l'autre. On appelle opérateur <u>nombre de particules</u> l'opérateur a.a.

$$(10) \qquad N = b^+b^- = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$ qui vaut 0 dans l'état vide, 1 dans l'état occupé.

[En physique, les << particules >> en question sont des individus insociables appelés <u>fermions</u> : dès qu'un fermion a occupé une place, il empêche tout autre fermion d'approcher, c'est pourquoi un nombre de fermions ne peut être que 0 ou 1 dans une place donnée]. On remarquera que

$$(11) \qquad b^+ + b^- = \sigma_x \quad , \quad i(b^+ - b^-) = \sigma_y \quad , \quad [b^-,b^+] = \sigma_z = I-2N$$

D'autre part, en calculant des <u>anticommutateurs</u> ($\{A,B\}=AB+BA$)

$$(12) \qquad \{b^-,b^-\}=\{b^+,b^+\}=0 \quad , \quad \{b^-,b^+\}=I$$

qui est une forme élémentaire des <u>relations d'anticommutation canoniques</u>. La situation présente est trop élémentaire, et le langage se comprendra mieux lorsque nous aurons vu d'autres exemples.

On notera que, lorsque l'on identifie l'espace de Hilbert H à $L^2(\mu)$ comme au début, l'opérateur de multiplication par X est égal à σ_x .

4. Sur notre espace de Hilbert de dimension 2, l'espace des mesures bornées réelles coïncide avec celui des opérateurs a.a., puisque tout opérateur a.a. a une trace finie. On voit sur la représentation (2) qu'il est de dimension réelle 4, et non 2 comme en probabilités classiques (il y a beaucoup plus d'"événements"). Un semi-groupe de noyaux markoviens sur $\mathfrak{M}$ est donné par une équation

$$(13) \qquad W_t = e^{tG}(W_0) \quad \text{ou} \quad \dot{W}_t = G(W_t)$$

où le générateur G est une application linéaire de $\mathfrak{M}$ dans $\mathfrak{M}$, et l'opérateur linéaire e^{tG} préserve la positivité et la trace. On a écrit $e^{tG}(.)$, $G(.)$ pour éviter la confusion possible avec un produit de matrices (2,2), alors qu'ici G serait plutôt une matrice (4,4). Nous allons rechercher la forme des générateurs possibles, en recopiant le livre de Davies.

Remarquons d'abord que l'ensemble des générateurs de semi-groupes de noyaux markoviens est un cône convexe. En effet, remplacer G par λG ($\lambda>0$) revient à changer de temps sur le semi-groupe, tandis que si G et G' sont les générateurs de deux semi-groupes markoviens P_t et P'_t, le semi-groupe d'opérateurs $e^{t(G+G')}$ est donné par la formule classique de Trotter-Kato (facile à justifier en dimension finie)

$$e^{t(G+G')} = \lim_n (P_{t/n}P'_{t/n})^n$$

et il est clair qu'il préserve la positivité et la trace.

- Un premier type de générateurs est donné par les <u>évolutions hamiltoniennes</u> : on considère un groupe unitaire $U_t = e^{itH}$, que l'on fait agir sur les mesures W par

$$(14) \qquad P_t(W) = W_t = U_t W U_t^*$$

Nous avons alors $\quad G(W) = i[H,W]$

Cette évolution préserve les lois pures : si $W=\varepsilon_\omega$, on a $W_t=\varepsilon_{U_t\omega}$. Du point de vue physique, c'est un comportement réversible, analogue aux flots déterministes de la mécanique classique.

- Voici un second type de générateurs . Nous nous donnons un opérateur K arbitraire, et nous posons

$$(15) \qquad G(W) = 2KWK^* - ((K^*K)W+W(K^*K)) = [KW,K^*]+[K,WK^*] \ .$$

Posons $Q_t = e^{tG}$. Il est clair que si W est a.a., $G(W)$ est a.a., donc $Q_t(W)$ également. Ensuite, on a $\mathrm{Tr}(G(W))=0$ pour tout W, d'où il résulte que Q_t préserve la trace. Reste à vérifier que Q_t préserve la positivité. Pour cela, on écrit G sous la forme $G_1 + G_2$, où

$$G_1(W) = KWK^* \ , \quad G_2(W) = -(JW+WJ) \text{ avec } J=K^*K$$

et l'on vérifie que e^{tG_1} et e^{tG_2} préservent la positivité. Pour le

premier, on écrit

$$e^{tG_1}(W) = W + tKWK^* + \frac{t^2}{2!}K^2WK^{*2} + \ldots$$

qui est une somme d'opérateurs positifs si $W \geq 0$ et $t \geq 0$ (on a un semi-groupe et non un groupe). Pour G_2 , on a

$$e^{tG_2}(W) = e^{-tJ}We^{-tJ}$$

qui est positif si W est positif, J étant a.a.. Tout ceci vaut sur un espace de dimension quelconque, et l'on peut démontrer de plus que les noyaux ainsi construits sont complètement positifs.

5. Voici des exemples plus concrets, sur l'espace à deux points. Cette fois ci, nous appellerons l'état $|- >$ état fondamental, $|+ >$ état excité, ce qui constitue une troisième interprétation physique. Nous allons donner un modèle d'évolution irréversible pour un système à deux états, que l'on appelle atome de Wigner-Weisskopf. Nous prendrons pour G un générateur

$$G = h + c_-g_- + c_+g_+ \qquad (16)$$

où h est un générateur du premier type (hamiltonien)

$$h(W) = i[H,W] \quad \text{avec} \quad H = \begin{pmatrix} \omega_- & 0 \\ 0 & \omega_+ \end{pmatrix}$$

et c_-, c_+ sont des constantes positives, g_-, g_+ des générateurs du second type (irréversibles)

$$g_-(W) = 2b^-Wb^+ - b^+b^-W - Wb^+b^-$$

$$g_+(W) = 2b^+Wb^- - b^-b^+W - Wb^-b^+$$

Appelons lois diagonales les lois de la forme $W = \begin{pmatrix} \lambda_- & 0 \\ 0 & \lambda_+ \end{pmatrix}$. On peut montrer que l'évolution (16) préserve les lois diagonales, les coefficients $\lambda_-(t)$, $\lambda_+(t)$ à l'instant t satisfaisant à l'équation différentielle suivante[1] (les coefficients $\omega_\pm$ n'interviennent pas)

$$\dot{\lambda}_- = -2c_+\lambda_- + 2c_-\lambda_+$$

$$\dot{\lambda}_+ = 2c_+\lambda_- - 2c_-\lambda_+$$

En particulier, si $c_+ = 0$ et $\lambda_-(0) = 0$, $\lambda_+(0) = 1$, on a $\lambda_+(t) = e^{-2ct}$, $\lambda_-(t) = 1 - e^{-2ct}$. L'évolution des probabilités décrit l'idée intuitive d'un état excité instable, tendant à revenir à un état fondamental stable (évolution d'un atome radioactif). Il s'agit sans doute du modèle le plus simple possible d'évolution non hamiltonienne (irréversible) que l'on rencontre dans la nature.

Pour toutes ces questions, voir le livre de Davies Quantum Theory of Open Systems, Academic Press.

1. La même que pour les chaînes de Markov à deux états.

II. DEUX EVENEMENTS

La lecture de ce paragraphe n'est pas indispensable pour la suite, mais elle constitue, il me semble, un excellent exercice, instructif à bien des égards. Il s'agit d'étudier " la tribu engendrée par deux événements ". Le contenu du paragraphe est emprunté à M.A. Rieffel et A. van Daele, A bounded operator approach to Tomita-Takesaki theory, Pacific J.M. 69, 1977, et à travers eux, en partie, à Halmos, Two subspaces, Trans. AMS 144, 1969, et même Dixmier, Rev. Sci. 86, 1948. Les principales applications des résultats présentés ici ne concernent pas les probabilités quantiques, mais la théorie des algèbres de von Neumann (quelques indications dans le dernier n° du paragraphe).

1. Soit Ω un espace de Hilbert complexe, et soient A,B deux événements (sous-espaces fermés). Les quatre sous-espaces $A\cap B$, $A\cap B^\perp$, $A^\perp\cap B$, $A^\perp\cap B^\perp$ sont stables par les projecteurs I_A et I_B , de même que leur somme K . Sur K les projecteurs I_A et I_B commutent, et les v.a. I_A et I_B ont donc une loi jointe sous toute loi ε_ω ($\omega\in K$) et tout mélange de telles lois. Tout se passe donc comme en probabilités classiques. Les phénomènes proprement quantiques se produisent sur le sous-espace $H=K^\perp$ (également stable par I_A et I_B). Nous nous restreindrons donc à H, autrement dit, nous supposons désormais que les sous-espaces sont "en position générale"

$$(1)\qquad A\cap B = A^\perp\cap B^\perp = \{0\} = A\cap B^\perp = A^\perp\cap B \ .$$

Nous désignons par p,q les projecteurs sur A et B . Nous posons alors

$$(2)\qquad c = p-q \ , \qquad s = p+q-I$$

Ce sont des opérateurs a.a. satisfaisant aux relations

$$(3)\qquad c^2+s^2=I \ , \quad cs+sc = 0 \quad (cs = [p,q]) \ .$$

Les notations c et s sont là pour rappeler la trigonométrie. Il est clair que les opérateurs I+c, I-c, I+s, I-s sont positifs, d'où les représentations spectrales

$$(4)\qquad c = \int_{-1}^{1} \lambda dE_\lambda \ , \qquad s = \int_{-1}^{1} \lambda dF_\lambda \ .$$

Tirons quelques conclusions de (1). La relation cx=0 entraîne px=qx, donc px=qx=0, donc $x\in A^\perp\cap B^\perp$, et enfin x=0 : c est injectif . Un raisonnement tout analogue montre que I-s, I+s sont injectifs. En utilisant la seconde moitié des relations (1) on verra que s, I-c, I+c sont injectifs. L'injectivité de c signifie que $\int I_{\{0\}}(s)dE_s=0$, et de même pour s . Dans la situation étudiée par Rieffel et van Daele, on suppose seulement la moitié de gauche de (1), de sorte que s, I-c, I+c ne sont pas nécessairement injectifs.

Posons

(5) $$j = \int_{-1}^{1} \mathrm{sgn}(\lambda)\,dE_\lambda \quad , \quad d = \int_{-1}^{1} |\lambda|\,dE_\lambda$$

Comme c est injectif, j est un opérateur a.a. de carré I : une symétrie. D'autre part, d est a.a. positif. D'après (4) , on a $d^2=I-s^2$: comme d est positif, c'est la racine carrée de $1-s^2$, ainsi $d=|c|=(1-s^2)^{1/2}$, et d commute avec c,s,p,q . Enfin $djp=cp=(p-q)p=(I-q)(p-q)=(I-q)dj=d(I-q)j$. Comme d est injectif, cela entraîne

(6) $$jp=(I-q)j$$, d'où en passant à l'adjoint $jq=(I-p)j$

et par addition

(7) $$js = -sj \quad , \quad jc = cj$$

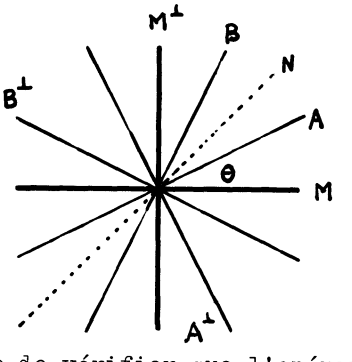
(la seconde relation est simplement rappelée : elle résulte de (5)). Nous dirons que j est la symétrie principale (principale, parce que son existence n'exige que la première moitié de (1))

Supposant la seconde moitié de (1), on construit de même la symétrie secondaire k et l'opérateur a.a. positif e , tels que

(8) $$s = ke = ek \quad , \quad e = |s| = (1-c^2)^{1/2}$$, e commute à c,s,p,q ,

(9) $$kc = -ck \quad , \quad ks=sk$$.

Les deux symétries j,k anticommutent. En effet, k anticommute à c , donc à tout polynôme impair en c , donc à la fonction impaire $j=\mathrm{sgn}(c)$. Nous faisons un dessin, en désignant par M le sous-espace $\{jx=x\}$, par N le sous-espace $\{kx=x\}$.



Il n'est pas difficile de vérifier que l'opérateur $kj = i_o$ est tel que $i_o^2=-I$, est une bijection de M sur $M^\perp$. Si l'on représente H comme $M \oplus M^\perp$ et que l'on identifie $M^\perp$ à M par i_o , tout opérateur borné sur H se représente comme une matrice (2,2) d'opérateurs sur M. En particulier

$$j = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \quad , \quad k = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \quad , \quad i_o = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$$ (I identité de M)

Si nous posons $\ell=-ijk$ (où i est l'opérateur de multiplication par ie^{ϕ}) la matrice de ℓ est la troisième matrice de Pauli, et il n'est pas difficile alors de vérifier que l'ensemble des opérateurs de la forme

$$a = u + vj + wk + z\ell$$

où u,v,w,z sont des fonctions boréliennes bornées de l'opérateur d (ou e) qui commute avec tous les autres, est exactement l'algèbre de v.N. engendrée par p et q, c'est à dire l'analogue quantique de la tribu engendrée par les deux événements. Cela illustre deux faits : la "tribu" engendrée par deux événements contient, dès que ceux-ci ne commutent pas, une infinité non dénombrable de projecteurs. D'autre part, le caractère fondamental du spin comme phénomène quantique élémentaire.

2. Les résultats de cette section ne nous serviront pas, mais il est difficile de quitter ce sujet sans signaler combien il est intéressant. Le lecteur pourra se reporter à l'article cité de Rieffel-Van Daele.

a) Tout ce que nous avons fait plus haut peut se faire pour un espace de Hilbert réel (sauf l'extrême fin où nous avons posé $\ell=ijk$). La figure de la page précédente suggère de considérer A et B, dans le "système d'axes" $M,M^{\perp}$, comme définis par des "équations" $y=(tg\Theta)x$, $y=(cotg\Theta)x$, où Θ est l'"angle" entre M et A. Cette trigonométrie a, semble t'il, des applications en statistique . Voici (d'après Halmos puis R-vD) le sens qu'on peut lui attribuer. Les opérateurs d et e commutent, laissent stable M : désignons par $\underline{d}$, $\underline{e}$ leurs restrictions à M , et convenons de poser

$$(\cos 2\Theta)=\underline{d} \ , \ (\sin 2\Theta)=\underline{e} \ ; \ (\cos^2\Theta)= \tfrac{1}{2}(I+\underline{d}), \ (\sin^2\Theta)= \tfrac{1}{2}(I-\underline{d}) \ ;$$
$$(tg\Theta)= \left(\frac{I-\underline{d}}{I+\underline{d}}\right)^{1/2} \ , \ (cotg\Theta) = \left(\frac{I+\underline{d}}{I-\underline{d}}\right)^{1/2}$$

tous opérateurs bornés sur M, sauf le dernier. Alors on peut montrer que A est le graphe de $(tg\Theta)$, B celui de $(cotg\Theta)$, lorsqu'on identifie $M^{\perp}$ à M par i_o, et donc H à M×M .

b) Dans la situation intéressante pour la théorie des algèbres de vN, H est un espace de Hilbert complexe, que l'on munit aussi de la structure réelle $(x,y)=Re\langle x,y\rangle$, et A et B sont deux sous-espaces réels, satisfaisant à la première moitié des relations (1) (mais on n'aura pas $A\cap B^{\perp}=\{0\}$), et liés par la relation $B=iA$ (ou encore $ip=qi$). L'opérateur s est alors $\mathbb{C}$-linéaire, mais c,j sont antilinéaires. Un élément de structure très important est le groupe unitaire U_t (noté Δ^{it})

$$U_t = (I-s)^{it}(I+s)^{-it}$$

qui commute avec j et laisse toute la figure invariante. Peut être aurons nous l'occasion d'y revenir.

III. SYSTEMES FINIS DE SPINS

Ce paragraphe est le développement logique des deux précédents : l'étude d'un spin était celle de l'espace mesurable quantique $L^2(\mu)$, où μ est une loi de Bernoulli, et nous allons considérer maintenant un système fini de v.a. de Bernoulli indépendantes.

Ce paragraphe est très important, car les analogies avec ce que nous ferons plus tard en temps continu sont étonnamment étroites. Néanmoins, le lecteur fera peut être bien de parcourir l'exposé III avant de revenir ici.

Un court appendice a permis au rédacteur d'apprendre un peu d'algèbre. Il n'est pas indispensable pour la lecture des exposés ultérieurs.

1. Donnons nous un entier M, et une famille finie ν_i de v.a. indépendantes ($i=1,\dots,M$), prenant la valeur 1 avec probabilité p, la valeur 0 avec probabilité $q=1-p$. Nous pouvons réaliser tout cela sur $E=\{0,1\}^M$, espace à 2^M points, et nous désignons par μ_p la loi de probabilité Nous posons

(1) $$X_i = \frac{\nu_i - p}{\sqrt{pq}}$$

v.a. de moyenne nulle et de variance unité. Les v.a. suivantes, où A parcourt l'ensemble des parties de $\{1,\dots,M\}$

(2) $$X_A = \prod_{i\in A} X_i \qquad (X_\emptyset = 1)$$

constituent une base orthonormale de l'espace $\mathcal{Y} = L^2(\mu_p)$. Nous désignerons par $\mathcal{Y}_n$ le sous-espace de $\mathcal{Y}$ engendré par les X_A, $|A|=n$.

Une v.a. Y s'écrit de manière unique sous la forme $\Sigma_A\, y_A X_A$ ($y_A \in \mathbb{C}$), et l'on a $< Y,Z > = \Sigma_A\, \bar{y}_A z_A$; on sait aussi calculer l'espérance de Y, égale à $y_\emptyset = < 1,Y >$. Quant à la multiplication, elle est complètement décrite par l'associativité, la commutativité, le fait que $X_\emptyset = 1$ est élément unité, et la règle (qui découle de (1))

(3) $$X_i^2 = 1 + cX_i \quad , \text{ avec } \quad c = \frac{1-2p}{\sqrt{pq}} \quad .$$

En particulier, lorsque $p=1/2$, $c=0$, on a la formule explicite

(4) $$X_A X_B = X_{A\Delta B} \quad (\text{différence symétrique}) \ .$$

Nous allons maintenant prendre une formulation purement algébrique de notre problème : considérons un espace de Hilbert complexe $\mathcal{Y}$ de dimension 2^M, avec une base orthonormale réelle notée X_A ($A\subset\{1,\dots,M$; dire que la base est réelle revient à dire que $\mathcal{Y}$ est muni d'une conjugaison, etc.). Pour chaque valeur de p, nous avons mis en évidence une structure possible d'algèbre associative et commutative sur $\mathcal{Y}$, admettant $X_\emptyset$ comme unité (nous écrirons toujours $X_\emptyset = 1$), et caractérisée par (3).

Le travail que nous allons faire, et qui est analogue en temps discret à celui que nous ferons plus tard en temps continu, consiste à décrire de manière purement algébrique ces divers opérateurs de multiplication sur $\mathcal{V}$, à partir des opérateurs de création et d'annihilation de type symétrique. Ensuite, nous définirons d'autres opérateurs, de type antisymétrique, qui nous permettront de définir d'autres structures d'algèbre sur $\mathcal{V}$, celles-ci non commutatives, et d'une extrême importance en physique et en mathématiques (algèbres de Clifford ; spineurs).

2. Nous commençons par les opérateurs du type symétrique : ils sont définis par

$$(5)\qquad \begin{aligned} a_k^+(X_A) &= X_{A\cup\{k\}} \quad \text{si } k\notin A\ ,\ 0 \ \text{ sinon} \qquad (\text{ création }) \\ a_k^-(X_A) &= X_{A\setminus\{k\}} \quad \text{si } k\in A\ ,\ 0 \ \text{ sinon} \qquad (\text{ annihilation }) \end{aligned}$$

Il est très facile de vérifier que $< X_B, a_k^- X_A> = < a_k^+ X_B, X_A >$, donc ils sont adjoints l'un de l'autre. L'opérateur

$$(6)\qquad N_k = a_k^+ a_k^- \ :\ N_k X_A = X_A \text{ si } k\in A\ ,\ 0 \text{ sinon}$$

est donc a.a.. Si l'on calcule $a_k^- a_k^+$, on trouve $I-N_k$, donc

$$(6)\qquad \begin{aligned} &[a_k^-, a_k^+] = I-2N_k\ , \quad [a_k^-, a_j^+] = 0 \text{ si } k\neq j \\ &[a_k^-, a_j^-] = [a_k^+, a_j^+] = 0 \quad \text{quels que soient } j,k \end{aligned}$$

En fait, on a $a_k^- a_k^+ + a_k^+ a_k^- = I$, et en regardant bien, on s'aperçoit que chaque couple (a_k^-, a_k^+) est une copie du couple (b^-, b^+) du § I, (9), tous ces couples commutant entre eux. Autrement dit, nous considérons un système de spins commutant entre eux. Les v.a. de spin sont (§ I, (11))

$$(7)\qquad q_k = a_k^+ + a_k^-\ ,\ p_k = i(a_k^+ - a_k^-)\ .$$

Les q_k commutent tous entre eux, et peuvent donc être considérés comme des v.a. classiques. Si l'on regarde comment q_k opère, on voit que

$$q_k X_A = X_{A\Delta\{k\}}$$

Comparant cela à (4), on voit que q_k est l'opérateur de multiplication par X_k correspondant à $p=1/2$. Il est facile d'expliciter l'opérateur de multiplication par X_k correspondant à $p\neq 1/2$, $c\neq 0$: en regardant comment cette multiplication opère sur X_A pour $k\in A$ et $k\notin A$, on trouve que

$$(8)\qquad X_k^{p,c} X_A = (q_k + cN_k) X_A\ .$$

Il est clair que sous la loi ε_1, les v.a. q_k+cN_k sont indépendantes, équidistribuées, avec la même distribution que $(\nu_i - p)/\sqrt{pq}$ en (1) : cela résulte de l'équivalence unitaire entre $\mathcal{V}$ muni des opérateurs q_k+cN_k et l'espace $L^2(\mu_p)$ muni des opérateurs de multiplication par X_k.

Quelle est dans ces conditions la loi des v.a. p_k, p_k+cN_k ? Désignons par $\mathfrak{F}$ (pour rappeler la transformation de Fourier) l'opérateur unitaire sur $\mathcal{Y}$ défini par

(9) $\qquad \mathfrak{F}X_A = i^{|A|}X_A$

opérateur qui préserve $X_\emptyset=1$, et qui satisfait à

(10) $\qquad \mathfrak{F}^{-1}a_k^+\mathfrak{F} = ia_k^- \ , \ \mathfrak{F}^{-1}a_k^-\mathfrak{F} = a_k^+ \ , \ \mathfrak{F}^{-1}N_k\mathfrak{F} = N_k$

et donc à $\mathfrak{F}^{-1}(q_k+cN_k)\mathfrak{F} = p_k+cN_k$, $\mathfrak{F}^{-1}(p_k+cN_k)\mathfrak{F} = -q_k+cN_k$. On voit (comme $\mathfrak{F}$ préserve 1) que les v.a. p_k+cN_k et q_k+cN_k ont même loi sous ε_1 .

Remarquer que sous la loi ε_1 , on a p.s. $N_k=0$, et que cependant q_k et q_k+cN_k n'ont pas la même loi.

Il est facile d'écrire au moyen des opérateurs de création et d'annihilation tous les opérateurs sur $\mathcal{Y}$. Remarquons en effet que tout vecteur $x\in\mathcal{Y}$ non nul peut être ramené à un multiple non nul de 1 par un produit d'annihilateurs convenable, puis que ce vecteur peut être transformé en un vecteur quelconque y par une somme de produits de créateurs convenables. Autrement dit, l'algèbre engendrée par les a_j^- et les a_k^+ opère sur $\mathcal{Y}$ de manière irréductible, et d'après un théorème (facile) d'algèbre, c'est alors l'algèble de tous les opérateurs sur $\mathcal{Y}$. Au moyen des relations

$$a_k^-a_k^+ = I - a_k^+a_k^- \qquad a_k^{+2}=a_k^{-2}=0 \ ,$$

et de la commutation d'opérateurs d'indices j,k différents, on peut exprimer tout produit de créateurs et d'annihilateurs comme une combinaison linéaire de produits $a_A^+a_B^-$ où tous les créateurs sont à gauche des annihilateurs, et où A et B sont deux parties quelconques de $\{1,\dots,M\}$. Comme le nombre de ces opérateurs est 2^{2M} , i.e. juste ce qu'il faut pour engendrer toutes les matrices $(2^M,2^M)$, ils forment une base de l'ensemble de tous les opérateurs sur $\mathcal{Y}$. Autre écriture (unique) : $a_A^+N_Ba_C^-$ avec A,B,C disjoints.

Nous interrompons provisoirement l'étude de ces opérateurs. Nous verrons plus tard qu'ils constituent une approximation discrète des opérateurs de création et d'annihilation de l'espace de Fock symétrique.

2. Nous passons aux opérateurs de création et d'annihilation antisymétriques, décrivant des systèmes de spins qui anticommutent, au lieu de commuter. Si k est un indice, B une partie de $\{1,\dots,M\}$, nous désignons par n(k,B) (resp. n'(k,B)) le nombre des éléments de B strictement inférieurs (resp. supérieurs) à k . Il est clair que $n(k,B)+n'(k,B) = |B|-I_B(j)$. On pose ensuite pour toute partie A $n(A,B) = \Sigma_{k\in A}\, n(k,B)$, et de même n'(A,B) : n(A,B) est le nombre d'inversions observé lorsqu'on écrit B à droite de A (A,B étant rangés par ordre croissant), et n'(A,B)=n(B,A). On a

$$n(A,B) + n(B,A) = |A||B| - |A\cap B| \ .$$

On pose

$$(11)\qquad b_k^+(X_A) = (-1)^{n(k,A)} X_{A\cup\{k\}} \quad \text{si } k\notin A \text{ , 0 sinon}$$
$$b_k^-(X_A) = (-1)^{n(k,A)} X_{A\setminus\{k\}} \quad \text{si } k\in A \text{ , 0 sinon}$$

Ces opérateurs sont adjoints l'un de l'autre. Contrairement aux opérateurs (5), qui ne sont qu'une approximation des opérateurs de création et d'annihilation de l'espace de Fock symétrique, ce sont exactement les opérateurs de création et d'annihilation de l'espace de Fock antisymétrique, construit sur un espace de Hilbert de dimension M, ici $H=\mathcal{Y}_1$.

Comme dans la section précédente, chaque couple (b_k^+,b_k^-) représente un spin - mais ces spins, au lieu de commuter, vont anticommuter. On vérifie en effet les relations d'anticommutation canoniques

$$(12)\qquad \{b_j^-,b_k^-\} = \{b_j^+,b_k^+\} = 0 \ , \quad \{b_j^-,b_k^+\} = \delta_{jk} I \ .$$

Notons aussi que $b_k^+ b_k^- = N_k$ (le même que tout à l'heure) et que la transformation de Fourier $\mathcal{F}$ opère sur les $b_k^\pm$ exactement comme (10).

Comme plus haut aussi, nous construisons des opérateurs autoadjoints

$$(13)\qquad r_k = b_k^+ + b_k^- \ , \quad s_k = i(b_k^+ - b_k^-)$$

qui satisfont eux aussi à des relations d'anticommutation simples

$$(12')\qquad \{r_j,s_k\} = 0 \ , \quad \{r_j,r_k\} = \{s_j,s_k\} = 2\delta_{jk} I$$

en particulier, $r_j^2 = s_j^2 = I$: r_j et s_j sont des symétries . Si A est une partie de $\{1,\dots,M\}$, que nous écrivons $A=\{i_1<i_2..<i_k\}$, nous posons $r_A = r_{i_1}\dots r_{i_k}$ et de même s_A . Comme dans la section précédente, l'algèbre d'opérateurs engendrée par les $b_k^\pm$ et l'identité (ou, ce qui revient au même, les r_k,s_k et I) contient tous les opérateurs sur $\mathcal{Y}$. En jouant à saute-mouton grâce à l'anticommutation, et en réduisant les puissances grâce à $r_j^2 = s_j^2 = I$, on voit que les opérateurs $r_A s_B$ forment une base de l'espace de tous les opérateurs sur $\mathcal{Y}$.

Recopions quelques formules sur les opérateurs ainsi construits

$$(14)\qquad r_k X_B = (-1)^{n(k,B)} X_{B\Delta\{k\}} \ , \text{ donc } r_A X_B = (-1)^{n(A,B)} X_{A\Delta B}$$

Ensuite, nous avons $s_k X_B = i(-1)^{n(k,B)} X_{B\Delta\{k\}} (I_{B^c}(k) - I_B(k))$. Ce dernier facteur (+1 si $k\notin B$, -1 si $k\in B$) peut s'écrire $(-1)^{n(k,B)+n'(k,B)+|B|}$. Donc

$$(15)\qquad \begin{aligned} s_k X_B &= i(-1)^{n'(k,B)} (-1)^{|B|} X_{B\Delta\{k\}} \\ s_A X_B &= i^{|A|} (-1)^{n'(A,B)+|A||B|+|A|-1} X_{A\Delta B} \end{aligned}$$

Enfin, nous avons

$$(16)\qquad r_A r_B = (-1)^{n(A,B)} r_{A\Delta B} = (-1)^{n(A,B)+n'(A,B)} r_B r_A \text{ (de même pour s)}$$

En particulier, $r_A^2 = s_A^2 = (-1)^{n(A,A)} I$, avec $n(A,A)=|A|(|A|-1)/2$. D'autre part

$$s_B r_A = (-1)^{|A||B|} r_A s_B \quad . \tag{17}$$

3. L'algèbre de Clifford.

Donnons nous un entier ν , et considérons une algèbre $\mathcal{G}$ sur $\mathbb{C}$, à unité 1, engendrée en tant qu'algèbre par 1 et par ν générateurs ε_k assujettis aux relations

$$\varepsilon_i^2 = 1 \ , \quad \varepsilon_i\varepsilon_j+\varepsilon_j\varepsilon_i = 0 \quad (i\neq j) \quad . \tag{18}$$

Alors, si nous définissons comme ci-dessus les produits ε_A , ces règles entraînent que les ε_A engendrent $\mathcal{G}$ en tant qu'espace vectoriel. Si les ε_A forment un système libre, nous connaissons la table de multiplication $\varepsilon_A\varepsilon_B=(-1)^{n(A,B)}\varepsilon_{A\Delta B}$, et l'algèbre est déterminée à isomorphisme près : on l'appelle algèbre de Clifford de dimension 2^ν, et nous la noterons $\underline{\underline{C}}(\nu)$.

Les résultats obtenus plus haut peuvent s'exprimer en disant que les r_i d'une part, les s_i d'autre part, sont les générateurs d'une algèbre d'opérateurs isomorphe à $\underline{\underline{C}}(M)$, et les (r_i,s_j) pris ensemble les générateurs d'une algèbre isomorphe à $\underline{\underline{C}}(2M)$. Mais nous avons vu que l'algèbre engendrée par les (r_i,s_j) est l'algèbre de tous les opérateurs sur $\mathcal{Y}$. Cela va avoir une conséquence importante.

Considérons une algèbre du type précédent, à 2M générateurs. L'application $I\mapsto 1$, $r_i\mapsto\varepsilon_i$, $s_i\mapsto\varepsilon_{M+i}$ se prolonge en un homomorphisme de l'algèbre engendrée par les (r_i,s_j) sur $\mathcal{G}$. Mais l'algèbre de tous les opérateurs sur $\mathcal{Y}$ n'admet pas d'idéal non trivial, donc le noyau est nul, et les ε_A sont forcément libres . Autrement dit, en dimension $\nu=2M$ paire, il n'y a qu'un seul type d'algèbre satisfaisant à (18), et il est isomorphe à l'algèbre de toutes les matrices $(2^\nu,2^\nu)$. Ce que nous avons montré s'énonce en langage algébrique en disant que $\underline{\underline{C}}(2M)$ est une algèbre simple. Nous construirons plus loin un opérateur σ sur $\mathcal{Y}$, tel que $\sigma^2=I$, anticommutant aux r_i,s_j : comme σ est une combinaison linéaire des r_As_B , il existe une algèbre du type (18) avec $\nu=2M+1$ telle que les ε_A ne soient pas libres, et cela signifie que $\underline{\underline{C}}(2M+1)$ n'est pas simple.

Les algèbres de Clifford ont une très grande importance en mathématiques et en physique. J'ai trouvé magnifique l'exposé de J. Helmstetter , Algèbres de Clifford et algèbres de Weyl, Cahiers Math. n°25, Montpellier 1982. Je le recommande vivement.

Remarques. a) Dans une algèbre de Clifford $\underline{\underline{C}}(\nu)$ à ν générateurs ε_i, les $\nu-1$ éléments $i\varepsilon_{12},\ldots i\varepsilon_{1n}$ satisfont aux relations d'anticommutation et à la propriété de liberté, donc engendrent une algèbre de Clifford isomorphe à $\underline{\underline{C}}(\nu-1)$, que l'on appelle l'algèbre paire.

b) Comment rendre les définitions précédentes aussi indépendantes que

possible du choix d'une base ? Soit H l'espace vectoriel complexe engendré par les ε_i , et soit (,) la forme bilinéaire complexe non dégénérée (pas une forme hermitienne !) pour laquelle les ε_i constituent une base orthonormale. La multiplication ne dépend que de H et de la forme (,). En effet, elle est décrite par la relation

$$uv + vu = (u,v)1 \quad \text{si } u,v \text{ appartiennent à } H$$

qui pour les éléments d'une base orthonormale quelconque redonnent (18). Les éléments de l'algèbre de Clifford engendrée par H sont alors appelés spineurs sur H .

(L'espace de Minkowski de la relativité restreinte est muni d'une forme bilinéaire réelle non dégénérée : lorsqu'on le complexifie, on obtient une structure du type précédent, et la théorie des spineurs et de l'algèbre de Clifford $\underline{\underline{C}}(4)$ se trouve ainsi étroitement liée à la théorie de la relativité. Nous n'avons pas à parler ici de ce genre d'applications de l'algèbre de Clifford).

c) L'algèbre de Clifford non triviale la plus simple est $\underline{\underline{C}}(2)$, et nous l'avons déjà rencontrée à propos du spin. Si nous prenons M=1, l'espace Ψ a pour base naturelle (1,X), et l'espace de tous les opérateurs sur Ψ a pour base (I,r,s,rs) , avec la table de multiplication $r^2=s^2=I$, rs=-sr. Si l'on choisit pour base (I,r,s,irs), la table de multiplication est celle des matrices de Pauli. Si l'on choisit (I,ir,is,-rs), la table est celle des quaternions (à coefficients complexes). L'algèbre paire est engendrée par I et rs, sa table de multiplication est celle de $\mathbb{C}$.

4. Définition d'un produit de Clifford sur Ψ . Nous avons promis au début du paragraphe de définir d'autres multiplications intéressantes pour nos v.a. de Bernoulli X_i : nous y arrivons.

Définissons une structure d'algèbre (nous n'exigeons pas l'associativité a priori) par la table de multiplication

$$X_A \cdot X_B = (-1)^{n(A,B)} X_{A\Delta B} \tag{19}$$

Un coup d'oeil à (14) montre qu'alors r_A est l'opérateur $X_A \cdot$: comme la composition des opérateurs est associative, la multiplication l'est aussi, et les propriétés $X_i^2=1$, $X_iX_j+X_jX_i=0$ montrent que Ψ est une algèbre de Clifford $\underline{\underline{C}}(M)$. Plus tard, nous étendrons cela à une famille continue de v.a. de Bernoulli, i.e. au mouvement brownien.

Nous avons plongé Ψ dans l'algèbre $\mathfrak{L}(\Psi)$ des opérateurs sur Ψ, par l'application $X_A \longmapsto r_A$ (prolongée par linéarité). En ramenant sur Ψ la norme $\| \|$, le passage à l'adjoint * , nous obtiendrons une C^*-algèbre fort intéressante. Nous allons en dégager la structure.

Rappelons d'abord, pour bien fixer le langage, que Ψ est un e.v. complexe, muni d'une conjugaison $x \longmapsto \overline{x}$ qui laisse fixe les X_A . Nous avons

donc sur Ψ le produit hermitien $\langle x,y\rangle$, et la forme bilinéaire (produit scalaire euclidien) $(x,y)=\langle\overline{x},y\rangle$. On a pour le produit de Clifford $\overline{uv} = \overline{u}.\overline{v}$ (nous omettrons les . la plupart du temps désormais).

On voit sur (14) que les opérateurs r_A sont <u>unitaires</u>, donc $r_A^* = r_A^{-1}$, ce qui nous donne

$$X_A^* = (-1)^{n(A,A)}X_A = (-1)^{|A|(|A|-1)/2}\,X_A \tag{20}$$

l'extension aux autres éléments se faisant par <u>anti</u>linéarité. Les propriétés $(uv)^*=v^*u^*$, $u^{**}=u$ sont héritées de $\mathfrak{L}(\Psi)$. Si l'on pose

$$\rho(u) = \overline{u}^* \tag{21}$$

on obtient une application linéaire de Ψ dans Ψ , telle que $\rho^2=I$: c'est le prolongement de (20) par linéarité. On a $\rho(uv)=\rho(v)\rho(u)$; la matrice de ρ dans la base (X_A) étant une diagonale réelle, ρ est un opérateur a.a..

<u>Norme</u>. La première chose que l'on voit , c'est que les r_A ont une norme d'opérateur égale à 1, donc $\|X_A\|=1$, et en particulier $\|X_i\|=1$. Il est intéressant de calculer aussi la norme de l'opérateur $X_f=\Sigma_i f^i X_i$. Si f est <u>réelle</u> et sa norme hilbertienne est 1, on peut faire entrer X_f dans une base à laquelle s'appliquent les calculs précédents, et on en déduit que $\|X_f\| = \|f\|$ pour f réelle. Pour f complexe telle que $\langle f,f\rangle=\|f\|^2=1$, posons $\alpha=\langle\overline{f},f\rangle$, nombre complexe de module $\leqq 1$, et soient F l'opérateur de multiplication par X_f , G l'opérateur a.a. positif F^*F. On a

$$G^2 = F^*(F^*F+FF^*)F - (F^*F^*)(FF) = 2G - |\alpha|^2 I$$

donc la plus grande valeur propre de G vaut $1+\sqrt{1-|\alpha|^2}$: elle est égale à $\|G\|$ et à $\|F\|^2=\|X_f\|^2$.

<u>Loi de probabilité</u>. L'espérance d'une v.a. $\Sigma_A \lambda_A X_A$ au sens de la situation probabiliste de départ est tout simplement égale à $\lambda_\emptyset$. Du point de vue algébrique, c'est une forme linéaire distinguée : elle est caractérisée par sa nullité sur les Ψ_n $(n>0)$ et la propriété $E(1)=1$.

Remarquons que tous les r_A sont des matrices de trace nulle, sauf $r_\emptyset=I$. Ainsi, si l'on identifie les éléments de Ψ aux opérateurs de multiplication de Clifford correspondants, on a $E[x]=2^{-M}\mathrm{Tr}(x)$ pour tout $x\in\Psi$. Comme le produit et la conjugaison dans Ψ correspondent à la multiplication et au passage à l'adjoint sur les opérateurs associés, on a

$$E[xy] = E[yx] \quad , \quad E[y^*y] \geqq 0 \quad \text{pour } x,y\in\Psi \ .$$

Nous avons donc une loi de probabilité au sens C^*-algébrique, qui est une trace. Le produit scalaire euclidien dans Ψ est $(x,y)=E[xy]$, et le produit scalaire hermitien $\langle x,y\rangle=E[x^*y]$.

Un dernier élément de structure intéressant sur l'algèbre Ψ est l'opérateur a.a. de carré I

(21) $$\sigma(X_A) = (-1)^{|A|} X_A$$

Il est très facile de vérifier que $\sigma(xy)=\sigma(x)\sigma(y)$ (c'est un automorphisme de l'algèbre de Clifford) et qu'il anticommute aux opérateurs p_i et q_j (mais les opérateurs p_i, q_j, σ n'engendrent pas une algèbre isomorphe à $\underline{\underline{C}}(2M+1)$: il n'y a pas la place). Si nous posons $\tau=\rho\sigma=\sigma\rho$, nous avons le petit tableau suivant, qui nous dit par quel coefficient l'opérateur de la colonne de gauche opère sur le "k-ième chaos" Ψ_k, suivant la classe de k mod. 4

(22)

	classe de k :	0	1	2	3	
opérateur	σ :	1	-1	1	-1	automorphisme
	ρ :	1	1	-1	-1	inversent le sens
	τ :	1	-1	-1	1	des produits
	$\mathcal{F}$:	1	i	-1	-i	transforme les r_i en s_i

On retrouve tous ces objets en temps continu, en remplaçant les v.a. de Bernoulli par le mouvement brownien, à propos du calcul stochastique sur les fermions.

5. Une remarque. Dans la première partie du paragraphe, nous avons rencontré non seulement les v.a. de Bernoulli symétriques, satisfaisant à $X_j^2=1$, mais des v.a. satisfaisant à $X_j^2=1+cX_j$ avec $c\neq 0$. Existe t'il des objets analogues dans le cas non commutatif ?

La première idée, pour construire des v.a. ne commutant pas possédant la propriété ci-dessus, consiste à copier ce que nous avons fait pour les v.a. commutatives, c'est à dire à introduire les opérateurs a.a. $\xi_j=r_j+c\mathbb{N}_j$. Ceux-ci satisfont à $\xi_j^2=1+c\xi_j$, mais leur composition ne présente aucune propriété algébrique simple. En revanche, si l'on pose

$$\eta_j = r_j + \frac{c}{2}(I+s_j) \quad , \quad \zeta_j = s_j + \frac{c}{2}(I+r_j)$$

ces opérateurs satisfont à

$$\eta_j^2 = I + c\eta_j \quad , \quad \zeta_j^2 = I + c\zeta_j$$

$$\{\eta_j,\eta_k\} = c(\eta_j+\eta_k) \ , \ \{\zeta_j,\zeta_k\} = c(\zeta_j+\zeta_k) \ , \ \{\eta_j,\zeta_k\} = \frac{c}{2}(\eta_j+\eta_k+\zeta_j+\zeta_k-\frac{c}{2}I)$$

L'expression exacte ne nous intéresse pas vraiment : ce qui nous intéresse, c'est qu'en regardant seulement les η_j , il existe une algèbre associative engendrée par des générateurs $1,\varepsilon_j$ satisfaisant aux relations

(23) $$\varepsilon_j^2 = 1+c\varepsilon_j \ , \ \varepsilon_i\varepsilon_j+\varepsilon_j\varepsilon_i = c(\varepsilon_i+\varepsilon_j) \ (i\neq j)$$

et telle que les ε_A soient libres (les $\eta_A\zeta_B$ ci-dessus formant manifestement une base de l'algèbre des opérateurs sur Ψ). Savoir si cette algèbre présente une utilité quelconque est une autre affaire !

APPENDICE.

Nous nous proposons ici de décrire, d'après l'exposé de Helmstetter cité plus haut, et aussi d'après le début d'un article très considérable de Sato, Miwa et Jimbo, Holonomic Quantum Fields I, Publ. RIMS Kyoto 14, 1977, comment l'algèbre de Clifford permet de représenter les éléments du groupe orthogonal d'un espace euclidien complexe H. Cela ne nous servira pas de manière immédiate, mais je pense qu'il peut y avoir des situations analogues intéressantes en temps continu.

Reprenons donc notre espace de dimension finie H, avec sa base X_i (orthonormale à la fois pour la structure hilbertienne complexe et pour la structure euclidienne complexe), et notre espace Ψ des spineurs sur H, avec sa base X_A , à 2^M éléments.

Soit g un élément inversible de Ψ ; posons pour $x\in\Psi$ $T_g x = gx\sigma(g^{-1})$. On vérifie sans peine, σ étant un automorphisme, que $T_g T_{g'} = T_{gg'}$, mais T_g n'est pas en général un automorphisme de l'algèbre (si g est impair, on a $T_g 1 = -1$). On dit que g appartient au groupe de Clifford si, de plus, T_g préserve H . S'il en est ainsi, T_g préserve le produit scalaire euclidien (,) sur H . En effet, soit $x\in H$. On a $\sigma(T_g x) = \sigma(g)\sigma(x)g^{-1}$ et $\sigma(x) = -x$, et d'autre part notre hypothèse entraîne que $\sigma(T_g x) = -T_g x$. Donc

$$\sigma(g)xg^{-1} = gx\sigma(g^{-1})$$

$$(T_g x, T_g x) = (T_g x)^2 = gx\sigma(g^{-1}).\sigma(g)xg^{-1} = g(x^2)g^{-1} = x^2 = (x,x) .$$

Exemple fondamental. Soit $g\in H$ tel que $g^2 = (g,g) = 1$. Alors $g^{-1} = g$ et $T_g x = -gxg$. Si $x\in H$ est colinéaire à g, on a $T_g x = -x$, et si x est orthogonal à g on a $xg = -gx$, $T_g x = x$. Donc T_g est la symétrie par rapport à l'hyperplan $g^\perp$ de H . On en déduit que tout g non isotrope ($(g,g)\neq 0$) appartient au groupe de Clifford, ($T_{cg} = T_g$ pour tout scalaire $c\neq 0$).

Le lemme suivant est classique :

LEMME. Toute transformation orthogonale sur H est un produit de symétries, et en conséquence est de la forme T_g , où g est un produit d'éléments de H non isotropes.

(Démonstration télégraphique, d'après Helmstetter : on raisonne par récurrence sur la dimension M de H . Si S est une transformation orthogonale et s'il existe x non isotrope tel que $Sx = x$, S laisse stable l'hyperplan $x^\perp$, et on peut utiliser l'hypothèse de récurrence dans cet hyperplan: Tout revient donc à trouver un produit de symétries U tel que US laisse fixe un vecteur non isotrope. Soit x un vecteur non isotrope normalisé ($x^2 = 1$) et soient $u = Sx - x$, $v = Sx + x$. Si u est non isotrope on peut prendre $U = T_u$. Si u est isotrope, on a $4(x,x) = (Sx-x, Sx-x) + (Sx+x, Sx+x) = u^2 + v^2 = v^2$, donc $v^{-1} = v/2$, et $xSx = -(Sx)x$. Ainsi T_v transforme Sx en $-v(Sv)v/2 = -x$, puis $T_x(-x) = x$, donc $T_x T_v Sx = x$ et on a gagné .)

Ainsi toute transformation orthogonale se représente sous la forme T_g, où $g = g_1 \ldots g_k$ est un produit d'éléments de H . Nous allons étudier

l'unicité de cette représentation.

Les produits $\mu(g)=g\rho(g)$ et $\nu(g)=g\tau(g)$ ne dépendent que de g. Or le premier vaut $g_1\ldots g_k g_k\ldots g_1 = \Pi_i (g_i,g_i)$, et le second vaut la même chose, multipliée par $(-1)^k$. On a évidemment $\mu(g)\mu(g')=\mu(gg')$, $\nu(g)\nu(g')=\nu(gg')$. Lorsque les g_i sont réels, $\mu(g)$ et $\nu(g)$ sont réels et $\mu(g)=|\nu(g)|$, de sorte que le second contient un peu plus d'information.

THEOREME. Tout élément du groupe orthogonal sur H est la restriction à H d'une transformation T_g, où g est un produit d'éléments de H non isotropes. Si l'on normalise g par la condition $\mu(g)=1$, g est déterminé à un facteur ± 1 près.

Démonstration. L'existence a déjà été vue. Pour l'unicité, désignons par Ξ l'élément de base $X_1\ldots X_M$, appartenant au "chaos d'ordre M" Ψ_M.

LEMME. Si g commute à tout élément de H, g est dans Ψ_0 pour M pair, dans $\Psi_0+\Psi_M$ pour M impair (si g anticommute à tout élément de H, on a $g=0$ pour M impair, $g\in\Psi_M$ pour M pair).

Ici g est un élément quelconque de Ψ, non nécessairement décomposable. Traitons par ex. le cas de la commutation. Ecrivons $g=\Sigma_A g_A$, où g_A est proportionnel à X_A. La relation $X_i g=gX_i$ entraîne $X_i X_A=X_A X_i$ pour tout A tel que $g_A\neq 0$; or les relations de commutation de X_i et X_A dépendent de i, sauf si A est vide ou plein, donc $g\in\Psi_0+\Psi_M$. On conclut en remarquant que Ξ commute avec H pour M impair, anticommute pour M pair. □

Passons au théorème. Supposons que $T_g|_H=I$, ou encore $gx=x\sigma(g)$ pour $x\in H$. Cela entraîne $(g+\sigma(g))x=x(g+\sigma(g))$. D'après le lemme, $g+\sigma(g)$ appartient à Ψ_0 si M est pair, à $\Psi_0+\Psi_M$ si M est impair - en fait à Ψ_0 dans les deux cas, car lorsque M est impair, $g+\sigma(g)$, qui est pair, n'a pas de composante suivant Ψ_M.

Donc $g+\sigma(g)\in\Psi_0$, et l'on peut écrire $g=c+i$ avec c scalaire et i impair. Alors i anticommute à tout $x\in H$, donc d'après le lemme $i=0$ ou $i\in\Psi_M$ si M est pair. Comme i est impair on a $i=0$ dans les deux cas, et g est un scalaire c.

Soit g un élément du groupe de Clifford (non nécessairement décomposable a priori). Soit γ un élément décomposable induisant la même transformation orthogonale. Le résultat qui vient d'être établi montre que $g\gamma^{-1}$ est un scalaire, donc en fait g était décomposable. Si l'on norme g convenablement, g est déterminé à un facteur $\pm$près, et cette ambiguité ne peut être levée.

REMARQUE. $T_g x = gxg^{-1}$ si g est pair, $-gxg^{-1}$ si g est impair. Lorsque M est pair (ce qui est le cas intéressant) on peut toujours se ramener à la première forme, car $\Xi x\Xi^{-1}=-x$ pour $x\in H$. On peut montrer que g s'écrit toujours $g_1\ldots g_k$ avec $k\leq n$, mais nous n'aurons pas besoin de ce résultat.

ELEMENTS DE PROBABILITES QUANTIQUES. III

Le couple canonique

Cet exposé est entièrement consacré à la notion de couple canonique. Il s'agit d'un modèle d'espace mesurable quantique, engendré par deux v.a. P,Q satisfaisant à une forme précisée de la relation d'Heisenberg

$$[P,Q] = \frac{1}{i}\hbar I \quad .$$

Ce modèle est fondamental pour la mécanique quantique, si important aussi pour les probabilités quantiques que je me rappelle avoir lu la phrase, un peu exagérée mais bien significative << les couples canoniques jouent en probabilités quantiques le rôle des variables aléatoires classiques >> .

Nous commençons par définir rigoureusement le couple canonique, sans commutateurs d'opérateurs non bornés, et par en donner le modèle classique (de Schrödinger). Nous donnons la forme précise des relations d'incertitude, puis introduisons, à propos de l'oscillateur harmonique quantique, un second modèle, avec les opérateurs de création et d'annihilation. Nous donnons quelques exemples de lois de probabilité sur un couple canonique, pures ou non (les lois gaussiennes). Nous reproduisons enfin certains résultats classiques sur les << fonctions de Wigner >> et leurs fonctions caractéristiques.

I. LE THEOREME DE STONE-Von NEUMANN

1. Le théorème de Stone-von Neumann est un résultat fondamental (considérablement généralisé par Mackey, mais nous ne nous occuperons pas ici de cette extension), qui est particulièrement attirant pour les probabilistes, parce qu'il s'exprime très bien dans leur langage usuel. En fait, il a été redécouvert par des probabilistes ! (O. Hanner, Deterministic and non-deterministic processes, Ark. Math 1, 1950 (temps discret) ; G. Kallianpur et V. Mandrekar , Multiplicity and representation theory of purely non-deterministic stochastic processes. Teor. Veroj. 10, 1965, et Ark. för Mat., 6, 1965). Voir aussi Lazaro-Meyer, Questions de théorie des flots, Sém. Prob. IX, Lecture Notes 465, Springer 1975. et ZW 18, 1971, p.116-118.

Voici la situation étudiée par ces auteurs. On considère un espace probabilisé classique $(\Omega,\mathcal{F},P)$, muni d'un flot (un groupe mesurable $(\theta_t)_{t\in\mathbb{R}}$ de transformations de Ω préservant P) et d'une filtration $(\mathcal{F}_t)_{t\in\mathbb{R}}$, liés par la relation

si f est $\mathcal{F}_t$-mesurable, $f\circ\theta_s$ est $\mathcal{F}_{s+t}$-mesurable .

On pose

$$U_t f = f \circ \theta_t \quad , \quad E_t f = E[f|\mathcal{F}_t]$$

de sorte que (U_t) est un groupe unitaire, (E_t) une famille de projecteurs, et que l'on a

$$U_t E_s = E_{s+t} U_t \quad . \tag{1}$$

On dit que le flot est un K-flot (ou est purement stochastique) si la tribu $\mathcal{F}_{-\infty}$ est dégénérée. Cela signifie que si l'on se restreint aux fonctions d'intégrale nulle, les espaces de Hilbert $\mathcal{H}_t = L^2_0(\mathcal{F}_t)$ constituent une famille spectrale sur l'espace $\mathcal{H} = L^2_0(\mathcal{F}_t)$.

Il est clair maintenant que la situation probabiliste précédente admet une version purement hilbertienne : un espace de Hilbert $\mathcal{H}$, muni d'une famille spectrale $(\mathcal{H}_t)_{t\in\mathbb{R}}$ et d'un groupe unitaire (U_t), les projecteurs spectraux E_t étant liés au groupe par (1). Un exemple intéressant de cette situation hilbertienne (lié à un flot aussi, mais en mesure infinie), que nous appellerons le modèle, est le suivant :

$\mathcal{H} = L^2(\mathbb{R})$ (mesure de Lebesgue) , $\mathcal{H}_t = L^2(]-\infty, t])$, $U_t f(x) = f(x-t)$.

Revenons à la situation probabiliste. On appelle hélice du flot un processus $(X_t)_{t\in\mathbb{R}}$, possédant les propriétés suivantes :

1) $X_0 = 0$; $X_t \in L^2(\mathcal{F}_t)$ pour $t \geq 0$; $E[X_t - X_s | \mathcal{F}_s] = 0$ pour $0 \leq s \leq t$ (propriété de martingale pour $t \geq 0$) ; $E[X_t^2] = t$.

2) $X_{t+s} - X_s = U_s X_t$ pour tout (s,t) (X est une "fonctionnelle additive")

On peut alors définir pour toute fonction $f \in L^2(\mathbb{R})$ une intégrale stochastique $\int f(s)dX_s = I_f$ étendue à $\mathbb{R}$ entier, et l'on a $E[I_f^2] = \|f\|_2$. De plus

$$U_t\left(\int f(s)dX_s\right) = \int f(s-t)dX_s \quad , \quad E_t\left(\int f(s)dX_s\right) = \int_{-\infty}^{t} f(s)dX_s$$

Si nous désignons par $L^2(X)$ l'espace des intégrales stochastiques I_f pour $f \in L^2(\mathbb{R})$, nous voyons donc que $L^2(X)$ est stable par les opérateurs U_t et E_t, et isomorphe au modèle. Le théorème de structure des flots filtrés affirme alors l'existence d'une famille (X^i) d'hélices, telle que l'on ait $L^2_0(\mathcal{F}) = \oplus_i L^2(X^i)$. La démonstration, voisine de celle de résultats classiques en théorie des processus de Markov, ne présente aucune difficulté pour un probabiliste (cf. par ex. ZW 18, 1971).

Si l'on revient à la situation hilbertienne générale, on s'aperçoit que l'on n'a qu'une simple traduction à faire, et on obtient le théorème suivant :

THEOREME. Soit $\mathcal{H}$ un espace de Hilbert, muni d'une famille spectrale $(\mathcal{H}_t)_{t\in\mathbb{R}}$ et d'un groupe unitaire (U_t) fortement continu, les projecteurs spectraux E_t étant liés aux (U_t) par la relation (1). Alors $\mathcal{H}$ est somme directe hilbertienne d'une famille de sous-espaces stables, isomorphes au modèle.

Le mot stable signifie : stable par les U_t et les E_t (et fermé). On dit que $\mathcal{H}$ est irréductible s'il ne contient aucun sous espace stable distinct de $\{0\}$ et de $\mathcal{H}$ entier. Montrons que le modèle est irréductible. D'après le théorème lui même, il suffit de montrer que le modèle ne contient qu'une seule hélice normalisée. Nous omettons la discussion un peu délicate des "p.p." : soit (X_t) une hélice normalisée. Pour $t\geq 0$, $X_t - X_0 = X_t$ est orthogonale à $L^2(]-\infty,0])$, donc $X_t(x)=0$ pour $x\leq 0$. Pour $t\geq 0$, $h\geq 0$, $t\leq x\leq t+h$ on a $X_{t+h}(x)-X_t(x)=X_h(x-t)$ et $X_t \in L^2(]-\infty,t])$ est nul en x , donc $X_{t+h}(x)=X_h(x-t)$. Il en résulte qu'il existe une fonction $F(x)$ (non nécessairement dans L^2) nulle sur $\mathbb{R}_+$ telle que

$$X_t(x) = F(x-t) \quad \text{pour } t\geq 0,\ x\geq 0$$

Ecrivant que pour $t>0$, $h>0$, $X_{t+h}-X_t$ est nulle sur $]-\infty,t]$, on voit que F est égale p.p. à une constante c . Alors

$$\text{pour } t>0 \ , \ X_t = cI_{]0,t]}$$

La normalisation nous donne $c=1$, puis on a pour $t<0$ $X_t=-I_{[t,0]}$. L'unicité est établie.

Nous allons modifier légèrement l'énoncé et les notations. Au lieu de prendre comme groupe unitaire $U_t f(x)=f(x-t)$ sur le modèle, nous prendrons

$$U_t f(x) = f(x+\hbar t) \tag{2}$$

dont le générateur P ($U_t=e^{itP}$) opère par

$$Pf(x) = \frac{\hbar}{i}\frac{d}{dx}f(x) \tag{3}$$ si cette dérivée existe au sens L^2 .

Au lieu d'utiliser la famille spectrale, nous faisons intervenir le groupe unitaire associé

$$V_t f = \int e^{itu} dE_u \quad , \quad V_t f(x) = e^{itx} f(x) \tag{4}$$

dont le générateur Q ($V_t=e^{itQ}$) opère par

$$Qf(x)=xf(x) \text{ si cette fonction appartient à } L^2 \ . \tag{5}$$

La traduction des relations de commutation (1), après ces changements de notation, est la relation de Weyl entre les deux semi-groupes :

$$U_t V_s = e^{i\hbar st} V_s U_t \tag{6}$$

et nous avons le théorème suivant :

THEOREME. *Soit $\mathcal{H}$ un espace de Hilbert complexe, muni de deux groupes unitaires fortement continus satisfaisant à (6). Alors $\mathcal{H}$ est somme directe hilbertienne de sous-espaces stables orthogonaux isomorphes au modèle. En particulier, $\mathcal{H}$ est isomorphe au modèle si et seulement s'il est irréductible.*

Nous dirons que deux opérateurs a.a. P,Q forment un couple canonique si les groupes unitaires qu'ils engendrent satisfont à (6). Nous supposerons le plus souvent, implicitement, que ces groupes opèrent de manière

irréductible. On notera bien que la notion est de nature " mesurable ", antérieure à la donnée d'une loi de probabilité quantique.

Exemple d'application du théorème. Prenons le modèle lui même, avec $\hbar=1$ pour simplifier, et posons $U'_t=V_{-t}$, $V'_t=U_t$ ($P'=-Q$, $Q'=P$). Alors les nouveaux groupes opèrent irréductiblement, et satisfont à (6), donc il existe un isomorphisme unique de $L^2(\mathbb{R})$ sur lui même transformant x en D, D en -x : c'est la transformation de Fourier.

COMMENTAIRE. Dans l'interprétation probabiliste ci-dessus, le paramètre t du groupe unitaire (U_t) avait une signification temporelle. Il n'en est pas ainsi en physique . Les groupes unitaires représentant une évolution en physique sont presque toujours de la forme $e^{itH/\hbar}$, avec un hamiltonien H représentant une énergie positive, tandis que l'opérateur P a un spectre non borné inférieurement.

II. CALCULS SUR LES COUPLES CANONIQUES

1. Les relations d'incertitude. Il est bon de retrouver les relations d'incertitude par un raisonnement direct sur le modèle, dans le cas explicite du couple canonique. Nous prendrons $\hbar=1$ pour alléger les notations.

Un élément ω de $L^2(\mathbb{R})$ tel que $E_\omega[Q^2]<\infty$, $E_\omega[P^2]<\infty$ est une fonction dérivable au sens L^2, telle que

$$\int|\omega(x)|^2dx < \infty \ , \ \int|\omega'(x)|^2dx < \infty \ , \ \int x^2|\omega(x)|^2dx < \infty .$$

La dérivabilité-L^2 entraîne en particulier que ω admet une version continue. Nous normalisons ω , et il s'agit de vérifier que

$$E_\omega[P^2]E_\omega[Q^2] = \langle\omega,P^2\omega\rangle\langle\omega,Q^2\omega\rangle = \|P\omega\|^2\|Q\omega\|^2 = (\int|\omega'|^2dx)(\int x^2|\omega|^2dx) \geq \frac{1}{4}$$

D'après l'inégalité de Schwarz, il suffit que $\int|x\omega(x)\omega'(x)|dx \geq 1/2$.
Si nous remplaçons ω par $|\omega|$, nous préservons la dérivabilité-L^2 et diminuons le module de la dérivée en presque tout point, donc nous pouvons supposer ω réelle. Nous avons alors $2\int_a^b x\omega\omega'dx = x\omega^2(x)|_a^b - \int_a^b \omega^2dx$ sur tout intervalle fini ; comme $x\omega\in L^2$, on peut trouver des $a_n\to-\infty$, $b_n\to+\infty$ tels que le premier terme à droite tende vers 0, et il reste $2\int x\omega\omega'dx = -\int\omega^2dx = -1$, d'où le résultat annoncé. Cette démonstration est due à H. Weyl.

Etats d'incertitude minimale. Nous conservons la convention $\hbar=1$. Recherchons les fonctions $\omega(x)$ réalisant le minimum dans la relation précédente. Nous avons déjà fait remarquer que le remplacement de ω par $|\omega|$ diminue les deux espérances, donc on peut supposer ω réelle. L'égalité étant atteinte dans l'inégalité de Schwarz $(\int\omega^2dx)(\int x\omega'^2dx)=(\int x\omega\omega'dx)^2$, on a $\omega'(x)=cx\omega(x)$ où c est une constante, d'où la forme de ω

$$\omega(x) = (2\pi c)^{-1/4}e^{-x^2/4c} \quad (c>0 \text{ du fait que } \omega\in L^2) .$$

Il pourrait a priori exister des fonctions complexes de même module,

réalisant elles aussi le minimum. Ecrivant $\omega=\Theta|\omega|$, et remarquant que $|\omega|$ calculée ci-dessus ne s'annule pas, on voit que Θ est dérivable. Alors on a $|\omega'|^2 \geqq |\omega|'^2$ là où $\Theta'\neq 0$, donc Θ est constante, et finalement se réduit à un facteur de phase de module 1.

Revenant aux incertitudes (variances au lieu des moments du second ordre) et rétablissant $\hbar$, on peut recopier dans un cours de mécanique quantique l'expression des << fonctions d'onde >> d'incertitude minimale, exprimées sur la représentation en position (q) ou en impulsion (p) ; $\langle q\rangle$ ou $\langle p\rangle$ est la notation des physiciens pour une moyenne.

$$(7)\qquad \begin{aligned} \phi_q(x) &= (2\pi c_q)^{-1/4}\exp\Big[\frac{-ix\langle p\rangle}{\hbar} - \frac{(x-\langle q\rangle)^2}{4c_q}\Big] \\ \phi_p(u) &= (2\pi c_p)^{-1/4}\exp\Big[\frac{-iu\langle q\rangle}{\hbar} - \frac{(u-\langle p\rangle)^2}{4c_p}\Big] \end{aligned} \qquad c_p c_q = \hbar^2/4$$

($\phi_p(u)$ est, à un facteur de phase près que nous avons supprimé, la transformée de Fourier de $\phi_q(x)$). On notera que la belle normalisation probabiliste, celle qui fait correspondre à une loi gaussienne réduite en q une loi gaussienne réduite en p, correspond à $\hbar^2/4=1$, donc à $\underline{\hbar=2}$ (c'est le même phénomène que pour le laplacien, les probabilistes utilisant $\frac{1}{2}\Delta$ et les analystes Δ).

2. Opérateurs de Weyl. Nous avons défini $U_r=e^{irP}$, $V_s=e^{isQ}$. Existe t'il une manière raisonnable de définir

$$W_{r,s} = e^{i(rP+sQ)} \quad ?$$

Un résultat formel, classique en théorie des groupes, dit que

$$\text{si } [A,[A,B]]=[B,[A,B]]=0,\ e^{A+B}=e^A e^B e^{-[A,B]/2}=e^B e^A e^{[A,B]/2}$$

Comme ici, $[P,Q]=-i\hbar I$, il est tentant de poser

$$W_{r,s} = U_r V_s e^{rs[P,Q]/2} = U_r V_s e^{-i\hbar rs/2} = V_s U_r e^{i\hbar rs/2}$$

ou sur le modèle, en ajoutant un paramètre t pour faire plus joli

$$(8)\qquad e^{i(rP+sQ+tI)}f(x) = W_{r,s,t}f(x) = e^{i(t+\hbar rs/2)}e^{isx}f(x+\hbar r)$$

Il est alors immédiat de vérifier que ces opérateurs $W_{r,s,t}$ sont effectivement unitaires, et se composent suivant la règle

$$W_{r,s,t}\,W_{r',s',t'} = W_{r+r',s+s',\ t+t'+\hbar(rs'-r's)/2}$$

qui pour $\hbar=1$ est la loi de composition des matrices

$$(r,s,t) = \exp\begin{pmatrix} 0 & r & t \\ 0 & 0 & s \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & r & t+rs/2 \\ 0 & 1 & s \\ 0 & 0 & 1 \end{pmatrix}$$

qui constituent le groupe de Heisenberg (le modèle fournit une représentation unitaire irréductible de ce groupe). Nous oublierons le paramètre t dans la suite, ainsi que le groupe d'Heisenberg. Nous retiendrons plutôt

la formule sans t

$$W_{r,s}\, W_{r',s'} = e^{i\hbar(rs'-r's)/2}\, W_{r+r',s+s'} \tag{9}$$

qui signifie que les $W_{r,s}$ forment une <u>représentation unitaire projective</u> du groupe additif $\mathbb{R}^2$: ce n'est pas une vraie représentation unitaire, mais elle fait intervenir un facteur de phase de module 1, de sorte que le groupe opère sur les événements, sur les lois de probabilité (en préservant les lois pures). Cette action du groupe $\mathbb{R}^2$ nous permet de dire comment se transforment les événements ou les lois de probabilité sous l'effet des changements de coordonnées sur l'espace des (p,q) (transformations de Galilée).

Une forme abrégée des relations de Weyl consiste à poser $W_z = W_{r,s}$ pour $z=r+is$. On a alors, en considérant $\mathbb{R}^2$ comme espace de Hilbert complexe

$$W_z W_{z'} = e^{i\hbar \mathrm{Im}\langle z,z'\rangle/2}\, W_{z+z'} \tag{10}$$

qui est la forme la plus importante des relations de Weyl en physique : elle se généralise en effet à un espace de Hilbert complexe quelconque, sans avoir besoin de le considérer comme complexifié d'un espace de Hilbert réel (ce qui lui donne un caractère plus intrinsèque).

Nous verrons dans un exposé suivant comment on peut étendre (10) en une représentation unitaire projective W_a du <u>groupe des déplacements</u> de H ((10) correspond aux translations seules).

Une extension relativement triviale des résultats ci-dessus consisterait à considérer un système de Weyl $W_{r,s}$ indexé par $\mathbb{R}^N \times \mathbb{R}^N$, correspondant à un système de N couples canoniques (P_i, Q_i) commutant entre eux. Malgré l'intérêt physique du cas N=3, nous n'en dirons rien.

3. <u>Définitions de divers opérateurs</u>. On a dit dans l'exposé I que l'addition d'opérateurs a.a. non bornés est toujours une opération délicate. Ici, les opérateurs P,Q du couple canonique ont un excellent domaine commun stable, l'espace $\underline{\underline{S}}$ de Schwartz (fonctions C^∞ dont toutes les dérivées sont rapidement décroissantes), et l'on peut donc formellement définir un opérateur $f(P,Q)$ pour tout polynôme f . Il est facile de reconnaître si cet opérateur est symétrique sur $\underline{\underline{S}}$; supposant cela vérifié, il se pose deux problèmes

i) $f(P,Q)$ admet il une extension autoadjointe (autrement dit, peut on considérer $f(P,Q)$ comme une v.a. ?),

ii) Cette extension est elle unique ?

Dans cette section, nous résoudrons le premier problème pour certains opérateurs simples, et ne dirons rien du second.

Opérateurs rP+sQ. Nous avons construit plus haut les opérateurs unitaires $W_{r,s}$, qui représentent formellement $e^{i(rP+sQ)}$. Il est donc tout naturel de définir rP+sQ comme

$$rP+sQ = \frac{1}{i}\frac{d}{dt}W_{tr,ts}|_{t=0} \tag{11}$$

Un calcul immédiat sur la formule explicite (8) montre que cette définition est correcte sur l'espace $\underline{\underline{S}}$. Ces opérateurs sont en principe des "observables" : je me rappelle avoir lu, dans un article de physique, qu'ils sont le meilleur exemple d'observables que personne ne sait obsèrver !

Soit μ une loi quantique pour le couple canonique : nous avons dit dans l'exposé I que (P,Q) n'ont pas de loi jointe, mais les opérateurs de Weyl en fournissent des substituts partiels. Par exemple, la fonction caractéristique, que nous étudierons en détail plus bas

$$\hat{\mu}(r,s) = E_{\mu}[e^{i(rP+sQ)}] \tag{12}$$

(elle n'est pas de type positif en général). En particulier, si $\omega(q)$ est une "fonction d'onde" normalisée

$$\int|\omega(q)|^2 đq=1 \quad (đq=dq/\sqrt{2\pi} \text{ (1) })$$

on a d'après (8)

$$\hat{\varepsilon}_{\omega}(r,s) = <\omega, W_{r,s}\omega> = \int \overline{\omega}(q)\omega(q+ħr)e^{iħrs/2}e^{isq}\, đq$$

$$= \int\overline{\omega}(x-ħ\tfrac{r}{2})\omega(x+ħ\tfrac{r}{2})e^{isx}\, đx \tag{13}$$

Par exemple, dans le cas de l'incertitude minimale, rencontré plus haut ($ħ=1$ pour simplifier)

$$\omega(q) = c_q^{-1/4}\, e^{-x^2/4c_q}$$

on trouve pour fonction caractéristique $e^{-c_q s^2/2 - r^2/8c_q}$, qui paraît curieusement dissymétrique en r et s - mais il n'en est rien : si l'on introduit c_p , lié à c_q par la relation d'incertitude minimale $c_p c_q = 1/4$, on a

$$\hat{\varepsilon}_{\omega}(r,s) = \exp(-\tfrac{1}{2}(c_p r^2 + c_q s^2)) \tag{14}$$

qui d'ailleurs est de type positif : cet état ressemble un peu plus aux lois classiques que la plupart des lois quantiques ! C'est l'exemple le plus simple de loi gaussienne quantique (on en verra d'autres).

Opérateur PQ+QP . Lorsque A et B sont deux opérateurs a.a. bornés, AB+BA est a.a.. Que se passe t'il ici ? Sur le modèle, et sur les bonnes fonctions, PQ+QP est l'opérateur $-iħ(xD+2I)$, et il est facile de construire le groupe unitaire correspondant : il s'agit des dilatations

$$\Lambda_t f(x) = e^{ħt} f(xe^{2ħt}) \quad .$$

1. Le choix de cette normalisation est le meilleur pour la formule de Plancherel.

L'oscillateur harmonique quantique. Voici l'occasion de rappeler certains points qui figurent au début de tous les cours de mécanique quantique. Pour le mouvement d'une particule de masse m sous l'action d'un potentiel V(x), la mécanique classique introduit les coordonnées $q=x$, $p=m\dot{x}$ (position et impulsion) et l'hamiltonien $H(p,q)= \frac{p^2}{2m} + V(q)$ (énergie). La mécanique quantique comporte un procédé de traduction (quantification) qui remplace $H(p,q)$ par l'opérateur

$$H(P,Q) = \frac{P^2}{2m} + V(Q) \quad \text{(sur les bonnes fonctions)}$$

que l'on espère pouvoir prolonger uniquement en un opérateur a.a. sur $L^2(\mathbb{R})$ (ou $L^2(\mathbb{R}^n)$ si l'on travaille à plusieurs dimensions : alors P est à interpréter comme l'opérateur vectoriel $-i\hbar\,\mathrm{grad}$, et P^2 comme $-\hbar^2\Delta$). Une gigantesque littérature est consacrée à ce problème, sous toutes sortes d'hypothèses sur le potentiel V ... Ce point étant supposé acquis, l'analogue du flot des équations d'Hamilton sur l'espace des (p,q) est constitué par le groupe unitaire $U_t = e^{itH/\hbar}$: si l'état initial de la particule est décrit par une fonction d'onde $\psi_0 \in L^2(\mathbb{R}_n)$, l'état à l'instant t est représenté par $\psi_t = U_t \psi_0$, l'équation d'évolution $\dot{U}_t = iHU_t/\hbar$ se lisant comme l'équation de Schrödinger

$$\dot{\psi}_t = iH\psi_t/\hbar \qquad \text{ou} \qquad -i\hbar\dot{\psi} = -\frac{\hbar^2}{2m}\Delta\psi + V\psi \quad .$$

On recherche la solution comme superposition de solutions stationnaires $\psi(x,t) = e^{itE/\hbar}\phi(x)$, et l'on tombe sur l'équation qui donne les fonctions propres

$$-\frac{\hbar^2}{2m}\Delta\phi + V\phi = E\phi$$

Nous n'allons pas du tout étudier ces problèmes, qui n'ont rien de probabiliste, mais étudier un cas particulier fondamental, celui de l'oscillateur harmonique quantique à une dimension.

Dans le cas de l'oscillateur harmonique classique, le potentiel est $V(x)=kx^2$, et les solutions sont des fonctions sinusoïdales de pulsation ω donnée par $k=m\omega^2$. L'hamiltonien quantique est donc $H(P,Q)= \frac{P^2}{2m} + \frac{m\omega^2}{2}Q^2$. Par des transformations simples, on peut se ramener à une forme réduite, que nous étudierons ici , celle où

$$H(P,Q) = \frac{1}{2}(P^2+Q^2) \quad \text{(on supposera aussi } \hbar=1 \text{) .}$$

Cet opérateur admet il au moins une extension a.a. ? Il existe pour aborder ce problème une méthode probabiliste, fondée sur la formule de Feynman-Kac, et extrêmement puissante. Posons $H_0= \frac{1}{2}P^2$. Alors le semi-groupe du mouvement brownien sur $\mathbb{R}$, que nous désignerons ici par (Π_t), est un semi-groupe de noyaux markoviens, et peut aussi être interprété comme semi-groupe fortement continu de contractions de $L^2(\mathbb{R})$, positives au sens hilbertien. Si nous désignons par $-L_0$, pour un instant, le générateur de ce semi-groupe (au sens précis du terme, avec son domaine naturel), il est

bien connu que L_0 réalise l'extension a.a. cherchée de H_0, et nous ne les distinguerons plus. De plus, on a

$$\Pi_t f(x) = E^x[f(X_t)]$$

où $(X_\cdot)$ est un mouvement brownien issu du point x, ce que rappelle la notation E^x pour l'espérance. Maintenant, si nous posons

$$\Gamma_t f(x) = E^x[f(X_t)\exp(-\int_0^t Vf(X_s)ds)]$$

nous obtenons un semi-groupe de noyaux positifs (sousmarkoviens si $V \geqq 0$). Si ce semi-groupe est fortement continu sur $L^2(\mathbb{R})$ - là se trouve le noeud du problème ! - son générateur $-L$ sera l'extension a.a. désirée de $-L_0-V$. Le cas où $V(x)=\frac{1}{2}x^2$ ne présente aucune difficulté.

Remarquons que H admet la fonction propre normalisée (dans $L^2(dx)$)

$$\phi_0(x) = \pi^{-1/4}\, e^{-x^2/2} \tag{15}$$

avec la valeur propre $1/2$. L'opérateur

$$N = H - \tfrac{1}{2}I \tag{16}$$

qui admet ϕ_0 comme vecteur propre avec la v.p. 0, sera dans la suite plus important que H. La fonction (15), qui correspond à la borne inférieure du spectre de H, est appelée état fondamental de l'oscillateur harmonique (ground state). On notera qu'elle est réelle et partout $\neq 0$.

La connaissance de ϕ_0 nous permet de voir d'une autre manière que H admet une extension a.a.. Soit γ la mesure gaussienne $\phi_0(x)^2dx$; alors la transformation $f \longmapsto f/\phi_0$ est un isomorphisme de $L^2(\mathbb{R})$ sur $L^2(\gamma)$. Si l'on travaille sur ce dernier espace comme espace de Hilbert fondamental (" ground state transformation "), Qf se lit encore xf, Pf devient $-i(Df-xf)$, Hf devient $-\frac{1}{2}D^2f + xDf + \frac{f}{2}$, Nf devient $-(\frac{1}{2}D^2-xD)f$. On reconnaît là l'opérateur d'Ornstein-Uhlenbeck changé de signe, de sorte que e^{-tN} est bien un semi-groupe markovien, et que N admet bien une extension autoadjointe. Cette méthode peut, elle aussi, être considérablement généralisée.

Nous reprendrons plus loin l'étude de l'oscillateur harmonique quantique, qui nous fournira dans un cas concret et simple tous les éléments de la théorie que nous verrons plus tard en dimension infinie : opérateurs de création et d'annihilation, vecteurs cohérents, etc. Pour le moment, revenons à quelques remarques simples sur le couple canonique.

4. Remarques sur le couple canonique.

a) La structure fondamentale de la mécanique classique, sous forme hamiltonienne, est une structure symplectique, qui permet d'écrire les équations d'Hamilton sous forme intrinsèque : sous la forme très élémentaire que nous rencontrons ici, cela se manifeste par des notations un peu

différentes des nôtres, que l'on rencontre fréquemment dans la littérature.

On peut munir l'espace $\mathbb{R}^2=\mathbb{C}$ ($p+iq=z$) de la forme alternée

$$\sigma(p,q\ ;\ p',q') = \begin{pmatrix} p & p' \\ q & q' \end{pmatrix} = pq'-qp' = \mathrm{Im}\langle z,z'\rangle$$

La définition correspondante de la transformée de Fourier d'une mesure μ (par exemple : ce pourrait être une distribution) est

$$\hat{\mu}(r,s) = \int \exp\left(i\begin{pmatrix} r & p \\ s & q \end{pmatrix}\right)\mu(dp,dq) \tag{17}$$

et celle des opérateurs de Weyl est $W_{p,q}=e^{i(qP-pQ)}$ (de sorte que la correspondance avec la notation antérieure est $r=q$, $s=-p$). La relation de commutation est inaltérée : $W_{p,q}W_{p',q'} = \exp\left(i\frac{\hbar}{2}\begin{vmatrix} p & p' \\ q & q' \end{vmatrix}\right)W_{p+p',q+q'}$.

Le caractère naturel de ces notations apparaît bien si l'on remarque que les transformations linéaires (dites <u>canoniques</u>)

$$P' = aP + bQ \quad , \quad Q' = cP + dQ$$

qui préservent la relation $[P,Q]=i\hbar I$ (ou les relations de Weyl) sont les transformations unimodulaires ($ad-bc=1$), i.e. celles qui préservent la forme σ .

b) Considérons une famille $(\tilde{W}_{rs})$ d'opérateurs unitaires sur un espace Ω, satisfaisant aux relations de commutation de Weyl, mais <u>non nécessairement irréductible</u>. D'après le théorème de Stone-von Neumann, Ω est somme-directe de sous-espaces stables Ω_n, qui sont des copies du modèle, et que nous identifierons à celui-ci. Si $\omega\in\Omega$ est un vecteur normalisé, nous pouvons le décomposer en $\omega=\Sigma_n\ c_n\omega_n$, chaque ω_n étant un vecteur normalisé de Ω_n, avec $\Sigma_n\ |c_n|^2=1$. On a alors

$$\langle \omega,\tilde{W}_{rs}\omega \rangle = \Sigma_n\ |c_n|^2 \langle \omega_n,W_{rs}\omega_n \rangle$$

Du côté droit, nous avons la fonction caractéristique d'une loi quantique sur le modèle, non pure en général : le mélange $\Sigma_n\ |c_n|^2\varepsilon_{\omega_n}$. Du côté gauche, nous avons $E[e^{i(r\tilde{P}+s\tilde{Q})}]$ sous la loi ε_ω sur Ω . Cela illustre bien la différence entre lois associées à un vecteur et lois pures, et montre aussi comment on peut utiliser des couples canoniques non irréductibles pour construire des lois sur le modèle.

> En dimension infinie, la notion de modèle unique, fournie par le théorème de Stone-von Neumann, fera défaut, et l'on devra parler plutôt de lois de probabilité sur la C^*-algèbre de Weyl.
>
> On notera aussi l'analogie entre cette construction, et la construction de processus canoniques en probabilités classiques : construire un processus continu (par ex.) sur un espace probabilisé auxiliaire plus riche, et se ramener au processus des coordonnées sur $C(\mathbb{R}_+)$ en prenant une mesure image.

Illustrons ce procédé. Soient (P_0,Q_0) et (P_1,Q_1) deux couples canoniques qui <u>commutent</u> . Par exemple, on peut les réaliser sur $L^2(\mathbb{R}^2)$, les opérateurs Q_j $(j=0,1)$ étant les multiplications par les coordonnées x_j , et

les P_j les dérivées partielles $i\hbar D_j$. Posons

$$\tilde{Q} = aQ_0 + bQ_1 \quad , \quad \tilde{P} = aP_0 - bP_1 \qquad (a,b \text{ réels})$$

Un calcul formel montre que $[\tilde{P},\tilde{Q}] = (a^2-b^2)i\hbar I$, et il est donc naturel de prendre $a^2-b^2=1$ (nous aurions pu considérer des coefficients plus généraux, mais ce ne sera pas utile). Les opérateurs (P,Q) forment alors un couple canonique, dont les opérateurs de Weyl sont

$$\tilde{W}_{rs}f(x_0,x_1) = e^{i\hbar rs/2}\, e^{is(ax_0+bx_1)}f(x_0+a\hbar r,\ x_1-b\hbar r)$$

et ce couple n'est pas irréductible (par exemple, le sous-espace des f nulles p.p. sur un ensemble de la forme $\{bx_0+ax_1 \in A\}$ est stable par les $\tilde{W}_{rs}$).

Soient ω_0 et ω_1 deux vecteurs normalisés du modèle, et soit ω le vecteur normalisé $\omega_0 \otimes \omega_1$ (autrement dit, $\omega_0(x_0)\omega_1(x_1)$). Nous avons

$$< \omega, \tilde{W}_{rs}\omega > = < \omega_0, W_{ar,as}\omega_0 > < \omega_1, W_{br,-bs}\omega_1 >$$

Si nous prenons en particulier pour ω_j deux copies du vecteur d'incertitude minimale (cf. (13),(14)) nous obtenons à gauche une fonction caractéristique de la forme

$$(17) \qquad \exp(-\tfrac{1}{2}(\lambda r^2+\mu s^2)) \qquad \text{avec} \quad \lambda=(a^2+b^2)c_p \ , \ \mu=(a^2+b^2)c_q$$

le coefficient a^2+b^2 étant >1 dès que $b\neq 0$. On a donc construit par ce procédé des <u>lois gaussiennes quantiques qui ne sont pas d'incertitude minimale</u> (on peut encore " faire tourner les axes " par une transformation canonique du type considéré en a)).

Ces lois gaussiennes quantiques forment exactement la classe des lois limites dans la généralisation naturelle du théorème limite central élémentaire (v.a. indépendantes équidistribuées avec moment du second ordre) Voir Cushen et Hudson, A Quantum-Mechanical Central Limit Theorem, J. Appl. Prob. 8, 1971, p. 454-469.

> On utilisera plus tard exactement le même procédé pour construire, non pas des v.a. gaussiennes quantiques, mais des mouvements browniens quantiques. Mais il y aura une différence importante : on ne pourra plus ramener la loi sur le modèle de départ (cf. les lois gaussiennes classiques : les lois de deux mouvements browniens de variances différentes sont étrangères).

5. <u>Opérateurs de création et d'annihilation</u>.

Posons, sur les bonnes fonctions (nous travaillons sur le modèle)

$$(18) \qquad a^- = \tfrac{1}{\sqrt{2}}(Q+iP) \quad , \quad a^+ = \tfrac{1}{\sqrt{2}}(Q-iP)$$

(- pour annihilation, + pour création). Ces opérateurs non bornés ne sont pas a.a., ni même normaux. On a pour $f,g \in \underline{S}$

$$< f, a^- g > = < a^+ f, g >$$

de sorte que $(a^\pm)^*$ a un domaine dense, et est une extension fermée de $a^\mp$. On peut montrer (mais nous n'en aurons pas besoin) que la fermeture de a^+ est exactement l'adjoint de a^- , et inversement - d'où les notations a, a^* fréquemment utilisées pour a^-, a^+.

Cette section est une préparation à la théorie de l'espace de Fock. Cependant, nous utilisons des normalisations différentes de celles qui nous seront commodes en dimension infinie. Nous prenons ici $\hbar=1$, là bas $\hbar=2$, et nous aurons de plus un choix de normes différent sur les sous-espaces propres. Le lecteur trouvera donc quelques différences dans les formules.

a) La présentation "algébrique" de la théorie de l'oscillateur harmonique (et plus généralement du couple canonique) que nous allons donner maintenant, figure dans tous les livres de mécanique quantique. Elle remonte à Dirac, semble t'il.

Nous avons d'abord les formules suivantes (sur le domaine dense $\underline{\underline{S}}$, stable par a^+ et a^-)

$$[a^-, a^+] = I$$

$$a^+ a^- = N = \frac{1}{2}(P^2+Q^2-I) \tag{19}$$

$$Na^- = a^-(N-I) \ , \ Na^+ = a^+(N+I)$$

L'écriture $N=a^* a$ suggère fortement que N , convenablement étendu, est a.a. positif, ce qui sera confirmé. La troisième formule montre que, si h est un vecteur propre de N (appartenant à $\underline{\underline{S}}$) avec valeur propre λ, $a^\pm h$ est encore vecteur propre, avec valeur propre $\lambda \pm 1$. Nous saurons donc fabriquer toute une échelle de vecteurs propres, à condition d'en avoir un seul. Celui-ci nous est fourni par le vecteur déjà vu en (15)

$$\phi_o(x) = \pi^{-1/4} e^{-x^2/2} \tag{20}$$

pour lequel on a $\phi_o'(x) = -x\phi(x)$, donc $a^-\phi_o = 0$. A partir de là, nous engendrons les fonctions propres de N (appartenant à $\underline{\underline{S}}$)

$$\phi_n = a^{+n}\phi_o \text{ (non normalisées)} \ , \ h_n = c_n \phi_n \text{ (normalisées)}. \tag{20}$$

Le vecteur ϕ_o jouera un rôle fondamental dans la suite : on l'appelle le vide.

Aucun des vecteurs ϕ_n n'est nul. En effet, la relation $a^+\phi = 0$ entraîne $0 = <a^+\phi, a^+\phi> = < \phi, a^- a^+ \phi > = < \phi, (I+a^+a^-)\phi > = < \phi, \phi > + < a^-\phi, a^-\phi >$, donc $<\phi,\phi>=0$. Déterminons les constantes c_n .

$$c_{n+1}^{-2} = < \phi_{n+1}, \phi_{n+1} > = < a^+\phi_n, a^+\phi_n > = < a^- a^+ \phi_n, \phi_n > =$$

$$= < (N+I)\phi_n, \phi_n > = (n+1)<\phi_n, \phi_n> = (n+1)c_n^{-2}$$

D'où (à des facteurs de module 1 près)

$$c_n = (n!)^{-1/2} \qquad h_n = (n!)^{-1/2} a^{+n} \phi_0$$

et la matrice de a^+ et de a^- dans la base orthonormale (h_n) (nous ne savons pas encore que c'est une base !) : on en déduira les matrices de P et Q dans cette même base, dont l'expression remonte au tout premier travail de Heisenberg sur la << mécanique des matrices >>

$$(21) \qquad a^+ h_n = \sqrt{n+1}\, h_{n+1} \ , \quad a^- h_n = \sqrt{n}\, h_{n-1} \quad (a^- h_0 = 0) \ .$$

On peut calculer explicitement les fonctions h_n sur le modèle : nous préférons travailler sur le modèle gaussien vu au n°3 , dans lequel l'espace de Hilbert est $L^2(\gamma)$, Q et P sont représentés par x, $-i(D-x)$, $a^- = D/\sqrt{2}$, $a^+ = (2x-D)/\sqrt{2}$, et $\phi_0 = 1$. Alors

$$(22) \qquad h_n(x) = (2^n n!)^{-1/2} (2x-D)^n 1 = (2^n n!)^{-1/2} H_n(x)$$

où les $H_n(x)$ sont les polynômes d'Hermite sous la forme usuelle en analyse (non en probabilités), de série génératrice

$$(23) \qquad \sum_n \frac{t^n}{n!} H_n(x) = e^{2tx-t^2} \ .$$

Dans le modèle usuel sur $L^2(\mathbb{R})$, il y aurait un facteur $\phi_0(x)$ à droite de (22). Si nous avions pris $\hbar = 2$, la loi gaussienne γ serait réduite, et nous aurions les polynômes des probabilistes, quelques $\sqrt{2}$ disparaissant.

Dans le modèle usuel, l'espace de Hilbert $\mathcal{H}$ engendré par les h_n (ou encore par les $K(x)\phi_0(x)$, où $K(x)$ est un polynôme) est stable par translation (la formule (23) nous dit comment calculer $\phi_0(x-t)$), c. à d. stable par le groupe $U_t = e^{itP}$. On vérifie sans peine qu'il est stable par le groupe $V_t = e^{itQ}$. D'après l'irréductibilité du modèle, c'est $L^2(\mathbb{R})$ entier. On en déduit que les h_n forment une base orthonormale, qui est une base de vecteurs propres pour N . Cela nous donne la représentation spectrale de N (ou plutôt, de son extension naturelle comme opérateur a.a. positif, ou comme v.a. à valeurs entières positives)

$$(24) \qquad N = \Sigma_n \, n E_n \qquad \text{où } E_n \text{ est le projecteur sur } \mathbb{C} h_n \ .$$

REMARQUE. Nous avons déjà vu deux réalisations du couple canonique : la première, celle de Schrödinger, sur $L^2(\mathbb{R})$; la seconde, liée à l'oscillateur harmonique, sur un $L^2(\gamma)$ gaussien. Nous venons d'en voir une troisième, sur l'espace ℓ^2 (correspondant à l'observable N), les opérateurs P,Q étant représentés comme les matrices de Heisenberg.

6. Vecteurs cohérents. Nous allons introduire maintenant une notion qui jouera un grand rôle en dimension infinie (mais encore une fois, les normalisations ne se correspondent pas tout à fait).

Nous associons à tout nombre complexe z le vecteur (dit exponentiel, ou cohérent

$$\mathcal{E}(z) = \Sigma_n \left(\frac{z}{\sqrt{2}}\right)^n \frac{\phi_n}{n!} = \Sigma_n \frac{z^n}{\sqrt{2^n n!}} h_n \tag{25}$$

(le coefficient $\sqrt{2}$ sera justifié un peu plus bas : il provient de la normalisation). Le vecteur unitaire associé est

$$\tilde{z} = e^{-|z|^2/4} \mathcal{E}(z) . \tag{26}$$

Sous la loi $\varepsilon_{\tilde{z}}$, l'observable N admet une loi de Poisson de moyenne $|z|^2/2$. Quelle est la loi de l'observable Q ? La fonction génératrice (23) nous donne (pour t complexe aussi)

$$\phi_o(x)e^{2tx-t^2} = \Sigma_n \frac{t^n 2^{n/2}}{\sqrt{n!}} h_n(x)$$

qui s'identifie à (25)-(26) pour $z=t/2$. Dans le modèle $L^2(\mathbb{R})$, le vecteur $\tilde{z}$ s'écrit donc

$$\tilde{z}(x) = \exp[zx - \tfrac{1}{4}(z^2+|z|^2)]\phi_o(x) \tag{27}$$

et dans le domaine $L^2(\gamma)$, $\tilde{z}$ est une exponentielle normalisée, et $\mathcal{E}(z)$ est simplement l'exponentielle complexe e^{zx} . La densité correspondante est, dans le modèle $L^2(\mathbb{R})$, et en posant $z=u+iv$

$$|\tilde{z}(x)|^2 = \pi^{-1/2} e^{-x^2} \exp[2ux-x^2] \tag{28}$$

qui est une gaussienne de moyenne u (de là le $\sqrt{2}$ de (25) !). Autrement dit, l'utilisation des vecteurs cohérents correspond un peu à une formule de Cameron-Martin, permettant de passer d'un mouvement brownien à un mouvement brownien avec dérive non nulle.

On appelle états classiques les mélanges

$$\tilde{\mu} = \int \varepsilon_{\tilde{z}} \mu(dz) \quad (\text{où } \mu \text{ est une loi sur } \mathbb{C}) . \tag{29}$$

D'après le livre de Davies, Quantum Theory of Open Systems, ces mélanges sont fort importants en optique quantique. Nous nous en servirons ici pour donner une construction des lois gaussiennes quantiques d'incertitude non minimale, plus directe que celle du n°4 . L'opérateur de densité W (opérateur positif de trace 1) associé au mélange est donné par

$$< \omega, W\psi > = \int <\omega,\tilde{z}><\tilde{z},\psi> \mu(dz) \tag{30}$$

d'où sa matrice $(W_{nm}) = < h_n, Wh_m >$

$$W_{nm} = (2^n n!\, 2^m m!)^{-1/2} \int \bar{z}^n z^m e^{-|z|^2/2} \mu(dz) .$$

Si μ est invariante par rotation, W est diagonale (i.e., est une fonction de l'observable N). Nous remarquons que dans tous les cas, il est facile de déterminer la loi de l'observable N par sa fonction génératrice $E_W[\lambda^N]$. En effet, sous $\varepsilon_{\tilde{z}}$ cette fonction génératrice est $e^{(\lambda-1)z^2/2}$ (loi de

Poisson), donc sous la loi (29) on a

$$E[\lambda^N] = \int e^{(\lambda-1)|z|^2/2}\,\mu(dz) \quad . \tag{31}$$

Enfin, calculons la "fonction caractéristique" jointe de P et Q sous la loi (29) : il suffit de la connaître sous $\varepsilon_{\tilde{z}}$ et d'intégrer en $\mu(dz)$

$$E_{\tilde{z}}[e^{i(rP+sQ)}] = <\tilde{z}, W_{rs}\tilde{z}> = e^{-(r^2+s^2)/4}\, e^{i(rv+su)} \tag{32}$$

d'après (8) et (27). Si z=u+iv , on trouve une " loi jointe " gaussienne de variance minimale, mais <u>de moyenne</u> (u,v). On en déduit que pour tout état classique, la fonction caractéristique est de type positif.

Nous allons appliquer ces calculs au cas où μ est elle même une loi gaussienne invariante par rotation

$$\mu(dz) = (2\pi)^{-1} a e^{-a|z|^2/2}\, dudv \qquad (\ z = u+iv\ )$$

Un calcul élémentaire, que nous ne ferons pas, montre que la loi de Q ou de P est gaussienne, de moyenne 0 et de variance $(a+2)/2a$: on retrouve donc les couples gaussiens quantiques d'incertitude non minimale (par une méthode qui ne s'étend pas en dimension infinie). Il est plus instructif de rechercher la loi de N . On a d'après (31)

$$E_W[\lambda^N] = \frac{a}{a+1-\lambda} = \frac{1-b}{1-\lambda b} \quad \text{pour} \quad b=1/1+a$$

Donc $P\{N=n\} = (1-b)b^n$, loi géométrique, et enfin

$$W = e^{-cN}/Tr(e^{-cN}) \qquad \text{avec} \quad b=e^{-2c} \quad .$$

Les probabilistes connaissent bien le <u>semi-groupe d'Ornstein-Uhlenbeck</u> e^{-tN}. L'opérateur de densité W est un élément de ce semi-groupe, normalisé pour lui donner une trace unité. L'interprétation physique est la suivante : lorsque l'on a affaire à un système physique d'hamiltonien H, il est admis que l'opérateur $e^{-H/kT}$ normalisé (si possible) pour avoir la trace 1 représente un état d'équilibre statistique d'un grand nombre de copies du système à la température T, k étant la constante de Boltzmann. L'hamiltonien H de l'oscillateur harmonique ne différant de N que par une constante, celle-ci disparaît par normalisation, et les lois gaussiennes du couple canonique apparaissent comme des états d'équilibre statistique de l'oscillateur, l'incertitude minimale ayant lieu pour T=0.

La partie de cet exposé qu'il est recommandé de lire avant l'exposé IV (espace de Fock) est terminée. Les paragraphes suivants ne doivent être considérés que comme des compléments.

III. FONCTIONS DE WIGNER

La théorie présentée dans ce paragraphe est un thème, d'importance assez mineure, qui se fait entendre depuis les débuts de la mécanique quantique. On ne le trouve pas couramment exposé dans les livres, soit parce qu'il n'est pas très utile, soit peut être parce qu'il est jugé un peu "hétérodoxe" par rapport à la philosophie générale des opérateurs. Pourtant ses créateurs (H. Weyl 1931, E. Wigner, Phys. Rev. 40, 1932) ne sont pas des marginaux ! Ce qui est attirant pour un probabiliste, c'est d'abord l'analogie avec des êtres bien connus, ensuite l'apparition de " lois jointes" non nécessairement positives, alors que celles-ci présentent aussi un certain intérêt en analyse.

Les références suivantes sont excellentes : Pool, J. Math. Phys, 7, 1966. Cushen-Hudson, J. Appl. Prob. 8, 1971. Combe-Guerra-Rodriguez-Sirugue-S. Collin, Proc. VII Int. Congr. Math. Phys., Physica 124A, 1984. Pour les physiciens, les exposés classiques sont : Moyal, Proc. Cambridge Phil. Soc. 45, 1949. Fano, Rev. Mod. Phys. 29, 1957. Baker, Phys. Rev. 109, 1958. Je les trouve difficiles à lire.

1. Nous allons travailler sur le couple canonique à une dimension, donc sur $\mathbb{R}^2$. Précisons nos notations pour la transformation de Fourier : si $\mu(dx,dy)$ est une mesure, sa transformée de Fourier (notée $\mathcal{F}\mu$ ou $\hat{\mu}$) est

$$\hat{\mu}(u,v) = \int e^{i(ux+vy)}\mu(dx,dy)$$

(on pourrait préférer la forme symplectique uy-vx). Pour avoir une belle formule d'inversion, nous identifions une fonction f(x) ou f(x,y) à la mesure $f(x)\not{d}x$ ou $f(x,y)\not{d}x\not{d}y$, où $\not{d}=d/\sqrt{2\pi}$. Alors pour $f\in\underline{S}$ la formule d'inversion est simplement

$$\hat{f}(u,v) = \int e^{i(ux+vy)}f(x,y)\not{d}x\not{d}y \quad ; \quad f(x,y) = \int e^{-i(ux+vy)}\hat{f}(u,v)\not{d}u\not{d}v$$

et la norme L^2 est préservée sans aucun coefficient. On peut aussi définir les transformations de Fourier partielles $\mathcal{F}_1$ et $\mathcal{F}_2$.

Remettons aussi sous les yeux du lecteur les définitions relatives aux opérateurs de Weyl. Si l'on s'est donné une loi quantique, d'opérateur de densité ρ , sa <u>fonction caractéristique</u> est

$$F(r,s) = E[e^{i(rP+sQ)}] = \mathrm{Tr}(\rho W_{r,s}) \quad . \tag{1}$$

Nous rappelons les relations de Weyl, sous leur forme réelle et complexe (celle-ci nous servira plus loin)

$$W_{rs}W_{r's'} = e^{i\hbar(rs'-sr')/2}\,W_{r+r',s+s'} \quad ; \quad W_zW_{z'} = e^{i\hbar\mathrm{Im}\langle z,z'\rangle/2}\,W_{z+z'}. \tag{2}$$

Nous prenons $\hbar=1$ dans toute la suite. Sous une loi pure ω, nous avons

$$(3) \qquad F_\omega(r,s) = \int \overline{\omega}(x)\omega(x+r)e^{irs/2}e^{isx}dx = \int\overline{\omega}(x-\tfrac{r}{2})\omega(x+\tfrac{r}{2})e^{isx}dx$$

Nous allons commencer par donner un sens plus large à cette formule. Soit $K(x,y)\in\underline{\underline{S}}(\mathbb{R}\times\mathbb{R})$. Nous posons, en suivant Pool :

$$TK(x,y) = K(x-\tfrac{y}{2},x+\tfrac{y}{2})$$

T est une bijection de $\underline{\underline{S}}$ sur $\underline{\underline{S}}$, qui est aussi unitaire. Ensuite, nous prenons une transformée de Fourier par rapport à la première variable

$$(4) \qquad \mathcal{F}_1 TK(s,y) = \int K(x-\tfrac{y}{2},x+\tfrac{y}{2})e^{isx}dx$$

et nous obtenons encore un isomorphisme unitaire de $\underline{\underline{S}}$ sur $\underline{\underline{S}}$, qui se prolonge en un isomorphisme de $L^2(\mathbb{R}^2)$ sur lui même. Revenant à (3), nous voyons que nous avons simplement appliqué cet isomorphisme à $K(x,y)=\overline{\omega}(x)\omega(y)$. D'où un résultat dont on se doutait bien, mais qui est devenu évident : $F_\omega(.,.)$ <u>détermine uniquement</u> $\overline{\omega}\otimes\omega$, <u>donc aussi</u> ω <u>à un facteur de module 1 près</u>.

Seconde propriété : puisque $F_\omega(r,s)$ est une "fonction caractéristique", il est naturel de se demander si elle est la transformée de Fourier de quelque chose, cet objet devant jouer le rôle de "loi jointe" pour (P,Q). Or la réponse est claire sur (4) : $\mathcal{F}_1 TK = (\mathcal{F}_1\mathcal{F}_2)(\mathcal{F}_2^{-1}TK)$, cette dernière parenthèse étant l'objet cherché. Comme la transformée de Fourier $\mathcal{F}=\mathcal{F}_1\mathcal{F}_2$ est un isomorphisme de $L^2(\mathbb{R}^2)$ sur lui même, l'"objet" est encore un élément de L^2. Dans le cas particulier de (3), c'est la <u>fonction de Wigner</u>

$$(5) \qquad f_\omega(p,q) = \int F_\omega(r,s)e^{-i(rp+sq)}drds = \int\overline{\omega}(q-\tfrac{y}{2})\omega(q+\tfrac{y}{2})e^{-ipy}dy \quad (1)$$

Cette fonction est réelle, mais non nécessairement positive. C'est pourquoi l'on dit souvent que les fonctions de Wigner définissent des "probabilités négatives" . Sauf erreur de ma part, il y a une différence plus profonde avec les probabilités ordinaires : la fonction de Wigner est dans L^2, mais non nécessairement dans L^1 (quelles hypothèses sur ω faut il pour cela ?).

Il est clair d'après la discussion ci-dessus que la correspondance que nous avons établie entre les "fonctions d'ondes" ω et leurs fonctions caractéristiques $F_\omega(r.s)$ ou leurs fonctions de Wigner $f_\omega(p,q)$ s'étend à tous les noyaux $K(x,y)\in L^2(\mathbb{R}\times\mathbb{R})$ (i.e. aux opérateurs de H-S sur $L^2(\mathbb{R})$)

$$(6) \qquad F_K(r,s) = \int K(x-\tfrac{r}{2},x+\tfrac{r}{2})e^{isx}dx \quad , \quad f_\omega(p,q) = \int K(q-\tfrac{y}{2},q+\tfrac{y}{2})e^{-ipy}dy$$

(intégrales au sens de Plancherel), les applications $K\longmapsto F_K$, f_K étant des isomorphismes de $L^2(\mathbb{R}\times\mathbb{R})$.

Avant de poursuivre la discussion mathématique, il faudrait indiquer pourquoi H. Weyl a été cité parmi les pères de ce sujet. Le problème qui

1. Intégrales au sens de Plancherel (la première) et de Lebesgue (la 2e).

préoccupe Weyl est de donner une règle précise permettant d'associer, à une observable classique f(p,q), une observable quantique f(P,Q) - ce n'est pas évident, même pour un polynôme, en raison de la non-commutativité ! La règle de Weyl consiste à écrire formellement

$$f(P,Q) = \int e^{-(rP+sQ)}\ \hat{f}(r,s)\,dr\,ds$$

$\hat{f}(r,s)$ est en général une distribution. Ce problème et celui des fonctions de Wigner sont souvent présentés ensemble, mais nous ne nous y intéressons pas ici (je n'y connais rien).

2. Dans ce n°, nous indiquons quelques propriétés de la fonction caractéristique d'une loi quantique (correspondant à l'opérateur de densité ρ) . Il sera commode de considérer (r,s) comme l'élément z=r+is de l'espace de Hilbert complexe $\mathbb{C}$!

a) La première remarque est que $F(0)=1$, $|F(.)| \leq 1$.

b) Etablissons la <u>continuité</u> de F(.). Nous traitons d'abord le cas de $F_\omega(z)$. On remarque que $s \longmapsto \omega(.+s)$ est continue dans L^2, donc $s \longmapsto \overline{\omega}(.)\omega(.+s)$ est continue dans L^1, et le résultat se voit sur la première expression (3).

La continuité, établie pour les lois pures ε_ω , s'étend alors aux mélanges par convergence dominée.

c) Nous avons ensuite la proposition suivante, qui rappelle le théorème de Bochner :

THEOREME 1. <u>Pour qu'une fonction</u> F(z) <u>sur</u> $\mathbb{C}$ <u>soit la fonction caractéristique d'une loi quantique, il faut et il suffit qu'elle soit continue</u>, <u>et que la fonction sur</u> $\mathbb{C}\times\mathbb{C}$

$$\Phi(z,z') = F(z'-z)e^{-i\mathrm{Im}\langle z,z'\rangle/2} \tag{7}$$

<u>soit un noyau de type positif</u> (i.e. les formes hermitiennes $\Sigma_{jk}\ \overline{\lambda}_j\lambda_k\Phi(z_j,z_k)$ sont positives , pour tout choix des $z_j \in \mathbb{C}$ en nombre fini).

<u>Démonstration</u>. Pour la nécessité, on écrit que l'opérateur a.a. positif $(\Sigma_j\ \lambda_j W_{z_j})^*(\Sigma_k\ \lambda_k W_{z_k})$ a une espérance positive, et on utilise la relation de Weyl sous la forme complexe rappelée en (2) plus haut . Pour la suffisance, on considère l'espace autoreproduisant associé à Φ, i.e. l'espace engendré par les masses unité δ_z avec le produit hermitien $\langle \delta_z, \delta_{z'}\rangle = \Phi(z,z')$, convenablement séparé et complété pour définir un Hilbert complexe H . On définit des opérateurs unitaires W_u ($u \in \mathbb{C}$) sur H par

$$W_u\delta_z = e^{i\mathrm{Im}\langle u,z\rangle/2}\ \delta_{u+z}$$

et l'on constate que ces opérateurs satisfont aux relations de Weyl (tout ceci est très facile). On a aussi $\langle \delta_z, W_u\delta_{z'}\rangle = e^{i\mathrm{Im}(\ldots)/2}F(u+z'-z)$, d'où la continuité de $u \longmapsto W_u$ pour la topologie faible des opérateurs, et pour z,z'=0, l'identification de F(u) à $\langle \delta_0, W_u\delta_0\rangle$. Le sous-espace stable

engendré par δ_0 contient les δ_z , donc il est dense, et H est séparable. Nous avons donc réalisé sur H une représentation des relations de Weyl, à laquelle on peut appliquer le th. de Stone-von Neumann, à un détail près : nous avons présenté celui-ci sous une hypothèse de continuité forte, et non faible, des deux groupes unitaires : il est tout à fait classique en analyse fonctionnelle que pour des semi-groupes de contractions (a fortiori des groupes unitaires) continuité forte et faible sont équivalentes.

D'après le th. de S-vN, H se décompose en une somme directe de copies du modèle. Par le même raisonnement qu'au paragraphe précédent, 4 b), cela revient à munir le modèle d'une loi quantique qui est un mélange, et la fonction F(u) est alors interprétée comme fonction caractéristique d'une telle loi. Nous empruntons à Cushen-Hudson une jolie conséquence du th. 1 .

COROLLAIRE. Si F et G sont deux fonctions caractéristiques (au sens quantique) leur produit FG est une fonction caractéristique au sens classique.

En effet, soient $\Phi(z,z')$ et $\Psi(z,z')$ les noyaux de type positif (7) correspondants : $\overline{\Psi}$ est aussi de type positif , et aussi le produit $\Phi\overline{\Psi}$ (résultat classique sur les n.t.p.). La multiplication enlève les exponentielles, et il ne reste que $FG(z'-z)$.

d) Nous n'avons pas encore montré que la connaissance de la fonction caractéristique $F_\rho(r,s)$ de la loi quantique ρ détermine celle-ci. En voici la démonstration la plus naturelle : connaître F_ρ détermine $\mathrm{Tr}[\rho W_z]$ pour tout z . Or les combinaisons linéaires des W_z sont denses dans l'espace $\mathfrak{L}$ de tous les opérateurs bornés sur $L^2(\mathbb{R})$, considéré comme dual de l'espace des opérateurs à trace, et muni de sa topologie faible : cela revient[1] à vérifier que le commutant des W_z est réduit aux multiples de l'identité, et signifie exactement que le couple canonique est irréductible (cela sera vu plus tard). Donc F_ρ détermine $\mathrm{Tr}(\rho a)$ pour tout opérateur borné a, et donc ρ.

Voici une démonstration plus élémentaire, qui rejoint l'idée de l'exposé I, fin du n°5 sur le rôle des opérateurs de H-S comme extensions naturelles des fonctions d'onde. Choisissons une base orthonormale (ω_i) dans lequel ρ soit diagonalisé ($\rho = \Sigma_i \lambda_i |\omega_i \rangle\langle \omega_i|$ en notation de Dirac, avec des λ_i positifs de somme 1). Alors ρ est représenté par un noyau

$$\rho f(x) = \int K(x,y) f(y)\,dy \quad \text{avec} \quad K(x,y) = \Sigma_i \lambda_i \overline{\omega}_i(x)\omega_i(y) \in L^2(\mathbb{R}\times\mathbb{R})$$

On a d'autre part

$$F_\rho(r,s) = \Sigma_i \lambda_i F_{\omega_i}(r,s) = \Sigma_i \lambda_i \int \overline{\omega}_i(x-\tfrac{r}{2})\omega_i(x+\tfrac{r}{2})e^{isx}\,dx$$

On reconnaît là, appliqué au noyau K , l'opérateur (4) : celui-ci étant bijectif, la connaissance de F_ρ détermine le noyau K, et donc ρ lui-même.

---

1. D'après le th. de densité de von Neumann.

e) Nous pouvons maintenant démontrer très simplement le très joli théorème de continuité du type de Lévy, établi par Cushen et Hudson :

THEOREME. Soit (ρ_n) une suite de lois quantiques pour le couple canonique, dont les fonctions caractéristiques $f_n(r,s)$ convergent simplement vers une fonction $f(r,s)$ continue en 0. Alors f est la fonction caractéristique d'une loi ρ, et les ρ_n convergent étroitement vers ρ (i.e. $tr(\rho_n a)$ tend vers $tr(\rho a)$ pour tout opérateur borné a).

Démonstration. D'après l'appendice à l'exposé I, nous savons que l'on peut se ramener par compacité au cas où les ρ_n convergent faiblement dans l'espace HS des opérateurs de Hilbert-Schmidt, vers un opérateur ρ, qui est positif et de trace ≤ 1 . Nous savons aussi que la convergence étroite équivaut à la propriété $tr(\rho)=1$.

D'après les résultats de Pool rappelés plus haut, les $f_n(r,s)$ convergent faiblement dans $L^2(\mathbb{R}\times\mathbb{R})$ vers la fonction caractéristique $F_\rho(r,s)$. Comme ils sont uniformément bornés et convergent simplement vers $f(r,s)$, on a $F_\rho(r,s)=f(r,s)$ p.p.. Comme f est supposée continue en 0, et que F_ρ l'est aussi, on a $F_\rho(0) = f(0) = 1$, et le théorème est établi. La démonstration est plus simple que celle du théorème de Lévy, parce que nous travaillons a priori sur une classe de lois plus restreinte (les lois sur $\mathbb{R}$ ne peuvent pas toutes se relever sur le couple canonique quantique !)

Soit $g(r,s)=\exp(-\frac{1}{4}(r^2+s^2))$: les gf_n sont des fonctions caractéristiques ordinaires, et convergent simplement vers gf continue à l'origine. D'après le théorème de Lévy classique, on a convergence uniforme sur tout compact. Divisant par g , on obtient le même résultat pour les f_n elles mêmes, et en particulier la continuité de f partout, et la relation $f=F_\rho$ partout.

Cushen et Hudson utilisent ce résultat pour donner une forme simple du théorème limite central pour les lois quantiques, exactement comme dans le cas classique. Nous renvoyons à leur article pour les détails.

REMARQUES. 1) Il y a un véritable intérêt mathématique à utiliser le formalisme symplectique dans ces questions. Voir la théorie de la "convolution gauche" dans Loupias et Miracle-Sole, Comm. M. Phys. 2, 1966 (après D.Kastler, Comm. M. Phys. 1).

2) Soit Π l'opérateur de parité sur $\mathbb{R}$ ($\Pi f(x)=f(-x)$). Une petite manipulation de la formule (5) montre que $f_\omega(p,q) = 2\langle\omega, W_{2p,-2q}\Pi\omega\rangle$ (cette formule, due à Combe et al., article cité, est plus jolie en formalisme symplectique) On l'étend aux mélanges : en dimension 1, et avec $\hbar=1$ toujours

$$F_\omega(r,s) = tr(\rho W_{r,s}) \quad , \quad f_\omega(r,s) = 2tr(\rho W_{2p,-2q}\Pi) \quad .$$

3) Moyal et d'autres se préoccupent de traduire l'équation de Schrödinger en équation d'évolution des fonctions de Wigner : Combe et al. ont souligné les analogies frappantes avec une équation de Kolmogorov .

ELEMENTS DE PROBABILITES QUANTIQUES. IV

Probabilités sur l'espace de Fock

Cet exposé est le plus important de la série ; il commence par des préliminaires un peu ennuyeux, où l'on introduit des objets de nature algébrique : les espaces de Fock symétrique et antisymétrique. Puis l'on donne l'interprétation probabiliste de l'espace de Fock symétrique : c'est l'espace L^2 du mouvement brownien. Ensuite, on s'aperçoit (en suivant Hudson et Parthasarathy[1]) que l'espace de Fock a une structure probabiliste extraordinairement riche : non seulement on peut y voir un mouvement brownien, mais beaucoup de mouvements browniens qui ne commutent pas, et des processus de Poisson, et bien d'autres choses. On retrouve ici tous les ingrédients rassemblés dans les exposés précédents.

I. ESPACE DE FOCK

1. Soient $\mathcal{H}$ et $\mathcal{K}$ deux espaces de Hilbert (complexes ou réels). Nous allons utiliser dans ce paragraphe un certain nombre de résultats triviaux concernant le produit tensoriel algébrique $\mathcal{H}\underline{\otimes}\mathcal{K}$ et son complété le produit tensoriel hilbertien $\mathcal{H}\hat{\otimes}\mathcal{K}$, mais cela n'exige aucune connaissance en algèbre. En pratique, nos espaces de Hilbert seront toujours des espaces $L^2(H,\lambda)$ et $L^2(K,\mu)$, et dans ce cas $\mathcal{H}\hat{\otimes}\mathcal{K}$ est tout simplement $L^2(H\times K,\lambda\times\mu)$, le produit tensoriel de deux éléments f et g étant la fonction $f\otimes g : (x,y) \longmapsto f(x)g(y)$ sur $H\times K$.

Un peu plus généralement, chaque fois que $\mathcal{H}$ sera interprété comme un espace de fonctions sur un ensemble H, $\mathcal{K}$ comme un espace de fonctions sur K, le produit tensoriel algébrique $\mathcal{H}\underline{\otimes}\mathcal{K}$ pourra s'interpréter comme l'espace vectoriel engendré par les fonctions $f\otimes g$ sur $H\times K$ ($f\in\mathcal{H}$, $g\in\mathcal{K}$) le produit scalaire étant donné, d'autre part, par

$$< f\otimes g,\ f'\otimes g' > = < f,f' >< g,g' > .$$

En mécanique quantique, si $\mathcal{H}$ et $\mathcal{K}$ sont les espaces de Hilbert décrivant deux systèmes physiques, $\mathcal{H}\otimes\mathcal{K}$ permet de décrire l'ensemble des deux systèmes. Les lois pures $\varepsilon_{f\otimes g}$ décrivent des états de cet ensemble dans lesquels les deux systèmes n'interagissent pas (cf. la notion classique d'indépendance, décrite par une loi produit).

Exemples. 1) Si l'on se donne une base o.n. $(f_i)_{i\in I}$ de $\mathcal{H}$ ($(g_j)_{j\in J}$ de $\mathcal{K}$), un élément x de $\mathcal{H}$ s'identifie à la fonction $(<f_i,x>)_i = (x_i)$ sur I, $\mathcal{H}$

1. Principalement Comm. Math. Phys. 93, 1984, p. 301-323.

s'identifie à $\ell^2(I)$, K à $\ell^2(J)$ de même, $H\otimes K$ admet les $f_i\otimes g_j$ comme base o.n., et s'identifie à $\ell^2(I\times J)$.

2) On peut identifier H à l'espace des formes antilinéaires sur $H=H$, en associant à $x\in H$ la fonction $\langle .,x\rangle$ (on n'a pas le choix : l'identification doit être linéaire. L'espace des fonctions $\langle x,.\rangle$ sera identifié à l'antidual H' de H). Alors $H\otimes H$ s'identifie à un espace de formes (anti)-bilinéaires sur $H\times H$, et $H\otimes H'$ à un espace de formes linéaires par rapport à la seconde variable, antilinéaires par rapport à la première, i.e. s'écrivant $\langle x,Ay\rangle$ où A est un opérateur. On vérifie sans peine que $H\otimes H'$ représente l'espace des opérateurs de Hilbert-Schmidt.

Faisons un petit catalogue de propriétés utiles.

a) Si $A : H\longrightarrow H'$ et $B : K\longrightarrow K'$ sont deux opérateurs bornés, on définit sans peine $A\otimes B : H\otimes K \longrightarrow H'\otimes K'$ satisfaisant à

$$A\otimes B(f\otimes g) = (Af)\otimes(Bg)$$

et l'on a des propriétés de composition immédiates. Nous aurons besoin de savoir que

$$\|A\otimes B\| \leqq \|A\|\|B\|. \tag{1}$$

ce qui entraînera que l'opérateur se prolonge par continuité aux complétés. Pour établir (1), on peut se ramener aux deux cas où $K=K'$, $B=I$ et où $H=H'$, $A=I$, le cas général s'obtenant par composition. Traitons le second. Soit $z=\Sigma_i\ x_i\otimes y_i \in H\otimes K$, de sorte que $(I\otimes B)(z)=\Sigma_i\ x_i\otimes By_i = z'$. Par le procédé usuel d'orthogonalisation, on peut supposer que les x_i forment un système orthonormal dans H, de sorte que les $x_i\otimes y_i$ et les $x_i\otimes By_i$ forment des systèmes orthogonaux dans leurs espaces respectifs. On a alors

$$\|z\|^2 = \Sigma_i\ \|y_i\|^2 , \quad \|z'\|^2 = \Sigma_i\ \|By_i\|^2$$

et la propriété est alors évidente.

b) On peut définir des produits tensoriels à plusieurs facteurs, et nous nous bornerons la plupart du temps à des facteurs identiques. Il y a une << associativité >> triviale du produit tensoriel, qui nous permet d'écrire simplement $H^{\otimes n}$. Nous aurons toujours affaire à un H de la forme $L^2(E,\lambda)$, donc $H^{\otimes n}$ sera, concrètement, un espace de (classes de) fonctions sur E^n. Si (e_i) est une base o.n. de H, les $e_{i_1}\otimes e_{i_2}\ldots\otimes e_{i_n}$ forment une base o.n. de $H^{\otimes n}$.

Avant de décrire le cas de n facteurs, faisons quelques remarques sur la cas n=2. On dispose de la notion de fonction de 2 variables (et de forme bilinéaire) symétrique ou antisymétrique, et l'on définit

$$x\circ y = \tfrac{1}{2}(x\otimes y+y\otimes x) \quad , \quad x\wedge y = \tfrac{1}{2}(x\otimes y-y\otimes x)$$

ces vecteurs engendrant respectivement les sous-espaces fermés $H\circ H$, $H\wedge H$ de $H\otimes H$ constitués des formes (anti)bilinéaires symétriques ou antisymétriques.

Si l'on gardait pour le produit scalaire de formes symétriques ou antisymétriques la même valeur que dans $\mathcal{H}\otimes\mathcal{H}$, on aurait

$$\langle x\circ y, x'\circ y'\rangle = \tfrac{1}{2}[\langle x,x'\rangle\langle y,y'\rangle + \langle x,y'\rangle\langle y,x'\rangle]$$

$$\langle x\wedge y, x'\wedge y'\rangle = \tfrac{1}{2}[\langle x,x'\rangle\langle y,y'\rangle - \langle x,y'\rangle\langle y,x'\rangle]$$

Il y a alors de très bonnes raisons de supprimer ces coefficients $\frac{1}{2}$, en se rappelant qu'une forme symétrique ou antisymétrique n'a pas la même norme dans son propre espace et dans le gros espace $\mathcal{H}\otimes\mathcal{H}$: cela nous permettra plus bas d'avoir une identification parfaite de la situation algébrique et de son interprétation probabiliste au moyen des chaos de Wiener.

> Il y a un seul cas où cela ne peut se faire décemment : celui où $\mathcal{H}$ est de dimension 1, que nous avons vu sous un autre nom dans l'exposé III. Alors $\mathcal{H}\circ\mathcal{H}=\mathcal{H}\otimes\mathcal{H}$ s'identifie à $\mathbb{C}$, et de même aux étages supérieurs. De là certaines différences dans les formules.

Si les (e_i) forment une base o.n. de $\mathcal{H}$, les $e_i\wedge e_j$ $(i<j)$ forment une base o.n. de $\mathcal{H}\wedge\mathcal{H}$, et les $e_i\circ e_j$ $(i\leq j)$ une base orthogonale de $\mathcal{H}\circ\mathcal{H}$, avec $\|e_i\circ e_j\|^2=1$ si $i\neq j$, 2 si $i=j$.

Tout cela s'étend sans peine à n dimensions : $\mathcal{H}^{\otimes n}$ s'interprète comme un espace de formes n-(anti)linéaires continues, et le groupe symétrique $\Sigma(n)$ opère sur $\mathcal{H}^{\otimes n}$: si σ est une permutation, l'opérateur $\overline{\sigma}$ associé sur $\mathcal{H}^{\otimes n}$ satisfait à

$$\overline{\sigma}(x_1\otimes\ldots\otimes x_n) = x_{\sigma(1)}\otimes\ldots\otimes x_{\sigma(n)} \quad ,$$

et l'on définit alors les opérateurs de symétrisation et d'antisymétrisation

$$\mathbf{s} = \frac{1}{n!}\sum\overline{\sigma} \quad , \quad \mathbf{a} = \frac{1}{n!}\sum\varepsilon(\sigma)\overline{\sigma} \quad (\ \varepsilon(\sigma),\ \text{signature de } \sigma\ )$$

qui sont des projecteurs de $\mathcal{H}^{\otimes n}$ sur les espaces $\mathcal{H}^{\circ n}$ et $\mathcal{H}^{\wedge n}$ de formes[1] respectivement symétriques et antisymétriques. En particulier, par symétrisation ou antisymétrisation du produit $x_1\otimes\ldots\otimes x_n$, on définit le <u>produit symétrique</u> $x_1\circ\ldots\circ x_n$ et le <u>produit extérieur</u> $x_1\wedge\ldots\wedge x_n$. On fait de ces deux espaces des espaces de Hilbert, <u>en convenant</u> (comme dans le cas où n=2) <u>que le produit scalaire de deux éléments du petit espace est égal à n! fois leur produit scalaire dans le gros espace</u> $\mathcal{H}^{\otimes n}$.

<u>Commentaire</u>. En mécanique quantique, si $\mathcal{H}$ est l'espace de Hilbert décrivant un objet, $\mathcal{H}^{\otimes n}$ est l'espace de Hilbert décrivant n copies du même objet. Mais les objets de la mécanique quantique sont vraiment des objets "élémentaires", i.e. doués d'un très petit nombre de propriétés, et il est interdit de leur en imaginer d'autres - en particulier, d'imaginer que les n objets ont été peints de couleurs différentes pour que l'on puisse les distinguer les uns des autres. Les n copies doivent être considérées comme formant un seul système indissociable. En pratique, on n'a rencontré dans la nature que <u>deux</u> types d'objets, quant à la manière dont ils s'associent

1. Nous noterons souvent $\mathcal{H}_n$ le n-ième espace, dans les deux cas.

avec des compagnons indistinguables : les bosons, pour lesquels l'espace à n objets est le produit tensoriel symétrique $\mathcal{H}^{\odot n}$, et les fermions, pour lesquels c'est le produit antisymétrique $\mathcal{H}^{\wedge n}$.

Ceci décrit des systèmes d'objets en nombre n fixé. Mais on a été amené (en particulier pour les besoins de la mécanique quantique relativiste, où des particules peuvent apparaître ou disparaître) à considérer des systèmes d'objets en nombre fini, mais a priori indéterminé, et susceptible de varier. L'espace de Hilbert correspondant est la somme directe hilbertienne des espaces à n objets, y compris un espace à 0 objet (engendré par un vecteur normalisé Ω_0 appelé vide). Cette somme directe hilbertienne est appelée espace de Fock symétrique ou antisymétrique suivant le cas, construit sur l'espace de Hilbert $\mathcal{H}$. Nous le noterons $\Phi^{\odot}(\mathcal{H})$ (bosons) ou $\Phi^{\wedge}(\mathcal{H})$ (fermions), mais le plus souvent simplement Φ . Nous nous intéresserons davantage au cas symétrique, qui est plus étroitement lié aux probabilités, et au calcul stochastique classique - quoique, nous le verrons, l'espace antisymétrique ait aussi des relations avec le mouvement brownien.

Dans les applications à la physique, $\mathcal{H}$ pourra être $L^2(\mathbb{R}^3)$, ou $L^2(\mathbb{R}^4)$, ou le produit tensoriel de l'un de ces espaces avec un espace de dimension finie (particules à spin). Dans les applications probabilistes, $\mathcal{H}$ sera $L^2(\mathbb{R}_+)$, moins souvent $L^2(\mathbb{R})$. Notre ambition n'est pas d'aborder les "problèmes mathématiques de la théorie quantique des champs" . Cependant, il est possible que certaines propriétés purement << mesurables >> soient transférables de $L^2(\mathbb{R}_+)$ à tous les espaces de Fock (comme cela se produit en théorie des mesures à accroissements indépendants, où les résultats établis, au moyen de la théorie des martingales, pour les p.a.i. non homogènes, peuvent être transférés à tous les espaces mesurables lusiniens).

Soulignons enfin que, lorsque $\mathcal{H}$ est donné comme un espace $L^2_{\mathbb{C}}$, il se trouve muni d'une structure plus riche : il y a un sous-espace réel, une conjugaison complexe $x \longmapsto \bar{x}$. Tout cela se transporte sur l'espace de Fock. En particulier, de manière concrète, nous distinguerons dans $\mathcal{H}^0$ le vecteur vide $\Omega_0 = 1 \in \mathbb{R}$, et la loi de probabilité ε_1 qui lui est associée sur l'espace de Fock (nous l'appellerons la loi naturelle)

2. Dans cette section, nous étudions un peu la structure algébrique de l'espace de Fock, et en particulier nous définissons les opérateurs de création et d'annihilation, qui ne nous quitteront plus de toute la suite.

Cas symétrique. Nous commençons par exhiber une base orthogonale de l'espace de Fock. Soit (e_i) une base o.n. de $\mathcal{H}$. Identifions $\mathcal{H}$ à un espace de formes (anti)linéaires sur lui-même, $\mathcal{H}^{\odot n}$ à un espace de formes n-linéaires symétriques (plus exactement anti..., mais oublions cela) sur $\mathcal{H}^n$. Or une forme n-linéaire symétrique est uniquement déterminée par le polynôme

sur $\mathcal{H}$ qui lui est associé, et le produit symétrique correspond à la multiplication de ces polynômes (pour les éléments très particuliers sur lesquels le produit symétrique a été défini jusqu'à maintenant). Nous adoptons donc les notations bien connues pour les algèbres de polynômes : on appelle multiindice un élément α de $I^{\mathbb{N}}$ (où I est l'ensemble d'indices de la base) qui n'est différent de 0 que pour des indices en nombre fini $i_1,\dots,i_k$ pour lesquels α vaut $\alpha_1,\dots,\alpha_k$; on pose $|\alpha|=\alpha_1+\dots+\alpha_k$, $\alpha! = \alpha_1!\dots\alpha_k!$. On identifie e_i au polynôme $X_i=<.,e_i>$, le produit symétrique $e_\alpha = e_{i_1}{}^{o\alpha_1}_{o}\dots e_{i_1}{}^{o\alpha_1}_{o} e_{i_2}{}^{o\alpha_2}_{o}\dots e_{i_2}\dots e_{i_k}{}^{o\alpha_k}_{o}\dots e_{i_k}$ au polynôme $X^\alpha = X_{i_1}^{\alpha_1}\dots X_{i_k}^{\alpha_k}$. Ces divers éléments constituent une base de l'algèbre des polynômes. Quant au produit scalaire que nous avons mis, il en fait une base orthogonale, avec $\|X^\alpha\|^2=\alpha!$. Lorsque $\mathcal{H}$ est de dimension finie, on retrouve ainsi le très utile produit scalaire ($<P,Q>$ = terme constant de $\overline{P(D)}Q$)[1] de la théorie des harmoniques sphériques, ce qui confirme que nos conventions sont raisonnables.

Dans le langage de l'espace de Fock, cette multiplication s'appelle[2] produit de Wick d'éléments (d'ordre fini) de l'espace de Fock. L'algèbre des polynômes ne décrit que la somme non complétée des $\mathcal{H}^{on}$: la somme complétée peut s'interpréter comme un espace de Hilbert de fonctions entières (ou plus naturellement, antientières !), qui n'est pas stable par multiplication. La coutume est de noter $:X^\alpha:$ les "monômes de Wick" écrits plus haut.

Cas antisymétrique. L'espace de Fock antisymétrique admet une base orthonormale formée des éléments e_A , où A est une partie finie $(i_1,\dots,i_n)$ de l'ensemble d'indices I : supposant celui-ci ordonné totalement de manière arbitraire, on peut supposer $i_1<i_2\dots<i_n$, et alors

$$e_A = e_{i_1}\wedge e_{i_2}\dots\wedge e_{i_n}$$

Ici encore, la somme non complétée des $\mathcal{H}^{\wedge n}$ porte une multiplication associative, celle de l'algèbre extérieure, pour laquelle

$$e_A e_B = (-1)^{n(A,B)} e_{A\cup B} \quad \text{si } A\cap B=\emptyset \ , \ = 0 \text{ sinon .}$$

Le coefficient n(A,B) a été défini dans l'exposé II, §3, qui présente en fait la théorie de l'espace de Fock antisymétrique lorsque $\mathcal{H}$ est de dimension finie. Celui-ci, contrairement à l'espace de Fock symétrique, est alors de dimension finie .

Opérateurs de création et d'annihilation.

Nous commençons par l'opérateur de création, qui fait monter d'un degré dans l'échelle : si $h\in\mathcal{H}$, on pose

$$(2) \qquad a_h^+(x_1{}^o\dots{}^o x_n) = h{}^o x_1{}^o\dots{}^o x_n \ , \quad b_h^+(x_1\wedge\dots\wedge x_n) = h\wedge x_1\dots\wedge x_n$$

1. P(D) est interprété comme un opérateur de dérivation $\Sigma_\alpha a_\alpha D^\alpha$.
2. Sauf erreur ! Ainsi $\mathcal{E}(f)$ (cf. (9)) est écrit $:e^f:$.

En particulier $a_h^+ 1 = h$, $b_h^+ 1 = h$. Pour calculer la norme de $a_h^+ : \mathcal{H}^{\odot n} \to \mathcal{H}^{\odot n+1}$, on peut se ramener au cas où $h = e_1$ dans une base orthonormale (e_n) de $\mathcal{H}$, et l'on voit alors que $\|a_h^+\| = \sqrt{n}\|h\|$, de sorte que a_h^+ considéré comme opérateur sur $\mathfrak{F}$ (admettant comme domaine la somme non complétée $\mathfrak{F}_0$ des $\mathcal{H}_n$) n'est pas borné. En revanche, dans le cas antisymétrique, b_h^+ est de norme $\|h\|$.

L'opérateur d'annihilation est défini pour chaque degré de l'échelle comme l'adjoint du précédent. Un calcul simple montre que sur $\mathcal{H}^{\odot n+1}$

$$(3) \qquad a_h^-(x_0 \odot \ldots \odot x_n) = \Sigma_i \langle h, x_i \rangle\, x_0 \odot \ldots \odot \hat{x}_i \odot \ldots \odot x_n ,$$

où le $\hat{}$ signifie l'omission de x_i . En particulier $a_h^- 1 = 0$. De même pour le cas antisymétrique

$$(4) \qquad b_h^-(x_0 \wedge \ldots \wedge x_n) = \Sigma_i\, (-1)^i\, x_0 \wedge x_1 \ldots \wedge \hat{x}_i \wedge \ldots \wedge x_n .$$

On notera que l'application $h \mapsto a_h^-$ ou b_h^- est antilinéaire.

Les opérateurs bornés b_h^-, b_h^+ sont évidemment adjoints l'un de l'autre. Dans la situation symétrique, une discussion est nécessaire[1]. Etendons d'abord (sans changer de notation) le domaine de $a_h^\pm$ à tous les $x = \Sigma_n x_n$ $(x_n \in \mathcal{H}_n)$ tels que $\Sigma_n \|a_h^\pm x_n\|^2 < \infty$, en posant $a_h^\pm x = \Sigma_n a_h^\pm x_n$. Il est très facile de vérifier que nous obtenons ainsi un opérateur fermé, qui est en fait la fermeture de l'opérateur $a_h^\pm$ précédemment défini. De plus, l'adjoint de $a_h^\pm$ est $a_h^\mp$. En effet, pour $x, y \in \mathfrak{F}_0$ (somme non complétée) $\langle x, a_h^- y \rangle = \langle a_h^+ x, y \rangle$. Cela s'étend à $y \in \mathfrak{F}_0$, $x \in \mathcal{D}(a_h^+)$, montrant que dans ce cas $\langle x, a_h^- . \rangle$ sur $\mathfrak{F}_0$ est continue : donc $a_h^+ \subset a_h^{-*}$. Inversement, si $x \in \mathcal{D}(a_h^{-*})$, $a_h^{-*}(x) = z$, on a pour $y \in \mathfrak{F}_0$ $\langle x, a_h^- y \rangle = \langle z, y \rangle$. Posant $x = \Sigma x_n$, $z = \Sigma z_n$ ($x_n, z_n \in \mathcal{H}_n$) on voit que $z_n = a_h^+ x_{n-1}$, donc $\Sigma_k \|a_h^+ x_k\|^2 < \infty$, $x \in \mathcal{D}(a_h^+)$, $a_h^+ x = z$. Cela justifie les notations a_h, a_h^* souvent utilisées pour a_h^-, a_h^+ (recopié dans Bratteli-Robinson).

D'après les évaluations de normes faites plus haut, si $x = \Sigma_n x_n$ $(x_n \in \mathcal{H}_n)$ avec $\Sigma_n n\|x_n\|^2 < \infty$, x appartient au domaine de a_h^+ et a_h^- .

Relations de commutation et d'anticommutation.

Un calcul immédiat sur la définition des opérateurs $a^\pm$, $b^\pm$ montre que, sur la somme non complétée $\mathfrak{F}_0$, on a

$$(5) \qquad [a_h^-, a_k^+] = \langle h, k \rangle I \qquad \text{et de même} \qquad \{b_h^-, b_k^+\} = \langle h, k \rangle I$$

Il est tentant de définir sur $\mathfrak{F}_0$ les opérateurs suivants, qui sont de bons candidats pour avoir une extension autoadjointe

$$(6) \qquad Q_h = a_h^+ + a_h^- \quad , \quad P_h = i(a_h^+ - a_h^-)$$

et pour satisfaire à une relation de commutation du type d'Heisenberg (avec $\hbar = 2$: normalisation probabiliste !)

$$(7) \qquad [P_h, P_k] = [Q_h, Q_k] = 0 \quad ; \quad [P_h, Q_k] = \frac{2}{i}\langle h, k \rangle I \quad .$$

1. Mais peut être omise sans inconvénient pour la suite.

Comme d'habitude, il faudra transformer cela en une forme plus précise des relations de commutation : la forme de Weyl (cf. n°3 plus bas).

Dans le cas antisymétrique, on aura des opérateurs autoadjoints bornés

$$(8) \qquad R_h = b_h^+ + b_h^- \quad , \quad S_h = i(b_h^+ - b_h^-) \quad .$$

Vecteurs cohérents

Ici, nous nous restreignons au cas symétrique. Pour $f \in \mathcal{H}$, nous posons

$$(9) \qquad \mathcal{E}(f) = \mathbf{1} + \Sigma_n \frac{f^{\circ n}}{n!} \quad (\|f^{\circ n}\|^2 = n!\|f\|^{2n} : \text{la série converge }).$$

Ces vecteurs sont appelés vecteurs cohérents . Ce sont les mêmes que dans l'exposé précédent (espace de Fock sur un espace $\mathcal{H}$ de dimension 1), mais ce n'est pas tout à fait évident avec les différences de normalisation ! Ici nous ne les normalisons pas au sens hilbertien, mais par la condition $< \mathbf{1}, \mathcal{E}(f) > = 1$. Notons la formule importante

$$(10) \qquad < \mathcal{E}(f), \mathcal{E}(g) > = \Sigma_n \, (n!)^{-2} < f^{\circ n}, g^{\circ n} > = \Sigma_n \, <f,g>^n / n! = e^{<f,g>} \quad .$$

L'espace vectoriel fermé engendré par les vecteurs cohérents contient tous les vecteurs $f^{\circ n} = \frac{d^n}{dt^n} \mathcal{E}(tf)|_{t=0}$: il est donc dense dans Φ . Les vecteurs cohérents servent très naturellement de << fonctions-test >> dans les calculs d'opérateurs non bornés sur l'espace de Fock. Par exemple, les opérateurs de création et d'annihilation opèrent sur les vecteurs cohérents par les règles suivantes

$$(11) \qquad a_h^- \mathcal{E}(f) = <h,f> \mathcal{E}(f)^{(1)}, \quad a_h^+ \mathcal{E}(f) = \frac{d}{dt} \mathcal{E}(f+th)|_{t=0} \quad .$$

Nous recopions aussi quelques formules, que le lecteur pourra vérifier, et qui serviront plus tard

$$(12) \qquad \begin{aligned} < a_h^- \mathcal{E}(f), a_k^- \mathcal{E}(g) > &= <f,h><k,g> < \mathcal{E}(f), \mathcal{E}(g) > \\ < a_h^+ \mathcal{E}(f), a_k^+ \mathcal{E}(g) > &= \{<f,k><h,g> + <h,k>\} < \mathcal{E}(f), \mathcal{E}(g) > \\ < a_h^+ \mathcal{E}(f), a_k^- \mathcal{E}(g) > &= <k,g><h,g> < \mathcal{E}(f), \mathcal{E}(g) > \quad . \end{aligned}$$

Il nous arrivera souvent de définir des opérateurs par leur valeur sur les vecteurs cohérents. Le lemme suivant est commode à cet effet. La démonstration est recopiée dans Guichardet, Symmetric Hilbert Spaces, LN 261.

LEMME. Soit (f_i) une famille finie d'éléments distincts de $\mathcal{H}$. Alors les vecteurs cohérents $\mathcal{E}(f_i)$ sont linéairement indépendants.

Dém. Supposons une relation $\Sigma_{i=1}^k \bar{\lambda}_i \mathcal{E}(f_i) = 0$, avec des $\lambda_i \neq 0$. Pour tout $g \in \mathcal{H}$ on a $\Sigma_1^k < \bar{\lambda}_i \mathcal{E}(f_i), \mathcal{E}(g) > = 0$, ce qui d'après (10) donne $\Sigma_1^k \lambda_i \exp <f_i, g> = 0$. Remplaçons g par g+th , dérivons k fois pour t=0 . Il vient

---

1. Les vecteurs cohérents sont les vecteurs propres communs des a^- : c'est le moyen le plus commode de vérifier qu'il s'agit bien de la même chose que dans l'exposé précédent.

$$\Sigma_1^k \lambda_i < f_i, . >^p = 0 \quad , \quad p=0,1,\ldots,k-1$$

Les λ_i étant $\neq 0$, le déterminant est identiquement nul. Or c'est un déterminant de Vandermonde $\prod_{i<j} <f_j-f_i, .>$. Pour tout couple i,j l'ensemble des x tels que $<f_j-f_i, x> \neq 0$ est un ouvert, dense puisque $f_i \neq f_j$. L'intersection de ces ouverts est donc non vide, ce qui contredit le résultat précédent.

3. Dans cette section, nous allons construire les opérateurs de Weyl. Nous recopions la présentation de Hudson-Parthasarathy.

Considérons le groupe des déplacements de l'espace de Hilbert $\mathcal{H}$. Un élément du groupe s'écrit $\lambda=(u,U)$, où u est un vecteur et U un opérateur unitaire, λ opérant de la manière suivante

$$\lambda.h = Uh+u \quad \text{pour} \quad h \in \mathcal{H} ,$$

de sorte que la loi de composition dans le groupe est, si $\mu=(v,V)$

$$\lambda\mu = (u,U)(v,V) = (u+Uv,UV).$$

Nous allons faire agir le groupe sur l'espace de Fock : d'abord sur les vecteurs cohérents

$$W_\mu \mathcal{E}(h) = e^{-<v,Vh>-|v|^2/2} \mathcal{E}(\mu.h) = e^{-A_\mu(h)} \mathcal{E}(\mu.h) \tag{13}$$

Nous vérifions la formule

$$\underset{(a)}{A_{\lambda\mu}(h)} = \underset{(b)}{A_\lambda(\mu.h)} + \underset{(c)}{A_\mu(h)} - iIm<u,Uv>$$

En effet , $a = < u+Uv, UVh > + |u+Uv|^2/2$

$$b = < u, U(Vh+v) > + |u|^2/2$$

$$c = < v, Vh > + |v|^2/2 = < Uv, UVh > + |v|^2/2$$

Donc la différence a-b-c vaut $-<u,Uv> + \frac{1}{2}(<u,Uv>+<Uv,u>) = -\frac{1}{2}[<u,Uv> - \overline{<u,Uv>}]$.

A partir de là, nous obtenons la relation de Weyl

$$W_\lambda W_\mu = e^{-iIm<u,Uv>} W_{\lambda\mu} \tag{14}$$

qui est très voisine de la formule (10) de l'exposé précédent : on a en plus l'opérateur unitaire U au lieu de I, et on a ici $\hbar=2$ (et le signe opposé devant i). Nous avons d'après (13)

$$< W_\mu \mathcal{E}(h), W_\mu \mathcal{E}(k) > = e^{-\overline{A_\mu(h)} - A_\mu(k)} < \mathcal{E}(\mu.h), \mathcal{E}(\mu.k) > .$$

Ce dernier produit scalaire vaut $e^{<\mu.h,\mu.k>}$ d'après (10). Il nous reste finalement l'exponentielle de

$$-\overline{<v,Vh>} - <v,Vk> - |v|^2 + <Vh+v,Vk+v> = < Vh,Vk > = <h,k>$$

exponentielle qui vaut $< \mathcal{E}(h), \mathcal{E}(k) >$ d'après (10). Les W_μ opèrent ainsi de manière unitaire sur le sous-espace dense des vecteurs cohérents, donc ils s'étendent de manière unique en des opérateurs isométriques, inversibles donc unitaires sur $\mathcal{F}$. La formule (17) nous dit que l'on a construit une représentation unitaire projective

du groupe des déplacements, et celle-ci nous servira, en prenant les générateurs de groupes à un paramètre, à construire autant d'opérateurs a.a. que nous le voudrons... à condition d'avoir vérifié un minimum de continuité. Celle-ci ne présente pas de difficulté, si l'on remarque que $f \to \mathcal{E}(f)$ est faiblement continue en restriction aux bornés (10), et que l'on utilise la formule explicite (13) sur les vecteurs cohérents.

COMMENTAIRE. La C^*-algèbre engendrée par les opérateurs W_λ correspondant aux translations seulement est appelée la C^*-algèbre des relations de commutation canoniques (CCR en anglais). Elle est étudiée en détail dans le second volume du livre de Bratteli-Robinson. En tant que C^*-algèbre abstraite, i.e. dans la topologie de la norme des opérateurs, c'est un être plutôt antipathique : on peut montrer que deux opérateurs W_u et W_v différents sont toujours à la distance 2 (la plus grande distance permise entre deux unitaires) : cf. B-R th. 5.2.8., p.20.

4. Pour interpréter les opérateurs W_λ lorsque $\lambda=(0,U)$ est une rotation pure, il nous reste à introduire une dernière notion algébrique, d'ailleurs très simple. Nous avons rappelé en 1 a) au début du paragraphe la notion de produit tensoriel $A \otimes B : \mathcal{H} \otimes \mathcal{K} \longrightarrow \mathcal{H} \otimes \mathcal{K}$ de deux opérateurs bornés $A : \mathcal{H} \to \mathcal{H}$ et $B : \mathcal{K} \to \mathcal{K}$. Cela se généralise à n facteurs sans la moindre difficulté, et permet en particulier de définir $A^{\otimes n} : \mathcal{H}^{\otimes n} \longrightarrow \mathcal{H}^{\otimes n}$: on définit alors A^{on} et $A^{\wedge n}$ par restriction aux sous-espace symétrique et antisymétrique. Nous étudierons ici le premier, mais tout (sauf l'utilisation des vecteurs cohérents) va de même pour le second. On a

$$A^{on}(x_1 o \ldots o x_n) = Ax_1 o \ldots o Ax_n \quad ; \quad \|A^{on}\| \leq \|A\|^n \text{ (cf. (1)) .}$$

Nous définissons maintenant l'opérateur $\Phi(A)$, dont le domaine est l'ensemble des $x=\Sigma_n x_n$ $(x_n \in \mathcal{H}^{on})$ tels que $\Sigma_n \|A^{on}(x_n)\|^2 < \infty$, et la valeur

(15) $\Phi(A)x = \Sigma_n A^{on} x_n$ ($\Phi(A)$ s'appelle la "seconde quantification de A")

$\Phi(A)$ est fermé, de domaine dense (si A est une contraction, $\Phi(A)$ est aussi une contraction). On a $\Phi(I)=I$, $\Phi(AB)=\Phi(A)\Phi(B)$, $\Phi(A)^*=\Phi(A^*)$. Si A est unitaire, $\Phi(A)$ est unitaire. Les vecteurs cohérents appartiennent toujours au domaine, et l'on a

(16) $$\Phi(A)\mathcal{E}(f) = \mathcal{E}(Af) .$$

Si l'on compare (16) et (13), on voit alors que $W_{0,U}$ est simplement l'opérateur unitaire $\Phi(U)$.

Une autre manière d'étendre à l'espace de Fock un opérateur A défini sur $\mathcal{H}$ consiste à poser formellement

(17) $$\lambda(A) = \frac{d}{dt}\Phi(e^{tA})\Big|_{t=0} \quad \left(\text{ ou } \frac{1}{i}\frac{d}{dt}\Phi(e^{itA}) \right)$$

de sorte que, en supposant toujours A borné, $\lambda(A)$ applique $\mathcal{H}^{on}$ dans

lui même, par la formule

$$(18) \qquad \lambda(A)(x_1 o \ldots o x_n) = (Ax_1) o x_2 \ldots o x_n + x_1 o (Ax_2) o \ldots o x_n + \ldots$$

Après quoi, on définit comme d'habitude le domaine de $\lambda(A)$ comme l'ensemble des $h=\Sigma_n h_n \in \mathfrak{F}$ tels que $\Sigma_n \|\lambda(A)h_n\|^2 < \infty$, et $\lambda(A)h=\Sigma_n \lambda(A)h_n$, ce qui donne un opérateur fermé. Les vecteurs cohérents appartiennent toujours au domaine, et il est facile en comparant (17) et (11) de voir que

$$(19) \qquad \lambda(A)\mathcal{E}(f) = a^+_{Af}\,\mathcal{E}(f)$$

ce qui peut d'ailleurs se voir aussi sur (18). On en tire

$$(20) \qquad < \mathcal{E}(f), \lambda(A)\mathcal{E}(g) > = <f, Ag> e^{<f,g>} .$$

Voici un petit formulaire concernant ces opérateurs, qu'il n'est pas utile de regarder en détail dès maintenant

$$(21) \qquad < \lambda(A)\mathcal{E}(f), \lambda(B)\mathcal{E}(g) > = \{<Af,g><f,Bg> + <Af,Bg>\} < \mathcal{E}(f), \mathcal{E}(g) >$$

$$(22) \qquad < a^-_h \mathcal{E}(f), \lambda(B)\mathcal{E}(g) > = <f,h><f,Bg>< \mathcal{E}(f), \mathcal{E}(g) >$$

$$(23) \qquad < a^+_h \mathcal{E}(f), \lambda(B)\mathcal{E}(g) > = \{<h,g><f,Bg> + <h,Bg> \} < \mathcal{E}(f), \mathcal{E}(g) > .$$

5. Un modèle discret de l'espace de Fock symétrique.

Maintenant que nous avons un peu décrit la structure algébrique de l'espace de Fock symétrique, nous pouvons comparer celui-ci au modèle fini de l'exposé II, § III. Nous avions là un espace de Hilbert $\mathcal{H}$ de dimension finie M , avec une base orthonormale (X_i) (qui pour nous étaient des v.a. de Bernoulli symétriques) choisie une fois pour toutes. $\mathcal{H}$ était plongé dans un espace de Hilbert $\mathcal{Y}$ de dimension finie 2^M, muni d'une décomposition $\mathcal{Y}=\oplus_n \mathcal{Y}_n$ en "chaos" ($\mathcal{Y}_0=\mathbb{C}$, $\mathcal{Y}_1=\mathcal{H}$). L espace $\mathcal{Y}$ comporte beaucoup moins d'éléments de base que l'espace de Fock $\mathfrak{F}$: au lieu de tous les X^α indexés par les multiindices, nous ne conservons que les X_A indexés par les multiindices ne comportant que des 1 et des 0, i.e. par les parties de $\{1,\ldots,M\}$. Ces éléments sont orthonormés dans $\mathcal{Y}$.

Rappelons la définition des opérateurs de création et d'annihilation

$$(24) \qquad a^+_k(X_A) = X_{A\cup\{k\}} \text{ si } k\notin A \text{ , 0 sinon ; } a^-_k(X_A) = X_{A\setminus\{k\}} \text{ si } k\in A, \text{ 0 sinon}$$

et les relations de commutation

$$(25) \qquad [a^-_k, a^-_j] = [a^+_k, a^+_j] = 0 \text{ ; } [a^-_k, a^+_j] = \begin{cases} 0 & \text{si } k\neq j \\ I-2N_k & \text{si } k=j \end{cases} \quad (N_k = a^+_k a^-_k)$$

Nous avons aussi les opérateurs a.a. associés

$$(26) \qquad q_k = a^+_k + a^-_k \text{ , } p_k = i(a^+_k - a^-_k) \qquad (q_k X_A = X_{A\Delta\{k\}}) .$$

Les opérateurs p_k, q_k sont, pour k fixé, des opérateurs a.a. de carré I qui anticommutent (opérateurs de spin : la troisième matrice de Pauli est

ici représentée par $2N_k$-I), les divers spins commutant entre eux. Si l'on pose

$$p_A = \prod_{k\in A} p_k \quad , \quad q_B = \prod_{k\in B} q_k \tag{27}$$

on a $q_Aq_B=q_{A\Delta B}$, $p_Ap_B=p_{A\Delta B}$, $p_Aq_B=(-1)^{|A\cap B|}q_Bp_A$.

Notre espace $\mathcal{Y}$ se trouve muni de diverses structures d'algèbres : la première, celle qui correspond au produit symétrique o , consiste à poser

$$X_A \circ X_B = X_{A\cup B} \text{ si } A\cap B=\emptyset \text{ , 0 sinon} \tag{28}$$

La seconde (qui correspond à ce que nous appellerons plus tard le <u>produit de Wiener</u> sur l'espace de Fock) est le <u>produit de Bernoulli</u> symétrique, défini par

$$X_AX_B = X_{A\Delta B} \tag{29}$$

(nous avons vu aussi dans l'exposé II un produit de Clifford sur $\mathcal{Y}$ - laissons le de côté pour l'instant - et d'autres produits de Bernoulli, correspondant à l'interprétation des X_i comme v.a. de Bernoulli non symétriques).

L'espace $\mathcal{Y}$, description d'un système de spins qui commutent, est un objet très familier pour les physiciens. Mais l'idée de considérer $\mathcal{Y}$ comme une <u>approximation</u> de l'espace de Fock , pour M grand (que nous empruntons à un exposé de J.L. Journé) semble beaucoup moins banale. Nous allons la développer un peu ici, et la relier au point de vue sur l'espace de Fock présenté par Guichardet, Symmetric Hilbert Spaces, LN in M. 261.

Nous commençons par poser, pour $f=\Sigma_k f_kX_k \in \mathcal{H}$

$$a_f^+ = \Sigma_k f_k a_k^+ \quad , \quad a_f^- = \Sigma_k \bar{f}_k a_k^- \tag{30}$$

ces opérateurs étant mutuellement adjoints (le lecteur définira tout seul q_f,p_f pour f réelle) . Introduisons les <u>vecteurs cohérents</u>

$$\mathcal{E}(f) = \Sigma_n \frac{f^{\circ n}}{n!} = \prod_k (1+f_kX_k) \text{ (produit de Bernoulli (29))} \tag{31}$$

Lorsque le nombre de points M devient très grand, k étant identifié au site kT/M sur l'intervalle [0,T], f_k étant pris de la forme $f(kT/M)\sqrt{T/M}$ où f est une bonne fonction sur [0,T], l'expression (31) "tendra vers" une exponentielle stochastique brownienne (en un sens à préciser). D'autre part on a des expressions très analogues à (11)

$$a_g^-\mathcal{E}(f) = (\Sigma_k \frac{\bar{g}_kf_k}{1+f_kX_k})\mathcal{E}(f) \quad , \quad a_g^+\mathcal{E}(f) = (\Sigma_k \frac{g_kX_k}{1+f_kX_k})\mathcal{E}(f) = \frac{d}{dt}\mathcal{E}(f+tg)|_{t=0} \tag{32}$$

Identifions un élément de $\mathcal{Y}$ à son développement $\Sigma_A f_AX_A$ dans la base orthonormale (X_A) : nous voyons que $\mathcal{Y}$ s'identifie à l'ensemble des fonctions $A \longmapsto f(A)$ <u>définies sur l'ensemble des parties de</u> $\{1,\ldots,M\}$. Par exemple, dans cette interprétation le vecteur cohérent $\mathcal{E}(f)$ est donné par

(33) $$\mathcal{E}(f)(A) = \prod_{k\in A} f(k)$$

Quant aux opérateurs de création et d'annihilation, ils sont tout retournés !

(34) $a_k^- f(A) = f(A\cup\{k\})$ si $k\notin A$, 0 sinon ; $a_k^+ f(A) = f(A\setminus\{k\})$ si $k\in A$, 0 sinon

comme on le voit à partir de (24) en prenant $f=X_B$.

REMARQUE. L'ensemble $\mathcal{P}$ des parties de $\{1,\dots,M\}$, muni de l'opération Δ, est un groupe compact ; le groupe dual s'identifie aussi à $\mathcal{P}$, si l'on associe à une partie A le caractère $\chi_A(B)=(-1)^{|A\cap B|}$. En particulier, munissons $\mathcal{P}$ de sa mesure de Haar (qui place la masse 2^{-M} en tout $A\in\mathcal{P}$) ; les v.a. $X_k=\chi_{\{k\}}$ sont des v.a. de Bernoulli symétriques indépendantes, et l'on a $\chi_A=\prod_{k\in A} X_i = X_A$. Ce qui précède est donc de l'analyse de Fourier sur $\mathcal{P}$. Par analogie avec les opérateurs de Weyl $W_{r,s}$, Combe, Rodriguez, Sirugue et S.-Collin ont défini des opérateurs $W_{A,B}$, réalisant une représentation unitaire projective de $\mathcal{P}\times\mathcal{P}$. Voir leur article, Weyl Quantisation of Spin Systems, dans LN in Phys. 106, Feynman Path Integrals Marseille 1978.

Passage au véritable espace de Fock. Voici le point de vue de Guichardet pour exposer la théorie de l'espace de Fock sur $L^2(T,\mu)$, où T est un bon espace mesurable, et μ est une mesure diffuse sur T. Soit $\mathcal{P}$ l'ensemble des parties finies de T (c'est un groupe pour la différence symétrique Δ). Nous munissons $\mathcal{P}$ d'une structure mesurable et d'une mesure λ de la manière suivante : identifions T à une partie borélienne de $\mathbb{R}_+$; alors l'ensemble $\mathcal{P}_n$ des parties à n éléments s'identifie au sous-ensemble de $\mathbb{R}_+^n$ formé des n-uples $\{s_1<\dots<s_n\}$, qui est borélien, et que l'on munit de la mesure λ_n induite par $\mu^{\otimes n}$ (si n=0, $\mathcal{P}_0=\{\emptyset\}$ et $\lambda_0=\varepsilon_\emptyset$). On pose alors $\lambda=\Sigma_n \lambda_n$. Du fait que μ est diffuse, il est facile de vérifier que $L^2(\mathcal{P}_n,\lambda_n)$ est isomorphe à $L^2(\mu)^{\odot n}$, donc $L^2(\mathcal{P},\lambda)$ est isomorphe à l'espace de Fock.

Dans cette représentation, on a pour $h\in L^2(T)$, $f\in L^2(\lambda)$

(35) $a_h^- f(A) = \int f(A\cup\{t\})\overline{h}(t)\mu(dt)$, $a_h^+ f(A) = \Sigma_{t\in A} h(t) f(A\setminus\{t\})$.

Nous reviendrons plus tard sur cette représentation. Notons seulement que le modèle fini que nous avons donné consiste à prendre au sérieux le cas où la mesure μ n'est pas nécessairement diffuse (ce qui se produit en particulier lorsque T est fini).(1)

Nous écrirons souvent dA pour $\lambda(dA)$ sur $\mathcal{P}$, pour alléger la notation.

---

1. Profitons de ce blanc pour indiquer que la théorie rigoureuse de l'espace de Fock est due à J.M. Cook, the Mathematics of Second Quantization, Trans. AMS 74, 1953 (longtemps après Fock, 1932-34). Tout ce qui a été exposé fait partie du folklore maintenant (mais notre présentation emprunte plusieurs détails à Neveu, Processus Aléatoires Gaussiens, Presses Univ. Montréal, 1968).

II. INTERPRETATIONS PROBABILISTES

1. Soit (X_t) une martingale, définie sur un espace probabilisé filtré [1] $(\Omega,\mathcal{F},P,(\mathcal{F}_t)_{t\geq 0})$, nulle en 0, telle que $E[X_t^2]<\infty$ pour tout t et que $\langle X,X\rangle_t=t$. Pour l'instant, nous travaillons sur des martingales réelles, pour simplifier. Appelons quadrant chronologique le sous-ensemble C_n de $\mathbb{R}_+^n$ formé (pour n>1) des n-uples tels que $s_1<s_2\ldots<s_n$ ($C_1=\mathbb{R}_+$), que nous munirons de la mesure de Lebesgue $\lambda_n=ds_1\ldots ds_n$. Pour toute fonction $f\in L^2(C_n)$, on sait définir l'intégrale stochastique multiple

$$J_n(f) = \int_{C_n} f(s_1,\ldots,s_n)dX_{s_1}\ldots dX_{s_n} \tag{1}$$

Ces v.a. appartiennent à $L^2(\Omega)$, avec une norme

$$\|J_n(f)\|^2 = \int_{C_n} |f(s_1,\ldots,s_n)|^2 ds_1\ldots ds_n$$

et pour $m\neq n$, $J_m(f)$ et $J_n(g)$ sont orthogonales. Dans le cas du mouvement brownien, ces résultats sont dus à Wiener, ont été généralisés par Ito (J. Math. Soc. Japan 3, 1951 ; Jap. J. Math 22, 1953). La démonstration pour des martingales générales est essentiellement la même (voir Sém. Prob. X , LN in M. 511, p.325-331).

Soit maintenant une fonction $f\in L^2(\mathbb{R}_+^n)$; on désire lui associer une intégrale stochastique $I_n(f)\in L^2(\Omega)$. Négligeant [2] ce qui se passe sur les diagonales (n-uples dont deux coordonnées au moins coïncident), qui forment un ensemble de mesure nulle, il est clair que le problème consiste à définir les intégrales stochastiques sur les copies du quadrant chronologique

$$\int_{s_{\sigma(1)}<s_{\sigma(2)}<\ldots} f(s_1,\ldots,s_n)dX_{s_1}\ldots dX_{s_n} \quad (\sigma \text{ permutation})$$

et la solution traditionnelle consiste à décider que cette intégrale est égale à

$$\int_{C_n} f(u_{\tau(1)},\ldots,u_{\tau(n)})dX_{u_1}\ldots dX_{u_n} \quad \text{avec} \quad \tau=\sigma^{-1} \tag{2}$$

et en particulier à poser, si f est une fonction symétrique appartenant à $L^2(\mathbb{R}_+^n)$

$$I_n(f) = n!J_n(f) \tag{3}$$

Nous avons alors $\|I_n(f)\|^2 = (n!)^2\int_{C_n} |f(s_1\ldots s_n)|^2 ds_1..ds_n = n!\|f\|^2_{L^2(\mathbb{R}_+^n)}$, qui est précisément la convention que nous avions adoptée pour la norme de f en tant qu'élément de $H^{\odot n}$, lorsque $H=L^2(\mathbb{R}_+)$. Par conséquent, si

---

1. Nous supposerons que $\mathcal{F}$ est engendrée par les v.a. X_t.
2. Il n'y a pas là de nécessité logique : nous décrivons seulement ce qui se fait d'habitude.

nous associons à un élément $\Sigma_n f_n$ de l'espace de Fock symétrique la somme des v.a. $\Sigma_n I_n(f_n)$ dans $L^2(\mathcal{F})$ (en convenant que pour n=0, $I_n(f_0)$ est la v.a. constante f_0), nous plongeons l'espace de Fock symétrique Φ^0 isométriquement dans $L^2(\mathcal{F})$.

Nous ferons la remarque suivante : la propriété (2) est une pure convention, et rien ne nous empêche de mettre un facteur ε_σ devant l'intégrale (2). Alors, la formule (3) aura lieu non plus pour f symétrique, mais pour f antisymétrique, et nous plongerons dans $L^2(\mathcal{F})$ l'espace de Fock $\Phi^\wedge$. Tout le problème est de savoir si cela présente quelque intérêt : je pense que oui, et je tâcherai de montrer que cela fournit une définition simple des algèbres de Clifford de dimension infinie utilisées en physique, avec un peu d'intuition probabiliste en plus.

Le problème se pose maintenant de savoir quelle est l'image de Φ (symétrique ou antisymétrique) dans $L^2(\mathcal{F})$, i.e. quel est le sous-espace de L^2 engendré par les intégrales stochastiques multiples $J_n(f)$. En particulier, on s'intéresse au cas où cet espace est $L^2(\mathcal{F})$ tout entier[1]. Comme les i.s. multiples (1) sont des i.s. $\int H_s dX_s$ de processus prévisibles (en interprétant l'i.s. multiple comme i.s. itérée), une condition nécessaire pour cela est que (X_t) possède la propriété de représentation prévisible. Les probabilistes ne semblent pas disposer d'une liste complète de martingales de carré intégrable X_t possédant la PRP et de crochet $\langle X,X\rangle_t=t$, mais en voici quelques représentants

- Le mouvement brownien standard : $X^0_t=B_t$.
- Les processus de Poisson compensés : $X^\rho_t = \rho(\nu^\rho_t - \frac{t}{\rho^2})$, où ν^ρ est un processus de Poisson à sauts unité, d'intensité $1/\rho^2$ ($\rho\in\mathbb{R}\setminus\{0\}$; lorsque $\rho\to 0$ on retrouve le mouvement brownien, qui peut être noté X^0_t .

Ceci fournit les seuls éléments de la liste qui soient à accroissements indépendants et homogènes dans le temps. Il est facile de construire des martingales Y non homogènes dans le temps, en posant

$$dY_t = dX^{\rho(t)}_t \qquad (\ \rho(t)=0 \text{ est permis }) \tag{4}$$

où $\rho(t)$ peut être une fonction borélienne arbitraire du temps. On peut conjecturer que cette liste est, en fait, complète.

COMMENTAIRE. Chaque fois qu'un espace de Hilbert H est interprété comme un espace $L^2(\Omega)$, il porte une observable à valeurs dans Ω (ou si l'on préfère, une mesure spectrale indexée par la tribu de Ω). Ainsi, le fait que l'espace de Fock admette des interprétations browniennes ou poissonniennes signifie que cet objet, purement algébrique, porte des observables à valeurs

---

1. Segal a construit, au moyen du mt brownien complexe, une intéressante interprétation probabiliste de l'espace de Fock dans laquelle celui-ci n'est pas L^2 tout entier.

dans l'espace des trajectoires continues ou des trajectoires ponctuelles.

Soit P la loi sur Ω ; puisque P est une loi de probabilité, la constante 1 appartient à L^2, et détermine une loi quantique ε_1 , sous laquelle la loi de notre observable est P . Dans le cas particulier de l'espace de Fock, la constante 1 correspond au vecteur-vide **1** dans toutes les interprétations probabilistes que nous avons vues. Ainsi sous la loi naturelle ε_1, les observables décrites plus haut ont une loi brownienne, resp. des lois de processus de Poisson de toutes intensités. Mais rien ne nous impose de rester pour toujours dans le vide ; l'idée de départ de ces notes a été de chercher une interprétation de ce genre, pour la mécanique stochastique de Nelson, à l'intérieur de la mécanique quantique traditionnelle : existe t'il une manière naturelle de choisir un vecteur f de $\mathfrak{F}$ (plus exactement, il faut rester en temps fini) tel que, sous la loi ε_f , l'observable "brownienne" se transforme[1] en une diffusion de Nelson ? Les autres processus sur l'espace de Fock se transformeraient sans doute de manière intéressante. Mais c'est là sans doute un travail de physicien, plutôt que de mathématicien (et de toute façon, il dépasse nos capacités en mécanique quantique).

Quoi qu'il en soit, nous revenons à Hudson-Parthasarathy : nous allons d'abord étudier l'interprétation brownienne, en montrant comment se traduisent les divers êtres algébriques introduits au paragraphe I. Puis nous construirons les processus de Poisson de divers paramètres, et parlerons un peu de l'algèbre de Clifford continue.

2. Vecteurs cohérents et exponentielles stochastiques.

Nous allons commencer par montrer que les vecteurs cohérents ont toujours une interprétation probabiliste simple.

Si (M_t) est une semimartingale nulle en 0, son exponentielle de Doléans $Z=\mathcal{E}(M)$ est par définition la solution de l'équation différentielle stochastique $Z_t=1+\int_0^t Z_{s-}dM_s$, et il est connu que Z admet un développement convergent en probabilité

$$Z_t = 1 + \int_0^t dM_{s_1} + \int_{s_1<s_2<t} dM_{s_1} dM_{s_2} + \ldots$$

qui est tout simplement la solution de l'e.d.s. par la méthode d'itération (cf. Emery, ZW 41, 1978, p. 256). En particulier, prenons $M_t = \int_0^t f(s)dX_s$, où (X_t) est une martingale de crochet $\langle X,X\rangle_t=t$. Abrégeant $fI_{[0,t]}$ en $f_{t]}$, et utilisant les notations du n° précédent, nous voyons que le n-ième terme de cette somme est $\frac{1}{n!} I_n(f_{t]}^{\otimes n})$. Si f appartient à $L^2(\mathbb{R}_+)$, la série converge dans $L^2(\Omega)$, et le processus (Z_t) est une vraie martingale bornée dans L^2, donc convergeant dans L^2 vers une v.a. Z_∞ . La v.a. Z_∞ est

1. Cela exigerait un certain élargissement du formalisme : remplacer $L^2(\mathbb{R}_+)$ par $L^2(\mathbb{R}_+,\mathbb{R}^3)$ pour disposer de plusieurs browniens, introduire un espace de Hilbert initial pour donner à X_0 une loi non dégénérée. Nous restons ici dans la situation la plus simple possible.

l'élément de $L^2(\Omega)$ qui correspond au vecteur $\Sigma_n \frac{1}{n!} f^{\otimes n}$, autrement dit, c'est l'interprétation probabiliste du vecteur cohérent $\mathcal{E}(f)$.

Traitons explicitement les cas particuliers vus plus haut

- Cas brownien . On a alors, d'après l'expression explicite de l'exponentielle de Doléans dans le cas des martingales continues

$$\mathcal{E}(f) = \exp\left(\int_0^\infty f_s dB_s - \frac{1}{2}\int_0^\infty f_s^2 ds\right) \tag{5}$$

- Cas poissonniens. Si (X_t) est le processus de Poisson compensé (4), de saut ρ_t (s'il se produit un saut) à l'instant t, on a pour l'exponentielle de Doléans

$$\mathcal{E}(f) = \exp\left(\int_0^\infty f_s dX_s\right) \prod_{s\in S} (1+f_s\rho_s)e^{-f_s\rho_s}$$

où S est l'ensemble aléatoire des sauts de $X_\cdot$. Si la fonction f/ρ est intégrable, cette expression se simplifie et devient

$$\mathcal{E}(f) = \exp\left(-\int_0^\infty f_s ds/\rho_s\right)\prod_{s\in S}(1+f_s\rho_s) \tag{6}$$

REMARQUES. a) L'espace de Hilbert $\mathcal{H}=L^2(\mathbb{R}_+)$ se trouve muni d'un sous-espace réel, et d'une conjugaison $f \mapsto \overline{f}$: on notera que la quantité $\int_0^\infty f^2 ds$ qui apparaît dans (5) est égale à $\langle \overline{f}, f\rangle$, et n'est pas une quantité intrinsèque de la structure hilbertienne de $\mathcal{H}$.

b) Toute identification de l'espace de Fock à un espace $L^2(\Omega)$ munit Φ de structures supplémentaires : par exemple, on peut parler de la réalité, de la positivité, de l'appartenance à L^p... de certains éléments de l'espace de Fock. Celui-ci se trouve en fait plongé dans le beaucoup plus gros espace de toutes les v.a. sur Ω, avec sa conjugaison complexe et sa multiplication. En particulier, on peut regarder comment se multiplient les éléments de l'espace de Fock, dans chacune des diverses identifications. Comme nous considérons plusieurs identifications à la fois, il faut bien préciser quelle identification on utilise pour définir ces divers produits.

Sur l'importance, en mécanique quantique, de ces méthodes non purement hilbertiennes, on pourra consulter les premières pages de Gross, Existence and Uniqueness of Physical Ground States, J.Funct. Anal. 10, 1972.

Par exemple, dans l'identification brownienne, les éléments des chaos de Wiener $\mathcal{H}_n$ appartiennent à tous les L^p, et la somme non complétée Φ_0 des $\mathcal{H}_n$ est une algèbre pour le "produit de Wiener". Les vecteurs cohérents appartiennent, eux aussi, à tous les L^p, leur produit de Wiener étant donné, d'après (5), par la formule

$$\mathcal{E}(f)\mathcal{E}(g) = \mathcal{E}(f+g)e^{\int f_s g_s ds} = e^{\langle \overline{f}, g\rangle}\,\mathcal{E}(f+g) \tag{7}$$

ce qui montre que le produit de Wiener est "presque" intrinsèque : sa définition n'exige que l'addition d'un seul élément de structure non hilbertien sur $\mathcal{H}$, la conjugaison complexe $f \mapsto \overline{f}$.

1. Dans l'interprétation brownienne associée au m^t brownien conjugué P_t, on a un autre produit de Wiener, pour lequel $\mathcal{E}(f)\mathcal{E}(g)=\exp(-\langle \overline{f}, g\rangle)\mathcal{E}(f+g)$.

la fonction caractéristique de P_h

$$E[e^{iP_h}] = \langle 1, W_{-h}1 \rangle = \langle 1, e^{-\|h\|^2/2}\mathcal{E}(-h)\rangle \quad (\S I,(13)\ ;\ 1=\mathcal{E}(0))$$
$$= e^{-\|h\|^2/2}$$

d'où il résulte que les v.a. P_h forment un espace gaussien, et en particules les v.a. P_t correspondant à $h=I_{[0,t]}$ forment un mouvement brownien standard.

D'autre part, un calcul direct sur l'exponentielle de Doléans montre que, pour h réelle

$$(10) \qquad a_h^-\mathcal{E}(f) = \langle h,f\rangle\mathcal{E}(f) = \nabla_h\mathcal{E}(f)\ ,\quad a_h^+\mathcal{E}(f) = \frac{d}{dt}\mathcal{E}(f+th)\Big|_0 = (\tilde{h}-\nabla_h)\mathcal{E}(f)$$

de sorte que, sur les vecteurs cohérents, on a

$$(11) \qquad a_h^+ + a_h^- = \tilde{h}\ ,\qquad i(a_h^+-a_h^-) = i(\tilde{h}-2\nabla_h)\ (\text{cf. } (9))$$

Les formules (11) justifient l'assertion faite au §I, (6) : Q_h et P_h étaient de bons candidats pour avoir des extensions a.a. : Q_h en tant qu'opérateur de multiplication, P_h en vertu de (9). On notera que si $h=I_{[0,t]}$, $Q_h=Q_t$ est simplement l'opérateur de multiplication de Wiener par B_t. D'autre part, e^{iQ_h} représente l'opérateur de multiplication de Wiener par $e^{i\tilde{h}}$. D'après (7) et la formule (13) du §I, on a

$$(12) \qquad e^{iQ_h}\mathcal{E}(g) = e^{-\|h\|^2/2}\mathcal{E}(ih)\mathcal{E}(g) = e^{i\int h_s g_s ds - \|h\|^2/2}\mathcal{E}(g+ih) = W_{ih}\mathcal{E}(g)$$

Jusqu'à maintenant, nous avons travaillé sur l'espace $\mathcal{E}$ des combinaisons linéaires de vecteurs cohérents. Est-ce que cela détermine les opérateurs par fermeture ? Par exemple : si $f\in\mathcal{D}(Q_h)$ ($f\in L^2$, $\tilde{h}f\in L^2$) existe t'il des $f_n\in\mathcal{E}$ tels que $f_n\to f$, $\tilde{h}f_n\to\tilde{h}f$? La réponse est oui, mais même dans ce cas simple, il faut réfléchir un peu : on se ramène par troncation à f bornée, on choisit des $f_n\in\mathcal{E}$ tendant vers f dans L^4 (par ex.), et on utilise le fait que $\tilde{h}\in L^4$ de sorte que $\tilde{h}f_n\to\tilde{h}f$ dans L^2.

COMMENTAIRES. a) Il existe un jeu consistant à regarder un dessin, et à y découvrir sept visages placés à l'envers dans les arbres, la rivière et le panier du pêcheur à la ligne. Il en va de même ici pour l'espace du mouvement brownien : partant du processus $B_t = Q_t$, qui est un objet bien familier, nous ne savions pas qu'il y avait, caché dans le paysage, un second mouvement brownien (P_t) exactement semblable.

La transformation qui permet de retourner le dessin est classique : c'est la <u>transformation de Fourier-Wiener</u> : étant donnée une v.a. g développée suivant les chaos de Wiener comme $g=\Sigma_n g_n$, on a $\mathcal{F}g = \Sigma_n i^n g_n$. Ou encore, $\mathcal{F}(\mathcal{E}(h))=\mathcal{E}(ih)$: c'est un isomorphisme de $L^2(\Omega)$ qui transforme Q_h en P_h , P_h en $-Q_h$. Cf. Hida, Brownian Motion, p. 182-184.

b) On notera, dans la discussion qui précède, le rôle particulier joué par les éléments réels de $\mathcal{H}$: l'espace de Wiener est un espace de trajectoires

c) Cela explique sans doute pourquoi le produit de Wiener est bien plus utilisé que les autres. Un résultat important affirme que, si A est une contraction de l'espace $\mathcal{H}$ (préservant le sous-espace réel), son extension $\Gamma(A)$ à l'espace de Fock préserve la positivité au sens de Wiener. Surgailis a étudié le problème analogue pour la positivité au sens de Poisson : il faut pour cela imposer à A des conditions beaucoup plus fortes. Sur tout cela, on pourra consulter B. Simon, the $P\phi_2$ euclidean quantum field theory, Princeton Univ. Press 1974, et D. Surgailis, On multiple Poisson stochastic integrals and associated Markov semi-groups. Prob. and M. Stat. 3, 1984. Voir aussi le travail d'exposition de J. Ruiz de Chavez, Sém. Prob. XIX, LN in M. 1123.

3. Interprétation brownienne des opérateurs $a^{\pm}$

Etant donné un élément h de $\mathcal{H}=L^2(\mathbb{R}_+)$, nous désignerons par $\underset{\sim}{h}$ la fonction de Cameron-Martin $\int_0^{\cdot} h_s ds$, et par $\tilde{h}$ la v.a. $\int_0^{\infty} h_s dB_s$. Dans tout ce n°, et sauf mention du contraire, nous supposons h réelle.

La formule de Cameron-Martin-Girsanov nous dit que l'espace de Wiener admet des translations par les fonctions de Cameron-Martin, qui transforment la mesure de Wiener en une mesure équivalente. Plus précisément, voici un groupe à un paramètre de transformations préservant la mesure

$$G_t f(\omega) = f(\omega + t\underset{\sim}{h})\exp(-t\tilde{h}(\omega) - \frac{t^2}{2}\|h\|^2) \qquad \text{(th. de Girsanov)}$$

d'où l'on déduit un groupe de transformations unitaires de $L^2(\Omega)$

$$T_t f(\omega) = f(\omega + t\underset{\sim}{h})\exp(-\frac{t}{2}\tilde{h}(\omega) - \frac{t^2}{4}\|h\|^2) \tag{8}$$

On a par un calcul explicite

$$T_{-2t}\mathcal{E}(g) = \exp(-t\int g_s h_s ds - \frac{t^2}{2}\|h\|^2)\mathcal{E}(g+th)$$

ce qui, comparé à (13), identifie T_{-2t} comme l'opérateur de Weyl $W_{th,I}$ (dans ce n°, les opérateurs de Weyl seront liés à des translations pures, de sorte que nous omettons désormais la mention de I).

Il est facile d'identifier le générateur du groupe à un paramètre T_t : nous désignons par ∇_h l'opérateur de dérivation suivant $\underset{\sim}{h}$, par $\tilde{h}$ l'opérateur de multiplication (de Wiener) par la v.a. $\tilde{h}$. Alors le générateur de (T_t) est $\nabla_h - \frac{1}{2}\tilde{h}$, et par conséquent

$$W_{th} = e^{-itP_h} \qquad \text{où} \quad P_h = i(\tilde{h} - 2\nabla_h) \tag{9}$$ (1)

Les opérateurs W_h pour h réelle commutent tous entre eux (§I, (14)), donc nous pouvons considérer les opérateurs a.a. P_h comme des v.a. au sens classique. Munissons l'espace de Fock de sa loi naturelle ε_1 , et cherchons

1. Nos conventions concernant l'opérateur de Weyl n'étaient pas les mêmes en dimension finie et infinie (cf. formule (14) du §I), c'est pourquoi nous avons un signe - dans la formule de gauche. On verra que $W_{ith} = e^{itQ_h}$.

réelles, et n'admet pas de translations complexes. Si h=u+iv, on peut bien définir formellement $\overline{V}_h=\overline{V}_u+i\overline{V}_v$, mais cela ne correspond à rien de concret. Il n'en va pas de même dans l'interprétation probabiliste construite par Segal au moyen du mouvement brownien complexe(1).

c) Nous avons $E[P_h^2]=\|h\|^2=E[Q_h^2]$, donc $(P_h/\|h\|,\ Q_h/\|h\|)$ est un couple canonique en état d'incertitude minimale ($\hbar^2/4=1$ dans cet exposé). Il résulte de la discussion de l'exposé III que si l'on fait une transformation canonique simple

$$U_h = aP_h+bQ_h\ ,\ V_h = cP_h+dQ_h\ ,\ a,b,c,d,h \text{ réels, } ad-bc=1$$

on obtiendra toutes sortes de lois gaussiennes quantiques, mais on ne sortira jamais du cas d'incertitude minimale. On a par exemple

$$e^{ibQ_h} = W_{ibh}\ ,\ e^{iaP_h} = W_{-ah}\ ,\ e^{itU_h}=W_{(-a+ib)th}$$

ces opérateurs commutent tous entre eux, donc on peut considérer les U_h comme formant un espace de v.a. classiques, gaussiennes sous la loi naturelle ε_1 , avec la loi

$$E[e^{iU_h}] = e^{-(a^2+b^2)\|h\|^2/2}$$

On obtient donc des mouvements browniens de variances arbitrairement grandes ou petites, et de même pour les mouvements browniens conjugués V_h, mais en conservant toujours le discriminant de la forme quadratique figurant dans la << loi jointe >>, qui reste égal à 1 .

Cela amène à un problème très intéressant : peut on trouver des lois quantiques (non pures) sous lesquelles les processus (P_t,Q_t) soient des mouvements browniens en état d'incertitude non minimale ? La réponse, contrairement à ce qu'on a vu en dimension finie, est NON : cette construction est possible, mais non sur l'espace de Fock. On y reviendra.

> Dans un langage abstrait, cela signifie que la C^*-algèbre des relations de commutation (i.e. la C^*-algèbre engendrée par les W_z) admet des représentations qui ne sont pas des sommes directes de copies de la représentation de Fock : le théorème de Stone-von Neumann n'est plus vrai en dimension infinie. Le raisonnement que nous avons employé en dimension finie pour revenir sur le modèle est alors en défaut.

4. Opérateurs nombres de particules ; constructions de processus de Poisson

L'opérateur nombre de particules N est (au signe près) le laplacien naturel sur l'espace de Wiener, rendu familier aux probabilistes par le << calcul de Malliavin >> : si $f = \Sigma_n f_n$ est le développement de $f\in L^2$ suivant les chaos de Wiener,

$$(13)\qquad f\in\mathcal{D}(N) \iff \Sigma_n\ n^2\|f_n\|^2 < \infty\ ,\ \text{et}\ \ Nf = \Sigma_n\ nf_n\ .$$

Les opérateurs $P_t=e^{-tN}$ forment le semi-groupe d'Ornstein-Uhlenbeck sur

1. The Complex-Wave representation of the free Boson field. Topics in Funct. Analysis, Adv. in M. Supplementary Studies vol. 3. Academic Press 1978.

l'espace de Wiener (cf. par exemple Sém. Prob. XVI, LN in M. 920, p. 95-132 : là bas L=N/2). On a sur le développement de Wiener $f=\Sigma_n f_n$

$$P_t f = \Sigma_n e^{-nt} f_n \quad ; \quad P_t \mathcal{E}(h) = \mathcal{E}(e^{-t}h)$$

On reconnaît là l'opérateur $\Gamma(A)$ associé à la contraction $A=e^{-t}I$ (§I, (16)), d'où l'on tire aussi que $N=\lambda(I)$ (§I, (17)). Signalons la formule de Mehler (Sém. Prob. XVI, p.96) - dont nous ne nous servirons pas

$$P_t f(\omega) = \int f(e^{-t}\omega + (1-e^{-2t})^{1/2} w)\mu(dw)$$

où μ est la mesure de Wiener.

A côté de l'opérateur N, nous introduirons toute une famille d'opérateurs N_t (nombres de particules jusqu'à l'instant t) et même N_α , qui seront formellement des intégrales stochastiques $\int \alpha_s dN_s$: N_t est l'opérateur nombre de particules sur l'espace de Fock associé à $L^2([0,t])$, prolongé naturellement au gros espace de Fock. Voici les définitions précises.

Soit α une fonction borélienne réelle. Considérons l'opérateur unitaire sur $L^2(\mathbb{R}_+)$

$$\lambda_\alpha \cdot f = e^{i\alpha} f \qquad (\lambda(\alpha)\lambda(\beta) = \lambda(\alpha+\beta))$$

(ceci correspond au choix d'une phase arbitraire en chaque point, et peut, du point de vue physique, être interprété comme une << transformation de jauge >> , d'où le nom de << gauge operators >> utilisé par Hudson-Parthasarathy pour les N_α). Les opérateurs $\Gamma(\lambda_\alpha)$ sont des opérateurs de Weyl sans terme de translation (§I, (15) et (13)), les opérateurs $\Gamma(\lambda_{i\alpha})$ forment un groupe unitaire fortement continu, que nous noterons $e^{i\alpha N_\alpha}$. En particulier, si $\alpha=1$ nous obtenons N , si $\alpha=I_{[0,t]}$ nous obtenons N_t . Nous avons $N_\alpha=\lambda(m_\alpha)$ (§I,(19)), m_α étant l'opérateur de multiplication par α dans $L^2(\mathbb{R}_+)$. Cet opérateur est borné sur chaque chaos H_n si et seulement si α est bornée. Notons la formule utile (et facile à vérifier)

$$(14) \qquad < \mathcal{E}(f), N_\alpha \mathcal{E}(g) > = (\int \alpha_s \bar{f}_s g_s ds) < \mathcal{E}(f), \mathcal{E}(g) > \qquad (\alpha \text{ bornée }).$$

Nous arrivons à l'un des résultats les plus frappants (pour un probabiliste) de Hudson-Parthasarathy : comment on construit un processus de Poisson compensé en ajoutant, à un mouvement brownien, un processus p.s. nul ! La somme d'opérateurs a.a. non bornés ne commutant pas (15) devra être proprement définie, c'est là l'essentiel de la démonstration.

THEOREME. Soit ϕ une fonction réelle, appartenant à $L^2(\mathbb{R}_+)$ (ou plus généralement à L^2_{loc}). Les opérateurs a.a.

$$(15) \qquad \Pi^\phi_t = N_t + \int_0^t \phi_s dQ_s$$

commutent tous. Considérés comme v.a. sous la loi naturelle ε_1, ils forment un processus de Poisson compensé d'intensité $\phi_s^2 ds$, à sauts unité.

Dém. Supposons ϕ bornée, et traitons un cas un peu plus général en remplaçant $I_{[0,t]}$ par α, fonction bornée à support compact. Introduisons un groupe à un paramètre de déplacements, indexé par $u\in\mathbb{R}$, et les opérateurs de Weyl correspondants $W(u)$

$$\Theta(u).f = e^{iu\alpha}f + \phi(e^{iu\alpha}-1)$$

$$W(u)\varepsilon(f) = \exp[\int(e^{iu\alpha_s}-1)\phi_s f_s ds + (\cos(u\alpha_s)-1))\phi_s^2 ds]\varepsilon(\Theta(u)\cdot f)$$

(§I, (13)). Ces opérateurs ne forment pas un groupe unitaire, mais en les multipliant par un facteur de phase $\exp[i\int\phi_s^2\sin(u\alpha_s)ds]$, on obtient un tel groupe

$$K(u)\varepsilon(f) = \exp[\int(e^{iu\alpha_s}-1)(\phi_s f_s+\phi_s^2)ds]\varepsilon(\Theta(u)\cdot f) \tag{16}$$

Donc l'opérateur $\frac{1}{i}\frac{d}{du}K(u)|_0 = C_\alpha^\phi$ est a.a.. En effectuant la dérivation, on trouve que sur les vecteurs cohérents

$$C_\alpha^\phi = N_\alpha + \int\alpha_s\phi_s dQ_s + (\int\alpha_s\phi_s^2 ds)I \quad . \tag{17}$$

La remarque suivante simplifie le calcul : on introduit les opérateurs de Weyl (cf. (12))

$$\exp(iuN_\alpha) = W_{\lambda(u)},\ \lambda(u)\cdot f = e^{iu\alpha}f\ ;\ \exp(iuQ_{\phi\alpha}) = W_{\mu(u)}$$
$$\mu(u)\cdot f = f+iu\phi\alpha\ .$$

Alors $W_{\lambda(u)}W_{\mu(u)} = W_{\lambda(u)\mu(u)}$ (I,(14)), où

$$\lambda(u)\mu(u)\cdot f = e^{iu\alpha}f + iue^{iu\alpha}\alpha\phi$$

ne diffère de $\Theta(u)\cdot f$ que par un terme du second ordre en u. Ainsi

$$\frac{1}{i}\frac{d}{du}W_{\Theta(u)}|_0 = \frac{1}{i}\frac{d}{du}W_{\lambda(u)\mu(u)}|_0 = N_\alpha + Q_{\alpha\phi}$$

sur les vecteurs cohérents. Le facteur de phase introduit pour former K(u) donne alors le dernier terme de (17).

On remarque ensuite que tous les opérateurs K(u) correspondant à une même fonction ϕ , mais à des fonctions α arbitraires, commutent. Il en résulte que les opérateurs C_α^ϕ peuvent être considérés comme des v.a. classiques, dont la fonction caractéristique sous la loi ε_1 est

$$E[\exp(iuC_\alpha^\phi)] = <1, K(u)1> = \exp[\int e^{iu\alpha_s}-1)\phi_s^2 ds]$$

Si nous prenons $\alpha=I_{[0,t]}$, nous trouvons une loi de Poisson de moyenne $\int_0^t\phi_s^2 ds$, ce qui nous donne pour (15) une loi de Poisson compensée. En faisant varier α, il est facile de vérifier que (15) est en fait un processus de Poisson compensé.

REMARQUES. Les divers processus de Poisson (15) correspondant à des intensités différentes ne commutent pas. Leurs commutateurs se déduisent aisément des relations

$$[N_\alpha, Q_h] = i\int\alpha_s h_s dP_s \quad , \quad [N_\alpha, P_h] = -i\int\alpha_s h_s dQ_s \quad . \tag{18}$$

La transformation unitaire de Fourier-Wiener transforme Q_h en P_h , P_h en $-Q_h$, et préserve N (et le vecteur-vide **1**). Elle transforme donc les processus $N_t+\int_0^t \phi_s dQ_s$ en $N_t+\int_0^t \phi_s dP_s$, et ceux-ci ont donc les mêmes lois sous ε_1 .

5. Interprétation poissonnienne.

Supposons que la fonction ϕ ne s'annule jamais, et posons $\rho=1/\phi$. Alors le processus stochastique

$$Y_t = \int_0^t \frac{1}{\phi_s} d\Pi_s^{\phi} \tag{19}$$

réalise explicitement sous la loi ε_1 , une martingale de carré intégrable, de crochet $\langle Y,Y\rangle_t=t$, et du type décrit en (4).

Cependant, cela ne suffit pas tout à fait : nous avons bien construit sur l'espace de Fock un processus d'opérateurs (19) qui sous la loi ε_1 a la loi demandée, mais nous n'avons pas vérifié que celui-ci engendre l'espace de Fock. Autrement dit, il s'agit de vérifier ceci : si nous partons d'une martingale X sur $(\Omega,\mathcal{F},P)$ ayant la loi (4) et engendrant la tribu $\mathcal{F}$, et si nous identifions Φ à $L^2(\mathcal{F})$, est ce que l'opérateur (19) se lit comme l'opérateur de multiplication par la v.a. X_t ?

Nous avons calculé plus haut en (16) et (17) l'effet de l'opérateur e^{iuY_t} sur un vecteur cohérent $\mathcal{E}(f)$ (nous supposerons $f\phi$ intégrable)

$$e^{iuY_t}\mathcal{E}(f) = \exp(-iu\int_0^t \alpha_s\phi_s^2 ds)\exp(iuC_\alpha^\phi)\mathcal{E}(f) \quad \text{avec } \alpha=\phi^{-1}I_{[0,t]}$$
$$= \exp[\int_0^t (e^{iu/\phi_s}-1)(f_s\phi_s+\phi_s^2)ds - iu\phi_s ds]\mathcal{E}(e^{iu\alpha}f + (e^{iu\alpha}-1)\phi)$$

Dans l'interprétation probabiliste explicite au moyen de la martingale X, nous avons d'après (6)

$$\mathcal{E}(f) = \exp(-\int f_s\phi_s ds)\prod_{s\in S}(1+f_s/\phi_s)$$

et l'exponentielle de Doléans du côté droit se calcule de même, ce qui donne pour le côté droit tout entier

$$\{\exp(-iu\int_0^t \phi_s ds)\prod_{s\in S_t} e^{iu/\phi_s}\}\mathcal{E}(f) \ , \ S_t=S\cap[0,t] \ .$$

Dans le facteur { } on reconnaît bien $e^{iuX_t} = \exp[\sum_{s\in S_t}\frac{1}{\phi_s} - \int_0^t \phi_s ds]$.

6. Expression " algébrique " des divers produits. (Heuristique).

Qu'y a t'il de commun aux diverses interprétations probabilistes de l'espace de Fock ? Dans la théorie algébrique du §I, nous avions développé $\mathcal{H}$ dans une base orthonormale discrète (e_i), tandis que maintenant nous le développons dans une << base continue >>

$$\mathcal{H} = \int_0^\infty f_s dX_s \quad \text{où la norme de cette i.s. est } (\int|f_s|^2 ds)^{1/2}$$

Autrement dit, les éléments de base formels sont les $e_t=dX_t/\sqrt{dt}$.

Un élément f de l'espace de Fock se développe en une série d'intégrales stochastiques multiples

$$f = f_0 + \Sigma_{n\geq 1} J_n(f_n) \quad , \quad f_n \in L^2(C_n) \text{ (le quadrant chronologique cf. n°1 de ce paragraphe)}$$

$$(20) \qquad J_n(f_n) = \int_{s_1<\ldots<s_n} f(s_1,\ldots,s_n)dX_{s_1}\ldots dX_{s_n} \in \mathcal{H}_n$$

$$\|J_n(f_n)\|^2 = \int_{s_1<\ldots<s_n} |f(s_1,\ldots,s_n)|^2 ds_1\ldots ds_n$$

Connaissant ce développement, on peut calculer E[f], qui est simplement la constante f_0 . Mais par ailleurs rien n'indique quelle est la martingale de carré intégrable (X_t) utilisée. Il s'agit simplement, encore une fois, d'une représentation du n-ième chaos $\mathcal{H}_n$ au moyen d'une "base continue" e_A indexée par l'ensemble des parties finies à n éléments A de $\mathbb{R}_+$, identifiées aux n-uples croissants $\{s_1<\ldots<s_n\}$: formellement, $e_A = dX_{s_1}..dX_{s_n}/\sqrt{ds_1\ldots ds_n}$, ou en notation plus condensée $dX_A/\sqrt{dA}$, tandis que $f(s_1,..,s_n)^n = f(A)$ se lit $\langle f, e_A \rangle/\sqrt{dA}$. Voir la remarque page suivante.

On est obligé de préciser la martingale X lorsque l'on veut <u>multiplier</u> entre eux les éléments des chaos. Pour expliquer comment cela se fait, prenons deux éléments du premier chaos , $\int f_s dX_s$ et $\int g_t dX_t$. Formellement, leur produit devrait s'écrire

$$(21) \qquad \int_{s<t} f_s g_t dX_s dX_t + \int_{s=t} f_s g_t dX_s dX_t + \int_{s>t} f_s g_t dX_s dX_t$$

Le premier terme ne pose aucun problème : c'est un élément du second chaos. Il en est de même du dernier, si on l'écrit $\int_{t<s} g_t f_s dX_t dX_s$ (cela suppose implicitement que $dX_t dX_s = dX_s dX_t$). Reste le terme du milieu : c'est là que se voit la martingale : dans le cas brownien, on a

$$(22_a) \qquad dX_t^2 = dt \quad (1)$$

et dans le cas poissonnien, avec saut de hauteur ρ_t à l'instant t

$$(22_b) \qquad dX_t^2 = dt + \rho_t dX_t$$

comme on le voit en écrivant que $dX_t = \rho_t d\nu_t - \frac{1}{\rho_t} dt$, où ν est un processus de Poisson d'intensité dt/ρ_t^2 : si ρ est bornée inférieurement, les intégrales stochastiques sont des intégrales de Stieltjes ; sur la diagonale on a simplement $\int f_s g_s \rho_s^2 d\nu_s = \int f_s g_s \rho_s [dX_s + \frac{1}{\rho_s} ds]$. Un peu de réflexion (qui n'a pas vraiment besoin ici d'être rendue tout à fait rigoureuse) montrera que les règles (22), plus l'associativité et la commutativité, permettent de multiplier les éléments de deux chaos d'ordre quelconque, de la même manière que ci-dessus pour le chaos d'ordre 1. Nous allons nous intéresser surtout au produit de Wiener.

1. On peut remarquer aussi que le produit de Wick correspond à $dX_t^2 = 0$.

REMARQUE. Les notations e_A, dX_A sont faites pour suggérer une analogie formelle avec le modèle discret de la fin de l'exposé II : si (B_t) est un mouvement brownien, P. Lévy écrivait déjà $B_t = \int_0^t \mathrm{sgn}(dB_t)\sqrt{dt}$, ce qui suggère que $dB_t/\sqrt{dt} = \mathrm{sgn}(dB_t)$ est une v.a. de Bernoulli.

Voici un petit dictionnaire pour passer du discret au continu. Il n'est justifié par aucun raisonnement rigoureux, mais il est utile. On imagine des sites le long de $\mathbb{R}_+$, aux points de la forme kdt, le k-ième site (au point t) portant une v.a. de Bernoulli symétrique

$$X_k = dX_t/\sqrt{dt}$$

Reprenons tout le langage du modèle discret : q_k , qui est l'opérateur de multiplication par X_k , se lit $dQ_t/\sqrt{dt}$; p_k se lit $dP_t/\sqrt{dt}$. On convient de traduire

(23) a_k^- par $da_t^-/\sqrt{dt}$, a_k^+ par $da_t^+/\sqrt{dt}$, $N_k=a_k^+a_k^-$ par dN_t

Les formules de l'exposé II deviennent alors :

(24) $da_t^+da_t^-=dN_tdt$ (=0 au premier ordre), $da_t^-da_t^+=dt-dN_tdt$ (=dt au premier ordre)

(25) $[\frac{da_t^-}{\sqrt{dt}} , \frac{da_t^+}{\sqrt{dt}}] = I-2dN_t$ (= I au premier ordre).

(26) $[\frac{dP_t}{\sqrt{dt}} , \frac{dQ_t}{\sqrt{dt}}] = \frac{2}{i}(I-2dN_t)$ ($= \frac{2}{i}I$ au premier ordre ; $\hbar=2$!)

Les relations au premier ordre seront correctes en calcul stochastique, et nous donneront les formules d'Ito ; les relations au second ordre sont plus douteuses, mais (24), par exemple, est susceptible d'une interprétation raisonnable, que nous verrons plus tard. Enfin, notons que les opérateurs $(dQ_t/\sqrt{dt}, dP_t/\sqrt{dt}, I-2dN_t)$ constituent formellement un trio de Pauli.

Pour les processus de Poisson $X_t=Q_t+\rho N_t$, le passage du discret au continu à partir de $X_k=q_k+cN_k$ (formule (8) de l'exposé II) se fait en remplaçant comme ci-dessus X_k par $dX_t/\sqrt{dt}$, N_k par dN_t , et en prenant $c=\rho/\sqrt{dt}$. La formule $X_k^2=1+cX_k$ devient $dX_t^2=dt+\rho dX_t$ comme il se doit.

La principale conséquence de l'analogie discrète concerne les <u>produits</u>. Nous avions défini sur le modèle discret une multiplication associative et commutative, donnée par la table

$$e_Ae_B = e_{A\Delta B}$$

Si on lit e_A comme $dX_A/\sqrt{dA}$ (si $A=\{s_1,\ldots,s_n\}$, $e_A=dX_{s_1}\ldots dX_{s_n}/\sqrt{ds_1\ldots ds_n}$) on voit apparaître la règle de multiplication

(27) $dX_AdX_B = dX_{A\Delta B}\, d(A\cap B)$ (on écrit dA pour $\lambda(dA)$ sur $\mathbb{P}$)

qui est la généralisation de (22_a), et décrit le produit de Wiener : pour multiplier $dX_{s_1}\ldots dX_{s_m}$ par $dX_{t_1}\ldots dX_{t_n}$, on sépare les éléments $u_1,..,u_k$ communs aux deux ensembles $\{s_1,..,s_m\}$ et $\{t_1,\ldots,t_n\}$,

$$s_{i_1}=t_{j_1}=u_1 \ , \ \ldots \ s_{i_k}=t_{j_k}=u_k$$

et on remplace $dX_{s_1} dX_{t_1}$ par du_1,... $dX_{s_k} dX_{t_k}$ par du_k (c'est la signification du d(A∩B) dans (27)). Nous donnerons plus bas la formule de multiplication des intégrales stochastiques, qui exprime cela de manière rigoureuse.

Mais surtout, nous avions défini sur le modèle discret une autre multiplication, associative, mais non commutative : la multiplication de Clifford

$$e_A e_B = (-1)^{n(A,B)} e_{A\Delta B}$$

(exposé II, § III, formule (19)) , qui indique la possibilité de définir sur l'espace de Wiener un produit de Clifford continu. Or nous avions vu, sur le modèle discret, la relation entre le produit de Clifford et les systèmes de spins qui anticommutent, i.e. les systèmes finis de fermions. Ici nous arriverons à rapprocher étroitement les espaces de Fock symétrique et antisymétrique, les bosons et les fermions. Ceci vient d'être réalisé par Hudson et Parthasarathy, par une autre méthode (utilisant le calcul stochastique ; cf. exposé V).

7. Formule de multiplication de Wiener.

Nous avons défini au début du paragraphe l'intégrale stochastique étendue à $\mathbb{R}_+^m$

$$I_m(f) = \int_{\mathbb{R}_+^m} f(s_1,\ldots,s_m) dX_{s_1} \ldots dX_{s_m}$$

Lorsque f est symétrique, cette intégrale est égale à $m! \int_{C_m} f(s_1..s_m) dX_{s_1} \ldots X_{s_m}$ (ou $m! \int_{\rho_m} f(A) dX_A$ dans la notation du n° précédent) ; si f^m n'est pas symétrique, on a simplement $I_m(f) = I_m(\mathbf{s}(f))$, où $\mathbf{s}(f)$ est la symétrisée de f. La symétrisation diminuant la norme L^2, on a toujours $\|I_m(f)\|^2 \leq m! \|f\|^2$.

Soient f et g deux fonctions appartenant respectivement à $L^2(\mathbb{R}^m)$ et $L^2(\mathbb{R}^n)$, et soit p un entier $\leq m \wedge n$; nous définissons une "contraction" $f \underset{p}{=} g$, qui est une fonction sur $\mathbb{R}^{m+n-2p}$:

$$(28) \quad f \underset{p}{=} g(s_1,\ldots,s_{m-p},t_1,\ldots t_{n-p}) = \int f(s_1,..,s_{m-p},u_1,..,u_k) \, g(u_k,\ldots,u_1,t_1,\ldots t_{n-p}) du_1 \ldots du_k$$

L'ordre des indices est ici indifférent, mais ce sera le bon ordre pour le cas antisymétrique. On remarquera que, si f et g sont symétriques, $f \underset{p}{=} g$ n'est pas symétrique en général.

Plus généralement, soient α, β deux injections de $\{1,\ldots,p\}$ dans $\{1,\ldots,m\}$ et $\{1,\ldots,n\}$ respectivement ; en remplaçant $s_{\alpha(i)}$ et $t_{\beta(i)}$ par u_i et en intégrant par rapport à $du_1 \ldots du_p$ (et en numérotant de 1 à m+n-2p les variables restantes, de la gauche vers la droite) on définit une contraction $f \underset{\alpha\beta}{=} g$ (qui nous servira plus bas en (31) seulement).

Voici la formule de multiplication des intégrales de Wiener . Cette formule est classique, et il en existe de nombreuses démonstrations. Voir par ex. Shigekawa, J.M. Kyoto 20, 1980, p.276, Lemme 4.1.

THEOREME. Soient f et g deux fonctions symétriques sur $\mathbb{R}_+^m$ et $\mathbb{R}_+^n$ respectivement. On a

$$I_m(f)I_n(g) = \Sigma_{p=0}^{m\wedge n}\ p!\binom{m}{p}\binom{n}{p} I_{m+n-2p}(f\overset{p}{=}g) \tag{29}$$

REMARQUE. Le cas où m=1 mérite d'être explicité. Dans ce cas, p ne prend que les valeurs 0,1 et m+n-2p = m+1,m-1. La formule se réduit alors à

$$\tilde{f}I_m(g) = a_f^+ I_m(g)+a_f^- I_m(g) \qquad (\text{cf. (11)}) . \tag{30}$$

Démonstration. Il suffit de démontrer la formule lorsque $f=a^{\otimes m}$, $g=b^{\otimes n}$, car on atteint par polarisation le cas où $f=a_1\circ\ldots\circ a_m$, $g=b_1\circ\ldots\circ b_n$ (produits symétriques) puis, par densité, toutes les fonctions symétriques. Dans ce cas, on a $f\overset{p}{=}g = \langle\bar{a},b\rangle^p a^{\otimes m-p}\otimes b^{\otimes n-p}$. L'intégrale $I_{m+n-2p}(.)$ ne changeant pas par symétrisation, on peut remplacer cela par $\langle\bar{a},b\rangle^p a^{\circ m}\circ b^{\circ n}$.

On sait que $\mathcal{E}(ra)=\Sigma_m\ r^m I_m(a^{\circ m})/m!$, $\mathcal{E}(sb)=\Sigma_n\ s^n I_n(b^{\circ n})/n!$ (cf. §I). La formule (7) nous dit aussi que

$$\mathcal{E}(ra)\mathcal{E}(sb) = \mathcal{E}(ra+sb)e^{rs\langle\bar{a},b\rangle} .$$

On développe l'exponentielle, puis $(ra+sb)^k$ par la formule du binôme, puis on identifie dans les deux membres les coefficients de $r^m s^n$, et on obtient le résultat désiré.

VARIANTES DE LA FORMULE. 1) Ici nous ne supposons pas nécessairement f et g symétriques. On a

$$I_m(f)I_n(g) = \Sigma_{p=0}^{m\wedge n}\ \frac{1}{p!}\Sigma_{\alpha,\beta}\ I_{m+n-2p}(f\overset{\alpha\beta}{=}g)$$

(α,β) parcourant l'ensemble des couples d'injections de $\{1,..,p\}$ dans $\{1,...m\}$, $\{1,...,n\}$. Cette formule est celle de Shigekawa[1], elle s'établit par récurrence en commençant avec (30). Sa signification combinatoire est très simple : elle ne fait qu'exprimer la décomposition de

$$dX_{s_1}\ldots dX_{s_m}dX_{t_1}\ldots dX_{t_n} I_{\{s_1\neq\ldots\neq s_m\}}I_{\{t_1\neq\ldots\neq t_n\}}$$

suivant le nombre p de coïncidences entre un s_i et un t_j.

2) Nous allons exprimer (29) dans le langage du calcul sur les parties finies de $\mathbb{R}_+$. Posons

$$J_m(f) = \int_{s_1<\ldots<s_m} f(s_1,\ldots,s_m)dX_{s_1}..dX_{s_m} = \int_{\mathcal{P}_m} f(A)dX_A = I_m(f)/m!$$

et $J_n(g)$ de même. D'après (29), on a pour $J_m(f)J_n(g)$ l'expression

$$\Sigma_p\ \frac{1}{(m-p)!(n-p)!}\int_{\{s_1\neq..\neq s_{m-p}\neq t_1..\neq t_{n-p}\}} h^p(s_1,..s_{m-p},t_1..,t_{n-p})dX_{s_1}\ldots dX_{s_{m-p}}dX_{t_1}..dX_{t_{n-p}}$$

avec $h^p(\ldots) = \frac{1}{p!}\int f_m(s_1,..,s_{m-p},u_1..,u_p)g_n(u_p,..,u_1,t_1,..t_{n-p})du_1\ldots du_p$

1. Référence en dernière ligne, page précédente.

Posons $H=\{u_1,\dots,u_p\}$, $K=\{s_1,\dots,s_{m-p}\}$, $L=\{t_1,\dots,t_{n-p}\}$; on a

$$h^p(K,L) = \int_{\rho_p} f(H\cup K)g(H\cup L)\,dH \quad , \ |K|=m-p,\ |L|=n-p,\ K\cap L=\emptyset\ .$$

Cette fonction n'est pas symétrique en toutes ses variables. Nous nous permettrons de désigner par la même lettre sa symétrisée, qui est pour $C\in\rho_{m+n-2p}$

$$h^p(C) = \binom{m+n-2p}{m-p}^{-1} \Sigma_{|K|=m-p,\ K+L=C}\ h^p(K,L)$$

où la notation $K+L=C$ signifie (ici et dans la suite) $K\cup L=C$, $K\cap L=\emptyset$.
Il nous reste alors

$$J_m(f)J_n(g) = \Sigma_{p=0}^{m\wedge n} \frac{1}{(m+n-2p)!} \int_{u_1\neq\dots\neq u_{m+n-2p}} h^p(u_1,..,u_{m+n-2p})dX_{u_1}..dX_{m+n-2p}$$

(intégrale d'une fonction symétrique). Le coefficient en tête est juste ce qu'il faut pour réduire l'intégration à ρ_{m+n-2p} . Ainsi

$$\Big(\int_{\rho_m} f(A)dX_A\Big)\Big(\int_{\rho_n} g(B)dX_B\Big) = \Sigma_{p=0}^{m\wedge n} \int_{\rho_{m+n-2p}} h(C)dX_C$$

où pour $|C|=m+n-2p$ on a $h(C)= \int_{\rho_p} dH\ \Sigma_{\substack{K+L=C\\ |K|=m-p}} f(H\cup K)g(H\cup L)$

Mais si l'on considère f (resp. g) comme une fonction sur ρ qui n'est différente de 0 que sur ρ_m (ρ_n)), il n'est pas nécessaire de spécifier du tout les indices, et l'on aboutit à la formule sympathique suivante, valable au moins pour des fonctions f,g nulles hors de la réunion d'un nombre fini de ρ_k

$$(31)\qquad \Big(\int_{\rho} f(A)dX_A\Big)\Big(\int_{\rho} g(B)dX_B\Big) = \int_{\rho} f\times g(C)dX_C$$

$$f\times g(C) = \int_{\rho} dM\ \Sigma_{K+L=C}\ f(M\cup K)g(M\cup L)$$

Cette formule est la version intégrale de (27).

Notons que les formules de multiplication Poissonniennes ont été traitées par D. Surgailis, On Multiple Poisson Stochastic Integrals and Associated Markov Semigroups, Probability and M. Stat 3, 1984, p. 227-231.

3) Nous allons nous livrer à un exercice ennuyeux, mais instructif : celui de vérifier sur la formule (31) un résultat connu a priori, l'<u>associativité</u> du produit de Wiener. Cela rendra plus facile le travail correspondant sur le produit de Clifford.

Il s'agit de comparer les deux quantités

$$(f\times g)\times h(C) = \int dMdN\ \Sigma_{\substack{K+L=N\cup P\\ P+Q=C}} f(M\cup K)g(M\cup L)h(N\cup Q)$$

$$f\times(g\times h)(C) = \int dMdN\ \Sigma_{\substack{K+L=M\cup Q\\ P+Q=C}} f(M\cup P)g(N\cup K)h(N\cup L)$$

Note : des calculs voisins, mais moins maladroits, figurent dans l'exp. V, §III.

Puisque C est fixe, et les mesures dM et dN sont diffuses, on ne modifie pas les intégrales en supposant que M et N sont disjoints, et disjoints de C . On transforme ces expressions respectivement en

$$\int dM dN \sum_{\substack{U+V+W=C \\ T\subset N}} f(M\cup T\cup U)g(M\cup T'\cup V)h(N\cup W) \quad (T'=N\setminus T)$$

$$\int dM dN \sum_{\substack{U+V+W=C \\ S\subset M}} f(M\cup U)g(N\cup S\cup V)h(N\cup S'\cup W) \quad (S'=M\setminus S)$$

Il suffit de vérifier l'égalité pour U,V,W fixés. Posant alors $\underline{f}(.)=f(.\cup U)$, $\underline{g}(.)=g(.\cup V)$, $\underline{h}(.)=h(.\cup W)$, et enlevant les _ ensuite, on se ramène au cas où $U=V=W=\emptyset$. Nous pouvons aussi supposer que f(A),g(A),h(A) ne sont $\neq 0$ que pour $|A|=p,q,r$ respectivement. La première intégrale n'est alors $\neq 0$ que si

$$|M|+|T|=p,\ |M|+|N|-|T|=q\ ,\ |N|=r$$

et il en résulte aisément qu'il existe trois entiers i,j,k tels que

$$r=j+k\ ,\ |T|=j,\ |T'|=k\ ,\ |M|=i\ ,\ p=i+j\ ,\ q=i+k$$

La première intégrale vaut donc

$$(32)\quad \int_{\mathcal{P}_i\times\mathcal{P}_{j+k}} dM dN \sum_{\substack{T+T'=N \\ |T|=j}} f(M\cup T)g(M\cup T')h(N)\ .$$

Nous utilisons le lemme suivant[1]:

LEMME. Pour toute fonction ϕ sur $\mathcal{P}_j\times\mathcal{P}_k$ on a (avec les réserves usuelles d'intégrabilité)

$$(33)\quad \int_{\mathcal{P}_{j+k}} dN \sum_{\substack{T+T'=N \\ |T|=j}} \phi(T,T') = \int_{\mathcal{P}_j\times\mathcal{P}_k} \phi(T,T')dTdT'\ .$$

Démonstration. Il suffit de traiter le cas où $\phi(T,T')=a(T)b(T')$. Le côté gauche vaut

$$\int_{u_1<\ldots<u_{j+k}} \binom{j+k}{j} \tilde{\phi}(u_1,\ldots,u_{j+k})du_1\ldots du_{j+k}$$

où $\tilde{\phi}$ est la symétrisée de $\phi(s_1,\ldots,s_j,t_1,\ldots,t_k)$ (déjà symétrique en les j premières et les k dernières variables). Le numérateur (j+k)! disparaît lorsqu'on intègre sur $\mathbb{R}^{j+k}$ entier au lieu du quadrant chronologique, et on obtient

$$\frac{1}{j!k!}\int \tilde{\phi}(u_1,\ldots,u_{j+k})du_1\ldots du_{j+k}$$

Mais maintenant nous pouvons à nouveau remplacer $\tilde{\phi}$ par ϕ, et comme $\phi=ab$ nous avons $(\frac{1}{j!}\int a(s_1,\ldots,s_j)ds_1..ds_j)(\frac{1}{k!}\int b(t_1,..,t_k)dt_1..dt_k)$, soit $\int_{\mathcal{P}_j\times\mathcal{P}_k} a(T)b(T')dTdT'$, le résultat désiré.

Ce lemme nous permet alors de donner à (32) la forme

$$(34)\quad \int_{\mathcal{P}_i\times\mathcal{P}_j\times\mathcal{P}_k} dAdBdC\ f(A\cup B)g(A\cup C)h(B\cup C)$$

1. On trouvera dans l'exp. V, § III, (7), une forme moins maladroite de ce lemme, avec d'autres applications.

dans laquelle l'association des indices j et k a disparu ; la seconde intégrale se réduirait à la même forme.

8. Multiplication de Clifford continue.

L'étude du "bébé Fock" dans l'exposé II a conduit à la définition d'une multiplication de Clifford, satisfaisant à

$$e_A * e_B = (-1)^{n(A,B)} e_{A\Delta B}$$

Les règles de correspondance entre le bébé Fock et le Fock adulte nous amènent donc à tenter de multiplier les chaos, suivant la formule analogue à (27)

$$(35) \qquad dX_A * dX_B = (-1)^{n(A,B)} dX_{A\Delta B}\, d(A\cap B) \ .$$

Cette fois-ci, nous adopterons la convention antisymétrique du début de l'exposé, autrement dit, nous poserons $\hat{I}_n(f)=n!J_n(f)$ pour f antisymétrique appartenant à $L^2(\mathbb{R}_+^n)$ - le $\hat{}$ ($\wedge$) est là pour éviter des confusions.

a) Nous commençons par expérimenter un peu, pour voir ce que signifie cette règle : nous allons calculer

$$J_1(h)*J_2(k) = (\int h_r dX_r)*(\int_{s<t} k_{st} dX_s dX_t) =$$

$$= \int_{r<s<t} h_r k_{st} dX_r dX_s dX_t - \int_{s<r<t} h_r k_{st} dX_s dX_r dX_t + \int_{s<t<r} h_r k_{st} dX_s dX_t dX_r$$

$$+ \int_{r=s<t} h_r k_{rt} dX_r^2 dX_t - \int_{s<t=r} h_r k_{sr} dX_s dX_r^2$$

$$= \int_{u<v<w} (h_u k_{vw} - h_v k_{uw} + h_w k_{vw}) dX_u dX_v dX_w + \int dX_u \left(\int_0^u h_r k_{ru} dr - \int_u^\infty h_r k_{ur} dr\right)$$

Convenons de désigner encore par k le prolongement antisymétrique de cette fonction à $\mathbb{R}_+^2$. Ce que nous calculons est alors $\frac{1}{2}I_1(h)*I_2(k)$. Nous avons d'autre part

$$h(u)k(v,w)-h(v)k(u,w)+h(w)k(v,w) = 3h\wedge k(u,v,w)$$

donc le premier terme de la dernière ligne vaut $\frac{1}{2}I_3(h\wedge k)$. Dans le second terme, l'intégrale intérieure est simplement $\int_0^\infty h_r k_{ru} dr$: c'est le demi-produit intérieur de h et k. En revenant au § I, formules (2) et (4), nos deux termes sont $\frac{1}{2}I_3(b_h^+ k)$ et $\frac{1}{2}I_1(b_h^- k)$.

Ce fait est général : nous avons travaillé sur le second chaos, mais le cas général n'est pas plus profond (seulement plus lourd à écrire) : nous avons établi que

(36) L'opérateur de multiplication $I_1(h)*$ est égal à $R_h = b_h^+ + b_h^-$.

b) Il faut maintenant transformer la règle de calcul (35) en une expression intégrale. Nous imiterons ce qui a été fait pour le produit de Wiener

en posant

$$(\int_{\rho} f(A)dX_A)*(\int_{\rho} g(B)dX_B) = \int_{\rho} f\times g(C)dX_C$$

(37)

$$f\times g(C) = \int_{\rho} dM \; \Sigma_{K+L=C} (-1)^{n(M\cup K, M\cup L)} f(M\cup K)g(M\cup L)$$

Le résultat principal est le suivant :

THEOREME. L'opération $*$ est associative.

Nous disposons pour démontrer ce théorème de la relation

(39) $$(-1)^{n(A,B\Delta C)}(-1)^{n(B,C)}=(-1)^{n(A,B)}(-1)^{n(A\Delta B,C)}$$

qui équivaut à l'associativité $e_A(e_Be_C)=(e_Ae_B)e_C$ en dimension finie, vue dans l'exposé II.

Nous allons suivre exactement la démonstration d'associativité faite pour le produit de Wiener :

$$(f\times g)\times h(C) = \int dN \; \Sigma_{P+Q=C} (-1)^{n(N\cup P, N\cup Q)}(f\times g)(N\cup P)h(N\cup Q)$$

$$= \int dMdN \; \Sigma_{\substack{K+L=N\cup P \\ P+Q=C}} (-1)^{n(N\cup P, N\cup Q)+n(M\cup K, M\cup L)} f(M\cup K)g(M\cup L)h(N\cup Q)$$

On pose $N=T+T'$, $P=U+V$, $K=U\cup T$, $L=V\cup T'$, $W=Q$. On abrège aussi $(-1)^{n(A,B)}$ en $(-^{A,B})$ pour avoir de la place. On arrive ainsi à

$$(f\times g)\times h(C) = \int dMdN \; \Sigma_{\substack{U+V+W=C \\ T+T'=N}} (-^{N\cup U\cup V, N\cup W})(-^{M\cup U\cup T, M\cup V\cup T'})f(M\cup U\cup T)g(M\cup V\cup T')h(W\cup N)$$

Exactement de la même manière

$$f\times(g\times h)(C) = \int dMdN \; \Sigma_{\substack{U+V+W=C \\ S+S'=M}} (-^{M\cup U, M\cup V\cup W})(-^{N\cup V\cup S, N\cup W\cup S'})f(M\cup U)g(N\cup V\cup S)h(N\cup W\cup S')$$

Il suffit de vérifier l'égalité des intégrales correspondant à U,V,W fixés. On pose $F(A)=f(A\cup U)$, $G(A)=g(A\cup V)$, $H(A)=h(A\cup W)$, ce qui fait disparaître U,V,W des fonctions (mais non des $n(.,.)$). On peut aussi supposer que F,G,H ne sont $\neq 0$ que sur ρ_p,ρ_q,ρ_r respectivement, et écrire $p=i+j$, $q=i+k$, $r=j+k$. La première intégrale à évaluer vaut alors (en appliquant le lemme comme au n° précédent)

$$\int_{\rho_i\times\rho_j\times\rho_k} dMdTdT' \; (-^{U\cup V\cup T\cup T', W\cup T\cup T'})(-^{U\cup M\cup T, V\cup M\cup T'})F(M\cup T)G(M\cup T')H(T\cup T')$$

La seconde intégrale est de la même forme,

$$\int_{\rho_i\times\rho_j\times\rho_k} dSdS'dN \; (-^{U\cup V\cup S'\cup N, W\cup S'\cup N})(-^{U\cup S\cup S', V\cup S\cup N})F(S\cup S')G(S\cup N)H(S'\cup N)$$

Pour rapprocher les deux formules, nous remplaçons dans la première

M par S, T par S', T' par N

et nous sommes alors ramenés à vérifier que

$$(-^{U\cup S\cup S', V\cup W\cup S\cup S'})(-^{V\cup N\cup S, W\cup N\cup S'}) = (-^{U\cup S\cup S', V\cup S\cup N})(-^{U\cup V\cup S'\cup N, W\cup S'\cup N})$$

seconde formule — première formule traduite

et la formule (39) s'applique, avec

$$A=UUSUS' \ , \qquad B=VUSUN \ , \qquad C=WUS'UN \ .$$

(Noter que S,S',N sont p.s. disjoints, et disjoints de U,V,W).

c) Les résultats les plus importants concernent le cas des opérateurs de multiplication de Clifford par les éléments du premier chaos, i.e. les opérateurs $R_h=I_1(h)*$. Comme dans toute algèbre de Clifford, on a

$$\{R_h,R_k\} = \langle \bar{h},k\rangle I \qquad . \tag{40}$$

Ensuite, rappelons la notation $\hat{I}_n(f)=n!J_n(f)$ pour f antisymétrique sur $\mathbb{R}_+^n$, et l'extension de cette notation aux fonctions non nécessairement antisymétriques par la convention $\hat{I}_n(f)=\hat{I}_n(\mathbf{a}(f))$. La relation $R_f=b_f^+ + b_f^-$ permet de demontrer par récurrence le résultat suivant : si (e_i) est une base orthonormale de $L^2(\mathbb{R}_+)$, on a

$$\hat{I}_p(e_{i_1}\wedge\ldots\wedge e_{i_p}) = I_1(e_{i_1})*\ldots*I_1(e_{i_p}) \ . \tag{41}$$

Posons alors pour $A=\{i_1<\ldots<i_p\}$

$$e_A=e_{i_1}\wedge\ldots\wedge e_{i_p} \quad , \quad X_A=\hat{I}_p(e_A) \tag{42}$$

Les e_A forment une base o.n. de l'espace de Fock antisymétrique, et les X_A forment une base orthogonale de l'espace L^2 du mouvement brownien. En utilisant (41), l'associativité du produit de Clifford, et (42), on aboutit à la formule familière

$$X_A*X_B = (-1)^{n(A,B)}X_{A\Delta B} \tag{43}$$

(mais A,B sont des parties de $\mathbb{N}$, non de $\mathbb{R}_+$!).

d) Dans le cas symétrique, nous sommes passés de la formule de multiplication détaillée (29) à la formule condensée (31). Ici, la formule condensée (37) nous a servi de définition ; on peut faire le passage en sens inverse pour obtenir une formule détaillée, qui (grâce à notre définition des contractions en (28)) est exactement semblable à (29) :

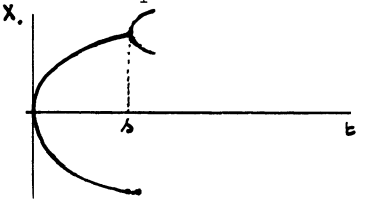
$$\hat{I}_m(f)*\hat{I}_n(g) = \Sigma_{p=0}^{m\wedge n} \, p!\binom{m}{p}\binom{n}{p}\hat{I}_{m+n-2p}(f\tilde{\otimes}_p g) \ . \tag{44}$$

e) Le mouvement brownien des fermions. Si h est un élément réel de $L^2(\mathbb{R}_+)$, l'opérateur R_h est a.a. borné, de norme $\|h\|$: en fait, $\frac{1}{\|h\|}R_h$ est une symétrie de l'espace de Wiener, une v.a. de Bernoulli sous la loi ε_1 puisqu'elle ne prend que les valeurs ± 1 et a une espérance nulle.

En posant $\beta_t=R_{I_{[0,t]}}$, on définit un processus stochastique quantique sans analogue classique (les v.a. à des instants différents ne commutent pas) que l'on appelle le mouvement brownien des fermions. C'est l'un des objets les plus intéressants du calcul stochastique quantique

(Voir par ex. les articles de Barnett, Streater et Wilde, The Ito-Clifford Integral I-IV : J. Funct. An 48, 1982 , J. London M. Soc. 27, 1983 , Comm. M. Phys. 89, 1983 et J. Oper. Thy 11, 1984). Du point de vue probabiliste, c'est simplement le mouvement brownien de tout le monde, mais opérant sur son propre L^2 par multiplication de Clifford au lieu de Wiener.

On peut s'en faire une représentation intuitive de la manière suivante :



Tant qu'on ne l'observe pas, la << particule >> à l'instant t est en résonance entre les points $\pm\sqrt{t}$. Si on la regarde à l'instant s, et que l'on a observé la valeur $+\sqrt{s}$ par exemple, la particule est perturbée, et son << mouvement >> se poursuit sur une parabole de sommet $(s,\sqrt{s})$ et non plus (0,0) (figure ci-dessus). L'opérateur de multiplication de Clifford par $(X_{t+h}-X_t)$ a une norme égale à $\sqrt{h}$, ce qui en probabilités classiques signifierait que le mouvement brownien fermionique est un processus continu en norme uniforme - ce qui ne se produit jamais pour une martingale classique.

L'espace de Fock ne porte pas un seul mouvement brownien fermionique $R_t = b_t^+ + b_t^-$, mais deux , car $S_t = i(b_t^+ - b_t^-)$ en est un aussi . Plus précisément, on peut établir la relation

$$R_h x = h * x \ , \quad S_h x = i\sigma(x) * h \qquad (h \text{ réelle}) \tag{44}$$

où σ est l'automorphisme de l'algèbre de Clifford défini dans l'exposé II, §III, (21) (multiplicateur $(-1)^p$ sur le p-ième chaos) ; à partir de là, en regardant soigneusement ce qui se passe sur chaque chaos (c'est ennuyeux, mais je crois l'avoir fait sérieusement), on vérifie que

$$\mathcal{F} R_h = S_h \mathcal{F} \qquad (h \text{ réelle}, \mathcal{F} \text{ transformée de Fourier-Wiener}) \tag{45}$$

de sorte que la relation entre R_h (m^t brownien fermionique) et Q_h (m^t brownien ordinaire) se trouve exactement reproduite entre S_h et P_h .

L'étude du bébé Fock suggère que les accroissements dR_u et dS_v anticommutent, et que l'on peut plonger le tout dans une vaste algèbre de Clifford. La fin de l'exposé II (formule (23)) suggère aussi l'existence d'algèbres de Poisson non commutatives, mais vraiment, cela suffit.

REMARQUE. Le mouvement brownien fermionique peut être observé à un instant fixe s comme sur le dessin, mais on peut aussi décider de le regarder aux instants 0,1/M,2/M,... Cela donne des courbes aléatoires formées d'arcs de paraboles ; lorsque $M \to \infty$, i.e. lorsqu'on regarde tout le temps le m^t br^n fermionique, on obtient un mouvement brownien ordinaire.

III. COMPLEMENTS DIVERS

Ce paragraphe est un fourre-tout, où l'on a rassemblé divers résultats intéressants sur l'espace de Fock, ne dépendant pas du calcul stochastique quantique de l'exposé V. Ils peuvent être lus maintenant, ou plus tard lorsque le besoin s'en fera sentir.

1. Opérateurs sur l'espace de Fock : forme normale.

Cette section est purement formelle : le contenu sera justifié plus tard de manière correcte. Elle est suggérée par le rapprochement entre le début du livre de Berezin, The Method of Second Quantization (Academic Press 1966 pour la traduction anglaise), et un tout récent preprint de H. Maassen, Quantum Markov Processes on Fock Space Described by Integral Kernels. Il s'agit seulement de montrer pourquoi il s'agit de la même chose dans un système de notations différent.

Nous avons vu plus haut que l'espace de Fock se représente agréablement dans une << base orthonormale continue >> $dX_{s_1}\dots dX_{s_n}/\sqrt{ds_1\dots ds_n}$ ou, en notation condensée, $dX_A/\sqrt{dA}$. Nous allons introduire une notation analogue pour définir des opérateurs sur l'espace de Fock. La notation de Berezin consiste à introduire des opérateurs

$$\Sigma_{n,m} \int K(s_1,\dots,s_m,t_1,\dots,t_n)a^+_{s_1}\dots a^+_{s_m}a^-_{t_1}\dots a^-_{t_n}ds_1..ds_m dt_1..dt_n$$

notation que nous préférons remplacer par des expressions du type << intégrale stochastique >> (K symétrique en les s_i et les t_j séparément)

$$(1)\qquad \Sigma_{n,m} \int K(s_1,\dots,s_m,t_1,\dots,t_n)da^+_{s_1}\dots da^+_{s_m}da^-_{t_1}\dots da^-_{t_n}$$

ou en notation condensée , comme on l'a fait pour les i.s. ordinaires

$$(2)\qquad \int_{P\times P} K(A,B)da^+_A da^-_B$$

Un tel opérateur est dit en forme normale, parce que les a^+ sont tous placés à gauche des a^- .

Que peut signifier une telle notation ? Nous allons nous laisser guider par le << bébé Fock >> dans les calculs formels suivants sur les << éléments de base >> . On doit avoir

$$da^+_A\, dX_L = dX_{A\cup L} \quad \text{si } A\cap L=\emptyset\ ,\ 0 \text{ sinon}$$

(le "vrai" calcul formel concerne $da^+_A/\sqrt{dA}$ et $dX_L/\sqrt{dL}$, le résultat étant $dX_{A\cup L}/\sqrt{dA\cup L}$, mais les facteurs $\sqrt{\dots}$ s'éliminent) . Ensuite

$$da^-_B\, dX_L = dX_{L\setminus B}\, dB \quad \text{si } B\subset L\ ,\ 0 \text{ sinon}$$

Par conséquent, en composant

$$(3)\qquad da^+_A da^-_B\, dX_L = dX_{A\cup(L\setminus B)}dB \quad \text{si } B\subset L,\ A\cap(L\setminus B)=\emptyset,\ 0 \text{ sinon}\ .$$

Maintenant, nous essayons de regrouper cela sous forme intégrale, en calculant

$$(\int_{\rho\times\rho} K(A,B)da_A^+da_B^-)(\int_\rho f(L)dX_L) =$$

$$\int_{\substack{\rho\times\rho\times\rho \\ B\subset L,\ A\cap(L\setminus B)=\emptyset}} K(A,B)f(L)dX_{A\cup(L\setminus B)}dB$$

Nous regroupons les termes correspondant à $A\cup(L\setminus B)=H$: B doit être disjoint de H, A contenu dans H, L s'écrit alors $B\cup(H\setminus A)$. Pour H fixe, la condition de disjonction de B et de H est réalisée dB-p.s.. Ainsi

$$(\int_{\rho\times\rho} K(A,B)da_A^+da_B^-)(\int_\rho f(L)dX_L) = \int_\rho (K\cdot f)(H)\, dX_H$$

(4)

$$K\cdot f(H) = \int_\rho \sum_{A\subset H} K(A,B)f(B\cup(H\setminus A))dB = \int_\rho \Sigma_{U+V=H}\, K(U,B)f(B\cup V)dB\ .$$

Ceci est précisément la formule donnée par Maassen. On remarquera que dans l'intégration en B, A parcourant l'ensemble des parties d'un ensemble fini H fixe, B et A seront presque sûrement disjoints, alors que dans le calcul formel (et sur le bébé Fock) c'était seulement $L\setminus B$ qui devait être disjoint de A. Le résultat est que les représentations de Berezin ou de Maassen << oublient >> l'opérateur N. On peut imposer $K(A,B)=0$ si $A\cap B\neq\emptyset$.

Conformément à nos notations de l'exposé V, nous noterons da_s° l'opérateur dN_s . Nous avons en transposant les résultats du bébé Fock

$$da_B^\circ\, dX_L = dX_L \quad \text{si } B\subset L\ ,\ 0 \text{ sinon}$$

(formule (6) p. II.11 et formule (23) p. IV.24). Par conséquent

$$da_A^+da_B^\circ da_C^-\, dX_L = dX_{A\cup(L\setminus C)}dC \quad \text{si } C\subset L,\ B\subset L\setminus C,\ A\cap(L\setminus C)=\emptyset\ ,\ 0 \text{ sinon}$$

(5)

$$(\int K(A,B,C)da_A^+da_B^\circ da_C^-)(\int f(L)dX_L) = \int K\cdot f(H)\, dX_H \quad \text{où } K\cdot f(H) \text{ vaut}$$

$$\int_\rho \Sigma_{A\subset H,B\subset H,A\cap B=\emptyset}\, K(A,B,C)f(C\cup(H\setminus A))dC = \int \sum_{U+V+W=H} K(U,V,C)f(C\cup V\cup W)dC$$

La formule (4) correspond au cas où $K(A,B,C)=K(A,C)$ si $B=\emptyset$, 0 sinon .

(Suite exposé V, § III).

2. Vecteurs-test sur l'espace de Fock.

Dans le " calcul de Malliavin ", on rencontre naturellement une classe de fonctions-test sur l'espace de Wiener : celle des fonctions qui, pour tout k et tout p, appartiennent au domaine-L^p de l'opérateur N^k. On peut montrer que ces fonctions-test forment une algèbre, et les utiliser pour construire une théorie des distributions sur l'espaces de Wiener (cf. S. Watanabe, Malliavin's Calculus in Terms of Generalized Wiener Functionals, Proceedings of the Bangalore Conf. on Random Fields, LN in C. and I. 49, p. 284). D'un point de vue hilbertien, le seul critère dont on dispose pour vérifier qu'un vecteur donné $f = \Sigma_n f_n$ ($f_n e^{\Phi_n}$) est un vecteur-test nous est fourni par l'inégalité

$$\|f_n\|_p \leq (p-1)^n \|f_n\|_2$$

qui nous dit que, si la série $\Sigma_n c^n \|f_n\|_2$ converge pour tout $c>0$, f est une fonction-test au sens de Malliavin. Ce critère est manifestement grossier, mais en l'absence de meilleure définition, nous dirons que f est un vecteur-test sur l'espace de Fock si $\Sigma_n c^n \|f_n\|^2$ converge pour tout $c>0$ (le remplacement de $\|f_n\|$ par $\|f_n\|^2$ ne change évidemment rien). Par exemple, les vecteurs cohérents sont des vecteurs-test.(1)

Voici un premier résultat :

THEOREME. L'espace des vecteurs-test est stable pour le produit de Wiener.

Remarque. Il ne peut rien y avoir de tel pour le produit de Poisson : Surgailis a montré en effet qu'il existe des éléments du second chaos dont le carré de Poisson est une v.a. qui n'appartient plus à L^2. En revanche, la même démonstration que ci-dessous s'applique au produit de Clifford.

Démonstration. Tout repose sur la formule (29). Nous allons considérer une fonction

$$f = \Sigma_n F_n \quad ; \quad F_n = I_n(f_n) \ , \ f_n \text{ symétrique}, \ \|f_n\|^2_{L^2(\mathbb{R}_+^n)} = \|F_n\|^2/n!$$

où la somme sera pour commencer supposée finie. Nous allons appliquer (29) pour calculer le développement $g = \Sigma_p G_p$ de $g=f^2$, et majorer $\Sigma_p c^p \|G_p\|^2$ en fonction de $\Sigma_n u^n \|F_n\|^2$ pour $u>0$ assez grand, dépendant de c seulement. Le passage aux sommes infinies se fait alors sans difficulté.

Nous avons d'après (29)

$$g = \Sigma_{\substack{m,n \\ k \leq m \wedge n}} k! \binom{m}{k} \binom{n}{k} I_{m+n-2k}(f_m \bar{\bar{k}} f_n)$$

Posons $m=k+\mu$, $n=k+\nu$, $p=m+n-2k=\mu+\nu$. Nous avons alors

$$G_p = \Sigma_k G_{pk} \quad \text{avec} \quad G_{pk} = \Sigma_{\mu+\nu=p} k! \binom{k+\mu}{k} \binom{k+\nu}{k} I_p(f_{k+\mu} \bar{\bar{k}} f_{k+\nu}) \ .$$

A droite, nous avons $\|I_p(f_{k+\mu} \bar{\bar{k}} f_{k+\nu})\|^2 \leq p! \|f_{k+\mu} \bar{\bar{k}} f_{k+\nu}\|^2$ (la symétrisation diminue la norme) $\leq p! \|f_{k+\mu}\|^2 \|f_{k+\nu}\|^2$. Si nous posons $\|F_n\|^2 = a_n$, nous avons donc

$$\|I_p(f_{k+\mu} \bar{\bar{k}} f_{k+\nu})\|^2 \leq \frac{p!}{(k+\mu)!(k+\nu)!} a_{k+\mu} a_{k+\nu} \ .$$

Nous écrivons que $a_n \leq M u^{-n}$, où $u \geq 4$ sera choisi plus tard, et M dépend de u (propriété de fonction-test de f). Nous avons alors

$$\|G_{pk}\| \leq \Sigma_{\mu+\nu=p} \frac{(k+\mu)!(k+\nu)!}{\mu!\nu!k!} M \left(\frac{p!}{(k+\mu)!(k+\nu)!} \right)^{1/2} u^{-k-p/2}$$

Nous écrivons tout ce gros paquet de factorielles sous la forme

---

1. Pourrait on faire du calcul de Malliavin en utilisant ces vecteurs-test ?

$$\left(\frac{(k+\mu)!}{k!\,\mu!}\right)^{1/2}\left(\frac{(k+\nu)!}{k!\,\nu!}\right)^{1/2}\left(\frac{p!}{\mu!\nu!}\right)^{1/2}$$

et nous majorons grossièrement les coefficients binômiaux par $2^{k+\mu}$, $2^{k+\nu}$, 2^p respectivement, ce qui nous laisse un majorant 2^{p+k} pour le tout. Comme il y a p termes dans la somme, nous avons

$$\|G_{pk}\| \leqq pM\left(\frac{2}{\sqrt{u}}\right)^p \left(\frac{2}{u}\right)^k$$

Comme nous avons supposé $u \geqq 4$, nous en déduisons $\|G_p\| \leqq 2pM\left(\frac{2}{\sqrt{u}}\right)^p$, et pour avoir une bonne majoration de $\Sigma_p\, c^p\|G_p\|$ il suffit de prendre u assez grand.

Nous allons maintenant étudier une famille d'opérateurs sur l'espace de Fock, qui laissent stable l'espace des vecteurs-test. Ces opérateurs A sont donnés par des matrices d'opérateurs bornés $A^n_m : \mathfrak{H}_n \longrightarrow \mathfrak{H}_m$, possédant la propriété suivante :

(6) Pour tout j il existe k et C tels que $\|A^n_m\| \leqq Cn^k m^{-j}$.

Vérifions que si h est un vecteur-test, Ah (est bien défini et) est un vecteur test. Pour procéder correctement, nous prenons d'abord h dans la somme non complétée $\oplus_n \mathfrak{H}_n$, avec une majoration de la forme

(7) $\|h_n\| \leqq Kn^{-\alpha}$ (α sera choisi plus loin)

et nous posons $j_m = \Sigma_n\, A^n_m h_n$.

On a $\|j_m\|^2 = \Sigma_{n,p} < A^n_m h_n, A^p_m h_p > \leqq \Sigma_{n,p}\, \|A^n_m\|\|A^p_m\|\|h_n\|\|h_p\|$

$$\Sigma_m\, m^{\ell}\|j_m\|^2 \leqq \Sigma_{m,n,p}\, m^{\ell}(Cm^{-j}n^k)(Cm^{-j}p^k)(Kn^{-\alpha})(Kp^{-\alpha}) \quad ((6) \text{ et } (7))$$

$$= C^2K^2(\Sigma_m\, m^{\ell-2j})(\Sigma_i\, i^{k-\alpha})^2$$

Nous commençons par fixer j assez grand pour que la première somme converge, puis dans (6) nous choisissons C et k en fonction de j, puis dans (7) nous choisissons $\alpha > k+1$. Le prolongement aux vecteurs-test est immédiat.

Plus précisément, l'opérateur A est continu de l'espace des vecteurs-test dans lui-même, pour la topologie associée aux semi-normes $q_\ell(h) = \sup_n n^{\ell}\|h_n\|$. On vérifie sans peine que les opérateurs matriciels du type précédent forment un espace stable par composition. En effet, si j étant donné on peut trouver C et ℓ tels que $\|B^p_m\| \leqq Cp^{\ell}m^{-j}$, puis C' et k tels que $\|A^n_p\| \leqq C'n^{-k}p^{-\ell-2}$, la série d'opérateurs $\Sigma_p\, A^n_p B^p_m$ est convergente et sa norme est majorée par $C''n^k m^{-j}$.

Les opérateurs de la forme a^+_h, a^-_h, N_h sont du type précédent. Par exemple, a^+_h est représenté par une matrice A telle que $A^n_m = 0$ si $m \neq n+1$, et $\|A^n_{n+1}\| \leqq c\sqrt{n}$, donc $\|A^n_m\| \leqq Cn^k m^{-j}$ avec $k = j+1/2$.

Il en résulte que tous les opérateurs polynômiaux en les a_f^+ , a_g^- , N_h sont représentés par des matrices satisfaisant à (6) - mieux, la matrice adjointe satisfait aux mêmes propriétés, et il n'est pas difficile de voir que si A' est l'opérateur (de l'espace des vecteurs-test dans lui même) défini par la matrice adjointe $A^{*n}_{m} = (A^{m}_{n})^{*}$, A et A' sont mutuellement adjoints sur l'espace des vecteurs-test.

ELEMENTS DE PROBABILITES QUANTIQUES. V

Calcul stochastique non commutatif

Avec cet exposé, nous abordons l'essentiel des travaux de Hudson et Parthasarathy, qui est le développement d'un calcul stochastique analogue au calcul d'Ito, mais relatif à des processus qui ne commutent pas, ainsi que la possibilité de résoudre des équations différentielles stochastiques linéaires (du type " exponentielle stochastique ") dont les solutions fourniront ensuite, par projection, des semi-groupes d'évolution.

Contrairement à Hudson-Parthasarathy, nous nous placerons dans la situation la plus simple possible : celle de l'espace de Fock sans adjonction d'aucun espace de Hilbert auxiliaire. C'est que notre but est purement pédagogique, et que nous n'avons aucune raison de rechercher la généralité maximale. Les extensions viendront plus tard.

§ I . INTEGRALES STOCHASTIQUES DE PROCESSUS D'OPERATEURS

1. Filtration de l'espace de Fock

Si f est un élément de $L^2(\mathbb{R}_+)$, nous avons désigné par $f_{t]}$, $f_{[t}$ les fonctions $fI_{[0,t]}$, $fI_{[t,\infty[}$; pour alléger, nous écrirons parfois f_t au lieu de $f_{t]}$ lorsqu'il n'y aura pas de confusion possible avec la valeur de f en t . Φ désignant toujours l'espace de Fock, nous conviendrons de désigner par $\Phi_{t]}$ (ou simplement Φ_t comme ci-dessus) et $\Phi_{[t}$ les sous-espaces engendrés respectivement par les vecteurs cohérents de la forme $\mathcal{E}(f_{t]})$, $\mathcal{E}(f_{[t})$. Si $s \leq t$, on pose $\Phi_{st} = \Phi_{[s} \cap \Phi_{t]}$.

Certaines propriétés purement algébriques de ces espaces peuvent se lire sur une interprétation probabiliste de Φ : pour fixer les idées, considérons l'interprétation brownienne : (X_t) est un mouvement brownien avec $X_0=0$, $\mathcal{F}$ et $\mathcal{F}_t$ sont les tribus engendrées respectivement par toutes les v.a. X_s et par les X_s, $s \leq t$, $\mathcal{G}_t$ est la tribu engendrée par les accroissements $X_u - X_t$ pour $u \geq t$; nous savons que Φ s'identifie alors à $L^2(\mathcal{F})$, et il est clair que $\Phi_{t]}$ et $\Phi_{[t}$ s'identifient respectivement à $L^2(\mathcal{F}_t)$ et $L^2(\mathcal{G}_t)$. Nous en déduisons alors les propriétés hilbertiennes suivantes :

- La famille $(\Phi_{t]})$ est croissante, continue à droite et à gauche, avec $\Phi_0 = \mathbb{C}.1$.
- La famille $(\Phi_{[t})$ est décroissante, continue à droite et à gauche, avec $\Phi_{[\infty} = \mathbb{C}.1$.
- Si $s<t$, Φ_{st} est engendré par les vecteurs cohérents $\mathcal{E}(fI_{[s,t[})$, et si

$s=t$, $\Phi_{ss}=\mathbb{C}.1$.

Le projecteur E_t sur $\Phi_{t]}$ est l'espérance conditionnelle par rapport à $\mathcal{F}_t$, dans l'interprétation brownienne (et dans toute autre interprétation probabiliste).

Soient $u\in\Phi_{t]}$, $v\in\Phi_{[t}$; notons $b(u,v)$, pour un instant, le produit de ces deux v.a. dans l'interprétation brownienne. L'indépendance de $\mathcal{F}_t$ et $\mathcal{G}_t$ entraîne que $\|b(u,v)\|^2=\|u\|^2\|v\|^2$, de sorte que $b(.,.)$ est une application bilinéaire continue de $\Phi_{t]}\times\Phi_{[t}$ dans Φ , satisfaisant à

$$b(\mathcal{E}(f_{t]}),\mathcal{E}(f_{[t}))=\mathcal{E}(f)$$

ce qui d'ailleurs caractérise uniquement b par densité. Il est facile de voir qu'il existe un isomorphisme unique de $\Phi_{t]}\otimes\Phi_{[t}$ sur Φ qui transforme $u\otimes v$ en $b(u,v)$: dans l'interprétation brownienne, cela signifie que la mesure de Wiener peut se représenter comme le produit d'une mesure de Wiener "avant t" par une mesure de Wiener " après t" , ce qui identifie $L^2(\mathcal{F})=\Phi$ à $L^2(\mathcal{F}_t)\otimes L^2(\mathcal{G}_t)$.

En réalité, tout ceci est indépendant du choix d'une interprétation probabiliste particulière : la seule chose qui compte est d'avoir affaire à des processus à accroissements indépendants. C'est pourquoi nous n'écrirons pas $b(u,v)$, ni même $u\otimes v$, mais simplement uv . Si l'on a $s<t$, si u,v,w sont des éléments de $\Phi_{s]}$, Φ_{st}, $\Phi_{[t}$ respectivement, on a $uv\in\Phi_{t]}$, $vw\in\Phi_{[s}$, et $(uv)w=u(vw)$. Tout cela reflète la structure de produit tensoriel continu de l'espace de Fock Φ , notion que nous ne tenterons pas de présenter de manière abstraite.

Opérateurs Φ_t-adaptés.

Nous éviterons, dans les discussions de cet exposé, de regarder de trop près les problèmes de domaines d'opérateurs non bornés. En règle générale, nous considérerons des opérateurs définis sur l'espace (noté $\mathcal{E}$ dans la suite) des combinaisons linéaires finies de vecteurs cohérents - les opérateurs étant prolongés par fermeture à partir de $\mathcal{E}$, chaque fois que cela est possible. Le prix payé pour cette solution de facilité est l'absence d'une bonne notion de composition d'opérateurs.

Nous dirons qu'un opérateur A défini sur $\mathcal{E}$ est Φ_t-adapté (en tant que probabilistes nous dirons aussi Φ_t-mesurable) si

(1) On a $A\mathcal{E}(f_{t]})\in\Phi_t$,

(2) On a $A\mathcal{E}(f)=(A\mathcal{E}(f_{t]}))\mathcal{E}(f_{[t})$,

pour tout $f\in L^2(\mathbb{R}_+)$.

Une famille (A_t) d'opérateurs, telle que pour tout t A_t soit un opérateur Φ_t-mesurable, sera appelée un processus adapté (d'opérateurs).

a) Ces définitions appellent plusieurs commentaires. D'abord, si A est borné, la Φ_t-mesurabilité signifie que dans la représentation $\Phi=\Phi_{t]}\otimes\Phi_{[t}$, A est de la forme $B\otimes I$, où B est un opérateur borné sur Φ_t. Par exemple,

l'opérateur d'espérance conditionnelle E_t n'est pas $\mathfrak{F}_t$-mesurable (en revanche, l'opérateur d'espérance conditionnelle brownien par rapport à la tribu engendrée par les X_s, $s \geq t$, est $\mathfrak{F}_t$-mesurable !)

Nous avons vu de nombreux opérateurs $\mathfrak{F}_t$-mesurables : les opérateurs Q_h, P_h, N_h , $a_h^{\pm}$, lorsque h est à support dans $[0,t]$.

b) Comme pour les opérateurs bornés, on a une définition satisfaisante de la $\mathfrak{F}_t$-mesurabilité des opérateurs autoadjoints : il suffit d'écrire que leurs projecteurs spectraux sont $\mathfrak{F}_t$-mesurables. D'une façon générale, les opérateurs bornés $\mathfrak{F}_t$-mesurables formant une algèbre de vN, il existe des moyens efficaces pour définir et étudier les opérateurs, non nécessairement bornés, << affiliés >> à cette algèbre. Nous n'utiliserons rien de tout cela, et nous en tiendrons à (1)-(2).

2. Espérances conditionnelles

Soit M un opérateur borné. Il existe un opérateur borné $\mathfrak{F}_t$-mesurable N unique satisfaisant à

$$(3) \qquad N(uv) = (E_t M u)v \quad \text{si} \quad u \in \mathfrak{F}_t \ , \ v \in \mathfrak{F}_{[t} \ ,$$

où E_t est, rappelons le, le projecteur orthogonal sur $\mathfrak{F}_t$. Si M est non borné, de domaine contenant $\mathcal{E}$, on se bornera à définir N sur $\mathcal{E}$ par cette formule : $N\mathcal{E}(f) = (E_t M \mathcal{E}(f_{t]}))\mathcal{E}(f_{[t})$.

Montrons que N joue le rôle d'espérance conditionnelle de M par rapport à $\mathfrak{F}_t$, sous la loi naturelle ε_1 . Soient K et H deux opérateurs bornés $\mathfrak{F}_t$-mesurables ; alors

$$(4) \qquad < 1, K^* N H 1 > = < 1, K^* M H 1 > .$$

En effet, posant $u=K1$, $v=H1$, cela s'écrit $< u, Mv > = < u, Nv >$, évident puisque $Nv = E_t Mv$. Il apparaît que la seule propriété du vecteur (normalisé) 1 intervenant dans ce raisonnement est $1 \in \mathfrak{F}_t$. Quant à l'unicité, la formule (4) détermine $<u,Nv>$ pour deux éléments arbitraires u,v de $\mathfrak{F}_t$, et cela détermine l'opérateur $\mathfrak{F}_t$-mesurable N .

Cela permet de définir (sous la loi naturelle) la notion de martingale d'opérateurs : c'est un processus adapté (M_t) d'opérateurs contenant tous $\mathcal{E}$ dans leur domaine, et possédant la propriété usuelle de compatibilité avec les espérances conditionnelles $E_1[.|\mathfrak{F}_t]$, soit

$$(5) \qquad \text{pour } s<t, \quad E_s M_t \mathcal{E}(f_{s]}) = M_s \mathcal{E}(f_{s]})$$

En pratique, on vérifiera cela sous la forme

$$(6) \qquad < \mathcal{E}(g_{s]}), M_t \mathcal{E}(f_{s]}) > = < \mathcal{E}(g_{s]}), M_s \mathcal{E}(f_{s]}) > .$$

Exemple. Les opérateurs de multiplication par X_t, dans une interprétation probabiliste quelconque, forment évidemment une martingale d'opérateurs.

Cela s'applique aux opérateurs Q_t (multiplication dans une interprétation brownienne), P_t (multiplication dans une autre interprétation brownienne), Q_t+cN_t (multiplication dans une interprétation poissonnienne). Par combinaison linéaire, nous en déduisons trois martingales d'opérateurs

$$(7) \qquad a_t^+ \ , \ a_t^- \ , \ a_t^\circ = N_t$$

que nous appellerons les <u>trois martingales fondamentales</u> sur l'espace de Fock. A partir de ces trois processus d'opérateurs , par rapport auxquels on définira plus loin des intégrales stochastiques, se développera le <u>calcul stochastique non commutatif</u>.

<u>Commentaire</u>. Pourquoi <u>trois</u> martingales fondamentales ? La première conjecture de Hudson-Parthasarathy était que deux ($a^\pm$) suffiraient à construire une bonne théorie. Puis la martingale a° a été découverte : y a t'il des raisons sérieuses de penser que la liste est maintenant complète ?

De telles raisons sérieuses ont été fournies par J.L. Journé, qui a examiné le modèle discret (le " bébé Fock " de la fin de l'exposé IV, §I) - c'est la réussite de cette démarche qui nous a amenés à utiliser systématiquement le modèle discret dans notre présentation de l'espace de Fock . Voici le raisonnement dans le cas discret.

Le << bébé Fock >> Ψ est l'espace $L^2(\Omega)$, où $\Omega=\{-1,1\}^M$ avec la loi P sous laquelle les coordonnées X_i sont des v.a. de Bernoulli symétriques indépendantes. Nous représentons la base (X_A) des polynômes de Walsh dans l'ordre

$$\underline{\underline{X_\emptyset \mid X_1 \mid X_2, X_{12} \mid X_3, X_{13}, X_{23}, X_{123} \mid X_4, X_{14} \cdots}}$$

Ecrivons $\Psi=\Psi_M$ pour être bien explicites. Les traits verticaux successifs délimitent des segments de la base, de tailles $1,2^1,2^2,2^3$...qui peuvent être identifiés aux bases naturelles de Ψ_0, Ψ_1,... Tout élément X_A de la base peut s'écrire X_AX_B , où A est une partie de $\{1,\dots,n\}$ et B une partie de $\{n+1,\dots,M\}$: ceci est l'analogue discret de la décomposition de l'espace de Fock en produit tensoriel continu, et nous pouvons dire qu'un opérateur U sur Ψ est Ψ_n-adapté s'il est de la forme $V\otimes I$, où V agit sur Ψ_n , autrement dit si dans la décomposition précédente de la base on a $U(X_AX_B)=V(X_A)X_B$. Pour fixer les idées prenons M=2. Une matrice 4×4 est Ψ_0-adaptée si elle s'écrit $\begin{vmatrix} a & & 0 & \\ & a & & a \\ 0 & & a & \end{vmatrix}$ et Ψ_1-adaptée si elle s'écrit $\begin{vmatrix} a & b & 0 & \\ c & d & & \\ & & a & b \\ 0 & & c & d \end{vmatrix}$.

Dans ces conditions, l'espérance conditionnelle d'une matrice U par rapport à Ψ_n s'obtient en isolant le bloc $2^n\times 2^n$ en haut à gauche, en le répétant le long de la diagonale, et en remplissant le reste par des zéros. Ainsi, si U est une matrice 4×4 $U=\begin{vmatrix} A & B \\ C & D \end{vmatrix}$ où $A=\begin{vmatrix} a & b \\ c & d \end{vmatrix}$, les deux matrices écrites ci-dessus sont les espérances conditionnelles $E[U|\Psi_0]$ et $E[U|\Psi_1]$.

Le problème de représentation des martingales consiste à représenter une différence $E[U|\mathcal{Y}_n]-E[U|\mathcal{Y}_{n-1}]$ comme une combinaison linéaire, à coefficients matriciels $\mathcal{Y}_{n-1}$-adaptés, d'un nombre minimum de matrices $\mathcal{Y}_n$-adaptées d'espérance conditionnelle $E[.|\mathcal{Y}_{n-1}]$ nulle. Le cas n=2 va rendre ici la situation claire .

Reprenant $U=\begin{vmatrix} A & B \\ C & D \end{vmatrix}$, nous avons $U-E[U|\mathcal{Y}_2] = \begin{vmatrix} 0 & B \\ C & D-A \end{vmatrix}$, et cette différence s'écrit sous la forme

$$\begin{vmatrix} 0 & B \\ C & D-A \end{vmatrix} = \begin{vmatrix} B & 0 \\ 0 & B \end{vmatrix}\begin{vmatrix} 0 & I \\ 0 & 0 \end{vmatrix} + \begin{vmatrix} C & 0 \\ 0 & C \end{vmatrix}\begin{vmatrix} 0 & 0 \\ I & 0 \end{vmatrix} + \begin{vmatrix} D-A & 0 \\ 0 & D-A \end{vmatrix}\begin{vmatrix} 0 & 0 \\ 0 & I \end{vmatrix}$$

Or les trois matrices d'espérance conditionnelle nulle représentent a_2^- , a_2^+ et N_2 . Le cas général est exactement semblable, A,B,C,D,I désignant des matrices $2^n\times 2^n$ au lieu de 2×2 .

3. Espérances conditionnelles (digression)

Il n'est pas usuel de savoir calculer des espérances conditionnelles d'opérateurs en probabilités quantiques. Nous allons montrer que les calculs faits plus haut s'étendent à des lois pures ε_ϕ , à condition d'imposer à ϕ une condition assez naturelle.

Soit ϕ un vecteur normalisé de Φ, Nous supposons que ϕ admet une factorisation (où b peut être supposé normalisé)

(8) $$\phi = ab \ , \ a\in\Phi_{t]} \ , \ b\in\Phi_{[t}$$

et nous définissons, pour tout opérateur borné M, un opérateur borné N_o sur $\Phi_{t]}$ par la formule suivante (suggérée par R.L. Hudson)

(9) $$< u,N_o v > \ = < ub \ , \ M(vb) > \ (ue^{\Phi_{t]}} \ , \ ve^{\Phi_{t]}}) \ .$$

Soit N l'extension $N_o\otimes I_{[t}$ de N_o à Φ . Montrons que N joue le rôle d'espérance conditionnelle $E_\phi[M|\Phi_t]$. Avec les notations de (4)

$$< \phi,K^*MH\phi > = < ab, K^*MH(ab) > = < K(ab), MH(ab) > = < ub, M(vb) >$$
où u=Ka, v=Hb

$$= < u,N_o v > = < ub, N(vb) > \ (b \text{ normalisé })$$
$$= < K(ab), NH(ab) > = < \phi, K^*NH\phi > \ . \qquad (9)$$

On remarquera que, si l'on doit avoir ainsi des espérances conditionnelles à tout instant t, la condition de factorisation (8) doit avoir lieu à tout instant, ce qui impose à ϕ d'être un << état multiplicatif >> dans l'espace de Fock. Il n'est pas très naturel dans ce cas d'imposer la normalisation de b , et cela fait apparaître un facteur $1/\|b\|^2$ à droite de (9).

Nous allons écrire explicitement N dans une interprétation probabiliste, ce qui mettra en évidence les analogies avec les formules de changement de loi de type classique.

Désignons par β la v.a. $E_t[|b|^2]$ ($|b|^2$ dépend de la multiplication, donc de l'interprétation probabiliste) et supposons que β soit >0 p.p.. Alors

(10) pour $v\in\Phi_{t]}$, $w\in\Phi_{[t}$, on a $N(vw) = \frac{1}{\beta}E_t[\bar{b}M(vb)]w$

Avant de démontrer cela, supposons que M soit l'opérateur de multiplication par une v.a. (encore notée M !). Alors la formule (10) exprime simplement que N est l'opérateur de multiplication par $\frac{1}{\beta}E_t[M\beta]$, l'espérance conditionnelle classique de M sous la loi βP.

Pour vérifier (10), on écrit

$$\begin{aligned} <\phi, K^*MH\phi> &= <K(a)b, M(H(a)b)> = <K(a), \bar{b}M(H(a)b)> \\ &= <u, E_t(\bar{b}M(vb))> = <ub, \tfrac{1}{\beta}E_t(\bar{b}M(vb))b> \\ &= <\phi, K^*NH\phi> \end{aligned}$$

où N est l'opérateur défini par (10).

4. Définition d'intégrales stochastiques d'opérateurs.

Soit (H_t) un processus adapté d'opérateurs . Sous des hypothèses raisonnables, nous allons définir des intégrales stochastiques des trois types

(11) $I_t^+(H) = \int_0^t H_s da_s^+$, $I_t^-(H) = \int_0^t H_s da_s^-$, $I_t^\circ(H) = \int_0^t H_s dN_s$.

Les objets ainsi désignés seront des opérateurs , dont le domaine contient les vecteurs $\mathcal{E}(f)$ ($f\in L^2_{loc}$ pour $I^\pm$, $f\in L^\infty_{loc}$ pour I°).

Nous ferons les hypothèses minimales suivantes (outre l'adaptation, et l'inclusion $\mathcal{E}\subset\mathcal{D}(H_t)$ pour tout t)

(12) Soit $u\in L^2(\mathbb{R}_+)$; posons $U_t=\mathcal{E}(u_{t]})$. Alors l'application $t\longmapsto K_t=H_tU_t$ est mesurable, et l'on a
$$\int_0^t \|K_s\|^2\, ds < \infty \text{ pour tout t fini .}$$

Si les opérateurs H_t admettent des adjoints H_t^*, satisfaisant aux mêmes hypothèses (12), nous dirons que (H_t) satisfait à l'hypothèse (12^*).

Comme dans toute théorie de l'intégrale stochastique, on est amené à faire des vérifications par passage à la limite à partir de processus simples, ou étagés. Il s'agira ici de processus de la forme

(13) $H_0=0$, $H_t=H_{t_i}$ pour $t_i<t\leq t_{i+1}$, $H_t=0$ pour $t>t_n$

où $0<t_1\ldots<t_n$ est une subdivision finie, et H_{t_i} est Φ_{t_i}-adapté pour $i=0,..,n-1$. L'i.s. se réduit dans ce cas à une somme finie, comme toujours.

Le cas de a^-. Nous commençons par le cas de $I_t^-(H)$, qui est presque trivial. Lorsque le processus H est supposé simple, l'intégrale stochastique peut s'écrire

$$I_\infty^-(H)\mathcal{E}(u) = \Sigma_i\, H_{t_i}(a^-_{t_{i+1}} - a^-_{t_i})\mathcal{E}(u)$$

Compte tenu de la propriété $a_h^-\mathcal{E}(u) = <h,u>\mathcal{E}(u)$, cela nous laisse

(14) $$I_\infty^-(H)\mathcal{E}(u) = \int_0^\infty H_s\mathcal{E}(u)u_s\, ds$$

formule qui peut s'étendre directement à tout processus d'opérateurs (même non adapté !) possédant la propriété :

$$\int_0^\infty \|H_s \varepsilon(u)\|^2 \, ds < \infty$$

Dans le cas adapté, on a $\|H_s \varepsilon(u)\|^2 = \|H_s \varepsilon(u_s)\|^2 \|\varepsilon(u_{[s}\|^2$, et ce dernier facteur est borné supérieurement et inférieurement : la propriété équivaut donc à (12) écrite pour $t=+\infty$ (la localisation pour t fini est immédiate).

Remarquons aussi une fois pour toutes que si l'on a défini un opérateur sur les vecteurs cohérents, le prolongement aux combinaisons linéaires de ceux-ci ne pose aucun problème, grâce au lemme d'indépendance (IV, §I, 2 ; p. IV.7).

La formule suivante est la première d'une longue liste ; elle ne fait que traduire (14)

$$(15) \qquad \boxed{< \varepsilon(u), I_t^-(H)\varepsilon(v) > = \int_0^t < \varepsilon(u), v_s H_s \varepsilon(v) > ds}$$

REMARQUE. La formule (14) sous forme différentielle s'écrit simplement

$$(H_s da_s^-)\varepsilon(f) = H_s(da_s^- \varepsilon(f)) = H_s(f_s ds \varepsilon(f)) = H_s \varepsilon(f) f_s ds$$

Dans une expression comme $H_s da_s^-$, l'adaptation entraîne que les deux symboles H_s et da_s^- peuvent être traités comme des opérateurs qui commutent (ils agissent sur deux parties différentes du produit tensoriel continu).

Considérons maintenant deux processus adaptés (H_t) et (K_t) satisfaisant aux hypothèses ci-dessus. Nous avons d'après (14)

$$< I_t^-(H)\varepsilon(u), I_t^-(K)\varepsilon(v) > = < \int_0^t H_r \varepsilon(u) u_r dr, \int_0^t K_s \varepsilon(v) v_s ds >$$

d'où par dérivation

$$(16) \qquad \boxed{\frac{d}{dt} < I_t^-(H)\varepsilon(u), I_t^-(K)\varepsilon(v) > = < I_t^-(H)\varepsilon(u), v_t K_t \varepsilon(v) > + < u_t H_t \varepsilon(u), I_t^-(K)\varepsilon(v) >}$$

Formellement, ceci est une formule d'Ito (nous le verrons plus tard) : en effet, le côté gauche nous permet d'évaluer $< \varepsilon(u), d[I_t^+(H^*) I_t^-(K)]\varepsilon(v) >$. C'est pourquoi la formule est importante.

L'inégalité de Schwarz appliquée à (14) nous donne immédiatement

$$(17) \qquad \sup_t \|I_t^- \varepsilon(u)\| \leqq \|u\| \left(\int_0^\infty \|H_s \varepsilon(u)\|^2 ds \right)^{1/2} .$$

Le cas de a^+. Dans cette section, nous allons supposer que le processus adapté H est simple : toutes les vérifications sont alors immédiates. Il restera un problème d'approximation d'un processus adapté général par des processus simples, que nous discuterons plus loin (remarque au n°5).

Supposons H non seulement simple, mais élémentaire, i.e. tel que seul un H_{t_i} soit non nul. On a alors

$$< \varepsilon(u), I_\infty^+(H)\varepsilon(v) > = < \varepsilon(u_{t_i]})\varepsilon(u_{[t_i}), H_{t_i} \varepsilon(v_{t_i})(a^+_{t_{i+1}} - a^+_{t_i})\varepsilon(v_{[t_i}) >$$

$$= < \varepsilon(u_{t_i]}), H_{t_i}(v_{t_i]}) > < \varepsilon(u_{[t_i}), (a^+_{t_{i+1}} - a^+_{t_i})\varepsilon(v_{[t_i}) >$$

Dans le second facteur, on fait passer l'opérateur a^+ de l'autre côté, ce qui fait apparaître

$$(\int_{t_i}^{t_{i+1}} \bar{u}_s ds) < \varepsilon(u_{[t_i}), \varepsilon(v_{[t_i}) >$$

Recomposant alors les deux facteurs, on obtient - d'abord pour H élémentaire, puis pour H simple (et plus tard sous l'hypothèse (12))

$$(18) \qquad \boxed{< \varepsilon(u), I_t^+(H)\varepsilon(v) > = \int_0^t < u_s \varepsilon(u), H_s \varepsilon(v) > ds}$$

Cette formule se déduit formellement de (15) en écrivant que $I_t^+(H)$ et $I_t^-(H^*)$ sont adjoints l'un de l'autre.

Il nous faut ensuite la formule analogue à (16) : celle-ci comporte un terme supplémentaire, qui correspond à un crochet non trivial dans une formule d'Ito. Ici encore, on traite le cas élémentaire, puis le cas simple, et l'extension est remise à plus tard.

$$(19) \qquad \boxed{\begin{aligned} \frac{d}{dt} < I_t^+(H)\varepsilon(u), I_t^+(K)\varepsilon(v) > &= < u_t I_t^+(H)\varepsilon(u), K_t \varepsilon(v) > + \\ &+ < H_t \varepsilon(u), v_t I_t^+(K)\varepsilon(v) > + < H_t \varepsilon(u), K_t \varepsilon(v) > \end{aligned}}$$

Indiquons aussi la formule analogue pour les intégrales des deux types.

$$(20) \qquad \boxed{\frac{d}{dt} < I_t^-(H)\varepsilon(u), I_t^+(K)\varepsilon(v) > = <u_t I_t^-(H)\varepsilon(u), K_t \varepsilon(v)> + <u_t H_t \varepsilon(u), I_t^+(K)\varepsilon(v)>}$$

Nous déduisons de (19) une inégalité analogue à (17), mais un peu plus délicate[1] ; nous posons

$$M_t = I_t^+(H) \ , \quad A_t = \|M_t \varepsilon(u)\|^2 \ , \quad A_t^* = \sup_{s \leq t} A_s \ , \quad B_t = \int_0^t \|H_s \varepsilon(u)\|^2 ds \ .$$

Nous avons alors d'après (19)

$$\|M_t \varepsilon(u)\|^2 = 2 \int_0^t \Re < u_s M_s \varepsilon(u), H_s \varepsilon(u) > ds + \int_0^t \|H_s \varepsilon(u)\|^2 ds$$

d'où en appliquant l'inégalité de Schwarz et prenant un sup

$$A_t^* \leq 2\|u\| A_t^{*1/2} B_t^{*1/2} + B_t$$

et enfin

$$(21) \qquad \sup_t \|I_t^+(H)\varepsilon(u)\| \leq (\|u\| + \sqrt{\|u\|^2 + 1}) (\int_0^\infty \|H_s \varepsilon(u)\|^2 ds)^{1/2} \ .$$

Le cas de N. On utilise la formule

$$< \varepsilon(u), (N_{t_{i+1}} - N_{t_i})\varepsilon(v) > = (\int_{t_i}^{t_{i+1}} \bar{u}_s v_s ds) < \varepsilon(u), \varepsilon(v) >$$

pour établir, pour H élémentaire puis simple, les formules suivantes

$$(22) \qquad \boxed{< \varepsilon(u), I_t^\circ \varepsilon(v) > = \int_0^t < u_s \varepsilon(u), v_s H_s \varepsilon(v) > ds}$$

1. Les trois formules analogues viennent d'un exposé de J.L. Journé.

$$(23)\quad \boxed{\begin{aligned}\frac{d}{dt}< I_t^o(H)\mathcal{E}(u),\ I_t^o(K)\mathcal{E}(v) > &= < u_t I_t^o(H)\mathcal{E}(u),\ v_t K_t \mathcal{E}(v) > + \\ &+ < u_t H_t \mathcal{E}(u),\ v_t I_t^o(K)\mathcal{E}(v) > + < u_t H_t \mathcal{E}(u),\ v_t K_t \mathcal{E}(v) >\end{aligned}}$$

Etablissons la formule analogue à (21), en prenant cette fois $M_t = I_t^o(H)$. Nous avons d'après (23)

$$A_r = \|M_r \mathcal{E}(u)\|^2 = 2\int_o^r |u_s|^2 \Re < M_s \mathcal{E}(u), H_s \mathcal{E}(u) > ds + \int_o^r |u_s|^2 \|H_s \mathcal{E}(u)\|^2 ds$$

$$|A_r - C_r| \leq 2\|u\| \left(\int_o^r \|M_s \mathcal{E}(u)\|^2 |u_s|^2 \|H_s \mathcal{E}(u)\|^2 ds\right)^{1/2} \ ;\ C_r = \int_o^r |u_s|^2 \|H_s \mathcal{E}(u)\|^2 ds$$

$$\leq 2\|u\| A_r^{*1/2} C_r^{1/2}$$

Passant au sup sur $r \in [0,t]$, nous avons $A_t^* \leq 2\|u\| A_t^{*1/2} C_t^{1/2} + C_t$, et comme pour (21), l'inégalité

$$(24)\quad \sup_t \|I_t^o(H)\mathcal{E}(u)\| \leq (\|u\| + \sqrt{\|u\|^2+1}) \left(\int_o^\infty |u_s|^2 \|H_s \mathcal{E}(u)\|^2 ds\right)^{1/2}$$

avec un second membre un peu moins sympathique que (21). Toutefois, on ne peut espérer mieux : on a en effet $C_t \leq 2\|u\| A_t^{*1/2} C_t^{1/2} + A_t^*$, de sorte que les deux quantités sont équivalentes. C'est pourquoi, dans les raisonnements concernant les i.s. relatives à (N_t), il est préférable d'utiliser des vecteurs cohérents $\mathcal{E}(u)$, où u est (localement) borné.

Pour conclure le formulaire, il reste les formules mixtes, analogues à (20). Seule la seconde est non triviale.

$$(25)\quad \boxed{\begin{aligned}\frac{d}{dt}< I_t^o(H)\mathcal{E}(u),\ I_t^{\pm}(K)\mathcal{E}(v) > &= < u_t H_t \mathcal{E}(u),\ v_t I_t^{\pm}(K)\mathcal{E}(v) > + \\ (\text{cas } -) \qquad &+ \begin{cases} < I_t^o(H)\mathcal{E}(u),\ v_t K_t \mathcal{E}(v) > \\ < u_t I_t^o(H)\mathcal{E}(u),\ K_t \mathcal{E}(v) > + < H_t \mathcal{E}(u), v_t K_t \mathcal{E}(v) > \end{cases} \\ (\text{cas } +) \qquad &\end{aligned}}$$

5. Interprétation au moyen d'i.s. classiques.

Savoir intégrer relativement à a^+, a^- revient à savoir intégrer un processus d'opérateurs (H_t) relativement aux mouvements browniens conjugués Q et P. Or certains probabilistes se sont intéressés à la définition d'intégrales stochastiques du type $\int_o^t H_s dB_s$ relativement au mouvement brownien, leur idée étant d'ailleurs de considérer l'élément différentiel $H_s dB_s$ comme un vecteur (H_s appliqué au vecteur dB_s) plutôt qu'un opérateur (H_s composé avec l'opérateur de multiplication par dB_s). Nous allons reprendre la discussion précédente sous cet angle.

Soit $\mathcal{E}(u) = U_\infty$ un vecteur cohérent (une exponentielle stochastique) ; on pose $\mathcal{E}(u_{t]}) = U_t$, et $K_t = H_t U_t$ - ceci est un processus adapté, de carré intégrable sur tout intervalle fini ; nous ne perdrons aucune vraie généralité en le supposant de carré intégrable sur $\mathbb{R}_+$ et (conformément aux lois du calcul stochastique) en le remplaçant par une projection prévisible. Lorsque H est un processus simple, nous pouvons définir de manière évidente

l'intégrale stochastique d'opérateurs $J_t=\int_0^t H_s dB_s$ pour $t\leq+\infty$, et l'on a d'après la propriété d'adaptation de (H_t)

$$J_\infty U_\infty = \Sigma_i \ (H_{t_i} U_{t_i})(B_{t_{i+1}} - B_{t_i})\varepsilon(u_{[t_i})$$

Or nous avons $\varepsilon(u_{[t_i})=U_\infty/U_t$, et nous obtenons l'expression suivante, qui acquiert un sens pour des processus adaptés non nécessairement simples

$$(26) \qquad J_\infty U_\infty = (\int_0^\infty H_s dB_s)\varepsilon(u) = U_\infty(\int_0^\infty K_s/U_s \ dB_s \ ) \quad \text{où} \quad K_s=H_s U_s$$

Posons $V_t=\int_0^t K_s/U_s \ dB_s$, $\mathcal{J}_t=U_t V_t = J_t\varepsilon(u_{t]})$. Nous avons $J_\infty U_\infty=\mathcal{J}_\infty$. D'autre part, on a

$$\mathcal{J}_t = \int_0^t U_s dV_s + V_s dU_s + d[U,V]_s \quad , \quad dU_t = u_t U_t dB_t$$

d'où pour $\mathcal{J}_t$ une équation différentielle stochastique linéaire non homogène

$$(27) \qquad \mathcal{J}_t = \int_0^t K_s(dB_s+u_s ds) + \int_0^t u_s \mathcal{J}_s dB_s \ .$$

On peut refaire ce raisonnement en remplaçant (B_t) par (X_t), une autre martingale fondamentale d'une interprétation probabiliste, mais il faut prendre garde à deux détails : d'abord, à bien choisir pour (K_t) une version càdlàg.

$$K_t = (H_{t_i} U_{t_i})\varepsilon(u_{]t_i,t]}) \text{ pour } t_i\leq t<t_{i+1}$$

et ensuite, à la possibilité d'annulation de (U_t), ce qui fait qu'une formule du genre de (26) cesse d'être correcte. En revanche, on a une formule analogue à (27)

$$(28) \qquad \mathcal{J}_t = \int_0^t K_{s-}(dX_s+u_s d[X,X]_s) + \int_0^t u_s \mathcal{J}_{s-} dX_s$$

Peut être est il bon de rappeler que la notion de produit de v.a. change avec l'interprétation probabiliste utilisée.

On peut utiliser l'équation (27) ou (28) pour établir des inégalités de normes, mais celles-ci sont moins bonnes que (17), (21) ou (24). Signalons que les versions antérieures indiquaient une inégalité fausse concernant le cas discontinu.

REMARQUE. L'équation (28) garde un sens pour un processus d'opérateurs non simple, si l'on note que les intégrales stochastiques à gauche ne dépendent (pour un processus de Poisson compensé) que de la classe du processus (K_t) pour la mesure $dP\times dt$.

On peut aussi, pour définir ces intégrales, utiliser un procédé de régularisation. Posons par convention $H_t=0$ pour $t<0$, puis définissons des régularisés continus, pour $\varepsilon>0$

$$H_t^\varepsilon = \frac{1}{\varepsilon}\int_0^\varepsilon H_{t-s}\,ds \quad \text{(a un sens sur le domaine commun } \mathcal{E}\text{)}$$

Pour ce processus adapté, il n'y a aucun problème de définition en (28) : le processus (K_t^ε) admet une unique version continue. Lorsque $\varepsilon \to 0$, on a ensuite un passage à la limite en norme L^2.

Cela résout aussi le problème d'approximation par des processus simples : pour les processus continus (H_t^ε), on utilise simplement des sommes de Riemann, permettant d'étendre toutes les formules fondamentales (15) à (24). Puis l'on passe encore une fois à la limite lorsque $\varepsilon \to 0$.

Cette démonstration n'est pas tout à fait satisfaisante, puisqu'elle utilise à la fois deux procédés de définition différents de l'i.s. d'opérateurs. La démonstration de Hudson-Parthasarathy (Comm. Math. Phys 93, 1984, Prop. 3.2 p. 305) est plus pure.

II. Développement du calcul stochastique

Le calcul stochastique non commutatif sur l'espace de Fock est le calcul des intégrales stochastiques par rapport aux trois processus a_t^ε ($\varepsilon = +,-,o$), auquel on ajoute le temps t . C'est un outil bien moins développé[1] que le calcul stochastique classique, car

- On n'a pas défini d'i.s. par rapport à des martingales plus générales que les processus ci-dessus ; on n'a pas de formules d'isométrie, mais seulement des inégalités ; on ne sait pas bien changer de loi.
- On ne dispose pas d'une variété suffisante de processus << à variation finie >> (on n'intègre que par rapport à dt).

On voit donc que le calcul stochastique quantique correspond à l'ancien calcul d'Ito - et encore : on ne sait bien ni ajouter et multiplier les opérateurs comme on fait pour les v.a., ni écrire une fonction C^2 d'opérateurs à la manière de la formule d'Ito. La principale application du calcul stochastique quantique, pour l'instant, consiste à résoudre des équations différentielles stochastiques linéaires (du type exponentiel) décrivant des processus d'opérateurs unitaires sur l'espace de Fock.

1. Définition des crochets

La formule la plus fondamentale du calcul stochastique classique est sans aucun doute la formule d'intégration par parties

$$d(XY) = X\,dY + Y\,dX + [dX,dY] \quad ,$$

de laquelle la formule d'Ito générale se déduit sans grande difficulté. Ici, en raison du danger de confusion avec les commutateurs, nous noterons simplement $dX_t dY_t$ le crochet $[dX_t, dY_t]$. Nous nous proposons de calculer les "crochets" $da_t^\varepsilon da_t^\eta$, où $\varepsilon,\eta = +,-,o$. Cela suppose d'abord que l'on sache définir les produits $a_t^\varepsilon a_t^\eta$. Plus précisément, nous voulons montrer

1. Barnett-Streater-Wilde ont une théorie plus complète pour les fermions.

$$< \mathcal{E}(u),\ a_t^{\varepsilon} a_t^{\eta} \mathcal{E}(v) > = < a_t^{-\varepsilon} \mathcal{E}(u),\ a_t^{\eta} \mathcal{E}(v) >$$

Ce point résulte de ce qui a été dit dans l'exposé IV, §III, 2 sur les opérateurs définis par de bonnes matrices (en particulier les a_t^{ε}) : ces opérateurs agissent non seulement sur les vecteurs cohérents, mais sur les vecteurs-test, et on a une bonne stabilité par composition, etc. Nous ne donnerons pas de détails.

Nous avons alors, d'après l'exposé IV, §I, (12)

$$< \mathcal{E}(u),\ a_t^+ a_t^- \mathcal{E}(v) > = < a_t^- \mathcal{E}(u), a_t^- \mathcal{E}(v) > = (\int_0^t \bar{u}_s ds)(\int_0^t v_s ds) < \mathcal{E}(u), \mathcal{E}(v) >$$

de même $< \mathcal{E}(u),\ a_t^- a_t^+ \mathcal{E}(v) > = (\int_0^t \bar{u}_s ds)(\int_0^t v_s ds + t\ ) < \mathcal{E}(u), \mathcal{E}(v) >$

et de même avec a° (formules (21)-(23) de l'exposé IV, en se rappelant que $N_t = \lambda(m_t)$, la multiplication par $I_{[0,t]}$). Nous comparons cela avec les formules tirées de cet exposé-ci, (15) et (18) par exemple

$$<\mathcal{E}(u), (\int_0^t a_s^+ da_s^-)\mathcal{E}(v)> = \int_0^t <\mathcal{E}(u), v_s a_s^+ \mathcal{E}(v)> ds = (\int_0^t v_s ds \int_0^r \bar{u}_r dr) <\mathcal{E}(u),\ \mathcal{E}(v) >$$

$$<\mathcal{E}(u), (\int_0^t a_s^- da_s^+)\mathcal{E}(v)> = (\int_0^t \bar{u}_s ds \int_0^s v_r dr) < \mathcal{E}(u), \mathcal{E}(v) > .$$

Nous obtenons deux formules d'intégration par parties (sur les vecteurs cohérents)

$$(1_a) \qquad a_t^+ a_t^- = \int_0^t a_s^+ da_s^- + a_s^- da_s^+ , \quad a_t^- a_t^+ = \int_0^t a_s^+ da_s^- + a_s^- da_s^+ + Ids$$

qui s'expriment comme des calculs de "crochets"

$$(2_a) \qquad da_t^+ da_t^- = 0, \quad da_t^- da_t^+ = dt$$

On pourra vérifier de même que

$$(1_b) \qquad (a_t^+)^2 = 2\int_0^t a_s^+ da_s^+ \quad , \quad (a_t^-)^2 = 2\int_0^t a_s^- da_s^- \quad ,$$

c'est à dire, pour les "crochets"

$$(2_b) \qquad da_t^+ da_t^+ = 0 \ , \ da_t^- da_t^- = 0 \quad .$$

Mais en réalité, ce n'est pas nécessaire : il suffit d'utiliser (2_a), et les crochets d'Ito $dQ_t^2 = dt$, $dP_t^2 = dt$, bien connus de tous. De même les relations $dP_t dt = dQ_t dt = 0$ nous donneront $da^{\pm} dt = 0$ sans autre travail. Quant à (32), elle s'exprime sur les mouvements browniens conjugués par

$$dQ_t dP_t = -dP_t dQ_t = i dt$$

qui nous redonne le commutateur $[\, dP_t, dQ_t \,] = \frac{2}{i} dt$. La table de multiplication doit être complétée par l'adjonction de $a^\circ = N$. Une vérification directe donne

$$(1_c) \qquad \begin{aligned} a_t^+ N_t &= \int_0^t a_s^+ dN_s + N_s da_s^+ \quad , & N_t a_t^+ &= \int_0^t N_s da_s^+ + a_s^+ dN_s + da_s^+ \\ a_t^- N_t &= \int_0^t a_s^- dN_s + N_s da_s^- & N_t a_t^- &= \int_0^t N_s da_s^- + a_s^- dN_s \end{aligned}$$

et $$(N_t)^2 = 2\int_0^t N_s dN_s + N_s$$

c'est à dire pour les " crochets"

(2c) $$da_t^+ dN_t = 0 \ , \ dN_t da_t^+ = da_t^+ \ ; \ da_t^- dN_t = da_t^- \ , \ dN_t da_t^- = 0 \ ; \ dN_t dN_t = dN_t$$

Comme d'habitude, il n'est pas nécessaire d'établir toutes ces formules : on peut utiliser les résultats probabilistes concernant les processus de Poisson compensés $X_t = Q_t + cN_t$, $Y_t = P_t + cN_t$, qui satisfont à $dX_t^2 = dt + cdX_t$, $dY_t^2 = dt + cdY_t$. De même, les probabilités nous donneront directement $dN_t dt = 0$.

Les quatre " crochets " non nuls sont évidemment les seuls à retenir. Si l'on ordonne naturellement $(-,o,+)$, ils correspondent à des couples croissants

(3) $$\boxed{da_t^- da_t^+ = dt \ , \ da_t^- da_t^o = da_t^- \ , \ da_t^o da_t^+ = da_t^+ \ , \ da_t^o da_t^o = da_t^o}$$.

2. Variations quadratiques.

La lecture de cette section n'est pas indispensable.

Nous désignons par $[0,T]$ un intervalle borné fixe, par t_i^n le point $i2^{-n}T$, souvent abrégé en t_i s'il n'y a pas d'ambigüité ; nous posons

$$\Delta_i^n t = 2^{-n}T = t_{i+1}^n - t_i^n \ , \text{ abrégé en } \Delta t \ ;$$

$$\Delta_i^n a^\varepsilon = a_{t_{i+1}^n}^\varepsilon - a_{t_i^n}^\varepsilon \quad (\varepsilon = -,o,+), \text{ abrégé en } \Delta_i a^\varepsilon \text{ ou } \Delta_i^\varepsilon \ .$$

L'étude de la variation quadratique consiste à déterminer la limite (sur le domaine $\mathcal{E}$) des sommes

(4) $$Q_n^{\varepsilon\eta} = \Sigma_i \ \Delta_i^n a^\varepsilon \ \Delta_i^n a^\eta \qquad (\varepsilon,\eta = -,o,+)$$

lorsque $n \to \infty$. Comme en probabilités classiques, cette limite est un " crochet" au sens du n° précédent. Désignant par $Q^{\varepsilon\eta}$ cette limite, nous nous occuperons aussi d'un problème (suggéré par Nelson en probabilités classiques, mais avec un conditionnement que nous ne ferons pas ici), qui consiste à étudier la limite de

$$(Q_n^{\varepsilon\eta} - Q^{\varepsilon\eta})/\Delta^n t$$

autrement dit, à pousser l'étude des variations quadratiques jusqu'au second ordre.

Nous commençons par remarquer que, pour tout vecteur cohérent $\mathcal{E}(u)$, les "sommes de Riemann"

$$\Sigma_i \ a_{t_i}^\varepsilon (a_{t_{i+1}}^\eta - a_{t_i}^\eta)\mathcal{E}(u) = \Sigma_i \ a_{t_i}^\varepsilon \Delta_i^\eta \mathcal{E}(u)$$

(n est sous-entendu) convergent en norme vers l'intégrale stochastique $(\int_0^T a_s^\varepsilon da_s^\eta)\mathcal{E}(u)$. En effet, la somme de Riemann elle même est l'intégrale stochastique d'un processus étagé convergeant vers (a_t^ε), et on applique alors à la différence l'inégalité (17) si $\varepsilon=-$, (21) si $\varepsilon=+$, et si $\varepsilon=o$ une inégalité

du type (28) . On a d'autre part

$$a_T^\varepsilon a_T^\eta = \Sigma_i\, a_{t_{i+1}}^\varepsilon a_{t_{i+1}}^\eta - a_{t_i}^\varepsilon a_{t_i}^\eta$$

$$Q_n^{\varepsilon\eta} = a_T^\varepsilon a_T^\eta - \Sigma_i a_{t_i}^\varepsilon (a_{t_{i+1}}^\eta - a_{t_i}^\eta) - \Sigma_i\, (a_{t_{i+1}}^\varepsilon - a_{t_i}^\varepsilon) a_{t_i}^\eta$$

(noter la commutation des derniers opérateurs). Cette expression tend en norme, sur les vecteurs cohérents, vers

$$a_T^\varepsilon a_T^\eta - \int_0^T (a_s^\varepsilon da_s^\eta + a_s^\eta da_s^\varepsilon) = \int_0^T da_s^\varepsilon da_s^\eta$$

exactement comme en calcul stochastique classique. Nous passons à l'étude au second ordre. Considérons d'abord la somme la plus intéressante

$$(5)\quad S_n^{+-}\mathcal{E}(u) = \frac{1}{\Delta t}\Sigma_i\, \Delta_i^+\Delta_i^-\mathcal{E}(u) = \Sigma_i (a_{t_{i+1}}^+ - a_{t_i}^+)(a_{t_{i+1}}^- - a_{t_i}^-)/(t_{i+1}-t_i)$$

Appliquons cet opérateur à $\mathcal{E}(u)$: nous obtenons

$$S_n^{+-}\mathcal{E}(u) = (\int_0^T u_s^n da_s^+)\mathcal{E}(u)$$

où u_s^n vaut $\frac{1}{t_{i+1}-t_i}\int_{t_i}^{t_{i+1}} u_r\,dr$ pour se $]t_i, t_{i+1}]$. Il est bien connu que u_s^n converge dans L^2 vers u_s sur $[0,T]$, et par conséquent l'expression précédente converge en norme vers $a_{uI_{[0,T]}}^+\mathcal{E}(u)$, c'est à dire, d'après la formule IV.I, (19), vers $N_T\mathcal{E}(u)$. <u>Ainsi,</u> S_n^{+-} <u>converge vers</u> N_T : c'est tout à fait satisfaisant, car nous avons vu en IV, §II, (24) que l'étude du bébé Fock suggère la formule $da_t^+ da_t^- = dN_t\,dt$ - exactement ce que nous venons de voir.

Notre satisfaction s'arrête là, car dans la théorie du bébé Fock, nous avons $a_k^- a_k^- = 0$. Ici, nous avons bien $da_s^- da_s^- = 0$, et le bébé Fock donne donc une prédiction correcte au premier ordre. Mais au second ordre, exactement la même démonstration que ci-dessus va nous donner

$$\lim_n S_n^{--}\mathcal{E}(u) = \lim_n \frac{1}{\Delta t}(a_{t_{i+1}}^- - a_{t_i}^-)^2\mathcal{E}(u) = \lim_n (\int_0^T u_s^n u_s\,ds)\mathcal{E}(u)$$

$$(6)\qquad = (\int_0^T u_s^2\,ds)\mathcal{E}(u)$$

La famille d'opérateurs M_t définis par $M_t\mathcal{E}(u) = (\int_0^t u_s^2\,ds)\mathcal{E}(u)$ nous fournit un nouvel exemple de martingale d'opérateurs.

Le cas de S_n^{-+} se ramène aisément à celui de S_n^{+-}, puisqu'on connaît le commutateur :

$$S_n^{+-} = \frac{1}{\Delta t}(\Sigma_i\, (a_{t_{i+1}}^+ - a_{t_i}^+)(a_{t_{i+1}}^- - a_{t_i}^-) - TI\,) = S_n^{-+} \longrightarrow N_T\ .$$

Reste l'étude de S_n^{++} : cette somme <u>ne peut pas avoir une limite</u>. En effet, si les quatre sommes avaient des limites, il en serait de même des sommes

$$\Sigma_i\, ((Q_{t_{i+1}} - Q_{t_i})^2 - (t_{i+1}-t_i))/(t_{i+1}-t_i) \qquad (\text{resp. } P_{t_{i+1}} - P_{t_i})$$

Or une telle somme comporte n variables de même loi que $\xi^2 - 1$, où ξ est

normale centrée réduite : sa norme dans L^2 tend donc vers l'infini. Cela indique que notre martingale (M_t) ci-dessus était pathologique : elle n'admet pas d'adjoint raisonnable.

Formellement, cette martingale est une intégrale stochastique $\int_0^t H_s da_s^-$, où $H_s\mathcal{E}(u)=u_s\mathcal{E}(u)$ est un opérateur défini sur beaucoup de vecteurs cohérents (les $\mathcal{E}(u)$ avec u continue), mais non fermable.

3. Equations différentielles stochastiques linéaires.

En probabilités classiques, on résout des équations différentielles stochastiques gouvernées par une ou plusieurs semimartingales directrices scalaires, et dont la solution est une semimartingale à valeurs dans une variété E (par exemple). Dans la théorie la plus ancienne, celle d'Ito, les semimartingales directrices sont le temps t , et un ou plusieurs mouvements browniens scalaires B_t^i jouant le rôle de << source de bruit >> . Fréquemment, la solution de l'e.d.s. fournit un processus de Markov sur la variété E : prenant une espérance (avec point initial variable), elle fournit un semi-groupe de Markov, de générateur donné.

En probabilités quantiques, c'est notre espace de Fock Φ , avec ses trois martingales fondamentales, qui joue le rôle de << source de bruit non commutatif >> ; plus précisément, Φ remplace un mouvement brownien scalaire Si l'on voulait l'analogue de n mouvements browniens, il faudrait utiliser $\Phi^{\otimes n}$ (qui est aussi l'espace de Fock sur $L^2(\mathbb{R}_+,\mathbb{R}^n)$). Nous nous bornerons à n=1, pour la simplicité.

Le rôle de la << variété >> E sera joué par un espace de Hilbert fixe $\mathcal{B}$ (n'abusons pas de la lettre H !) appelé espace de Hilbert initial. Le rôle du processus de diffusion solution de l'e.d.s. (autrement dit, une famille de v.a. définies sur l'espace Ω du brownien, à valeurs dans E) est joué par une famille d'opérateurs X_t sur l'espace de Hilbert $\mathcal{B}\otimes\Phi$, admettant pour domaine commun $\mathcal{B}\underline{\otimes}\mathcal{E}_b$, l'ensemble des combinaisons linéaires finies de vecteurs $b\otimes\mathcal{E}(u)$, où b parcourt $\mathcal{B}$, et u est un élément borné de $L^2(\mathbb{R}_+)$. Ici encore, nous évitons la généralité (et la complexité) maximale : au lieu de $\mathcal{B}$ entier, b pourrait être restreint à un sous-espace dense $\mathcal{B}_0$, comme font Hudson-Parthasarathy. Nous écrirons assez souvent bh ($b\in\mathcal{B}$, $h\in\Phi$) au lieu de $b\otimes h$, pour alléger les notations.

La notion de processus adapté d'opérateurs s'étend sans difficulté à la situation présente : c'est une famille d'opérateurs H_t , admettant comme domaine commun $\mathcal{B}\underline{\otimes}\mathcal{E}_b$, et possédant les propriétés

$H_t(b\mathcal{E}(u_{t]}))\ e\ \mathcal{B}\otimes\Phi_t$, $H_t(b\mathcal{E}(u))=(H_t(b\mathcal{E}(u_{t]}))\mathcal{E}(u_{[t})$ pour tout t . Pour un tel processus, tel en outre que $t\mapsto H_t(b\mathcal{E}(u))$ soit dans $L^2_{loc}(\mathbb{R}_+,\mathcal{B}\otimes\Phi)$,

on peut définir des intégrales stochastiques

$$(7) \qquad I_t^\varepsilon(H) = \int_0^t H_s d(I\otimes a_s^\varepsilon) \qquad (\varepsilon=-,o,+)$$

aussi notées, simplement, $\int_0^t H_s da_s^\varepsilon$. La théorie est parallèle à celle du paragraphe précédent, sans aucune difficulté nouvelle, et nous la laissons au lecteur. Nous inventerons un quatrième << indice ε >> (invisible à l'oeil nu) pour représenter l'intégration par rapport à ds . Avec cette convention, nous avons la formule suivante (qui généralise (15), (18) et (22))

$$(8) \quad < a\mathcal{E}(u), \Sigma_\varepsilon I_t^\varepsilon(H^\varepsilon) b\mathcal{E}(v) > = \int_0^t < a\mathcal{E}(u),\ v_s H_s^-(b\mathcal{E}(v)) > ds$$
$$+ \int_0^t < a\mathcal{E}(u)u_s, v_s H_s^o(b\mathcal{E}(v)) > ds + \int_0^t < a\mathcal{E}(u)u_s,\ H_s^+(b\mathcal{E}(v)) > ds$$
$$+ \int_0^t < a\mathcal{E}(u), H_s(b\mathcal{E}(v)) > ds \ .$$

REMARQUE. Il est intéressant de noter que, si l'on connaît le processus d'opérateurs $J_t = \sum_\varepsilon I_t^\varepsilon(H^\varepsilon)$, on connaît aussi les quatre processus adaptés H_t^ε . Supposons <u>continues à droite</u>, pour simplifier, toutes les fonctions $s \longmapsto < a\mathcal{E}(u), H_s^\varepsilon(b\mathcal{E}(v)) >$; fixons t, et posons

$$v'_s = v_s \text{ pour } s\leq t \ , \ v'_s = 0 \text{ pour } s>t$$

Alors on a

$$\frac{d}{ds} < a\mathcal{E}(u'),\ J_s(b\mathcal{E}(v'))>|_{s=t+} = < a\mathcal{E}(u_{t]}),\ H_t(b\mathcal{E}(v_{t]})) >$$

grâce à l'hypothèse de continuité à droite faite plus haut ; en utilisant l'adaptation, cela détermine $H_t(b\mathcal{E}(v_{t]}))$, puis $H_t(b\mathcal{E}(v))$. Cela permet de déterminer H , puis de se ramener au cas où $H=0$. Posons ensuite

$$v''_s = v_s \text{ pour } s\leq t \ , \ 1 \text{ pour } t<s\leq t+1, \ 0 \text{ pour } s>t+1$$

Nous avons (en supposant $H=0$)

$$\frac{d}{ds} < a\mathcal{E}(u'),\ J_s(b\mathcal{E}(v''))>|_{s=t+} = < a\mathcal{E}(u_{t]}), H_t^-(b\mathcal{E}(v_{t]})) >$$
$$\frac{d}{ds} < a\mathcal{E}(u''),\ J_s(b\mathcal{E}(v'))>|_{s=t+} = < a\mathcal{E}(u_{t]}), H_t^+(b\mathcal{E}(v_{t]})) >$$

qui permet d'extraire H^-, H^+, et de se ramener au cas où H, H^-, H^+ sont nuls. On a dans ce cas

$$\frac{d}{ds} < a\mathcal{E}(u''),\ J_s(b\mathcal{E}(v''))>|_{s=t+} = < a\mathcal{E}(u_{t]}),\ H_t^o(b\mathcal{E}(v_{t]})) >$$

toujours grâce à la continuité à droite.

Nous aurons besoin aussi d'intégrales stochastiques des types suivants

$$(9) \qquad \int_0^t H_s(L\otimes da_s^\varepsilon) \quad \text{et} \quad \int_0^t (L\otimes da_s^\varepsilon) H_s \ ,$$

où L est un opérateur borné (pour simplifier) sur $\mathcal{B}$: ce sont des intégrales du type (7), relatives aux processus adaptés $H_s(L\otimes I)$ et $(L\otimes I)H_s$ respectivement : alors que H_s commute avec l'élément différentiel $I\otimes da_s^\varepsilon$, il ne commute pas en général avec $L\otimes da_s^\varepsilon$.

Nous pouvons maintenant décrire le genre d'équations différentielles stochastiques linéaires (tout est linéaire en mécanique quantique) que nous considérerons :

$$(10) \qquad U_t = I + \int_0^t \Sigma_\varepsilon \; U_s(L_\varepsilon \otimes da_s^\varepsilon)$$

où les L_ε sont quatre opérateurs bornés sur $\mathfrak{B}$. L'inconnue (U_t) est un processus d'opérateurs adapté, possédant la propriété de continuité à droite faible utilisée à la page précédente. On s'intéressera tout spécialement au cas où la solution de (10) est un processus d'opérateurs unitaires (ou, plus précisément, prolongeables en opérateurs unitaires). En même temps que (10), on considérera (sans donner de détails) l'e.d.s. analogue où U_s est placé à droite.

Existence de la solution. Soit $\alpha=(\varepsilon_1,\dots,\varepsilon_n)$ un multiindice de longueur $|\alpha|=n$. Introduisons les opérateurs

$$(11) \qquad L_\alpha = L_{\varepsilon_1}\dots L_{\varepsilon_n} \quad , \quad I_t^\alpha = \int_{s_1<s_2\dots<s_n\leq t} da_{s_1}^{\varepsilon_1}\dots da_{s_n}^{\varepsilon_n}$$

Si M est une borne des $\|L_\varepsilon\|$, on a $\|L_\alpha\|\leq M^n$. D'autre part, on vérifie aisément (par récurrence) que (I_t^α) est un processus adapté, continu à droite, d'opérateurs sur Φ, satisfaisant à une inégalité de la forme

$$(12) \qquad \|I_t^\alpha \varepsilon(u)\|^2 \leq C^n t^n/n! \quad \text{pour} \quad t\in[0,T]$$

où C est une constante ≥ 1, dépendant des normes de u dans L^2 et L^∞, et de T (il faut appliquer une inégalité de Schwarz lors de l'intégration par rapport à dt) : cf (17), (21), (24). Il est alors facile de calculer les approximants de la solution de (10) par la méthode d'itération

$$(13) \qquad U_t^{(n)}(bh) = bh + \Sigma_{|\alpha|\leq n} L_\alpha(b) I_t^\alpha(h) \qquad (h=\varepsilon(u))$$

et la présence de n! au dénominateur de (12) montre que ces sommes convergent en norme - on vérifie sans peine que la limite est une solution de (10). Il est peut être intéressant de souligner que (13) reste dans le produit tensoriel algébrique $\mathfrak{B}\otimes\Phi$, pour n fini.

REMARQUE. On voit que les opérateurs L_ε n'ont pas vraiment besoin d'être bornés : il suffit qu'ils admettent un domaine commun stable $\mathfrak{B}_0$ (auquel appartiendra le vecteur b) et que $\|L_\alpha(b)\|$ soit à croissance suffisamment lente.

D'autre part, les approximants de l'équation analogue à (10), mais avec U_s à droite, diffèrent de (13) par le fait que les indices $\varepsilon_1\dots\varepsilon_n$ dans la définition de I_t^α sont remplacés par $\varepsilon_n\dots\varepsilon_1$, en ordre inverse.

Unicité de la solution. Il s'agit de vérifier qu'une famille (W_t) d'opérateurs sur $\mathfrak{B}\otimes\mathfrak{E}_b$, adaptée, continue à droite, et telle que

$$W_t = \int_0^t \Sigma_\varepsilon \; W_s(L_\varepsilon \otimes da_s^\varepsilon)$$

est nécessairement nulle. Pour cela, on raisonne comme plus haut, en vérifiant par récurrence une inégalité de la forme

$$\|W_t(b\mathcal{E}(u))\|^2 \leqq C^n t^n/n! \quad \text{pour } t\in[0,T]$$

qui, pour $n\to\infty$, entraîne $W_t=0$.

Quelques identités. L'existence de la solution ayant été établie, nous recopions quelques relations établies au paragraphe I. Les relations sont longues à écrire, mais immédiates quant à leur substance. D'abord :

$$(14)\quad \frac{d}{dt}\langle a\mathcal{E}(u),U_t(b\mathcal{E}(v))\rangle = v_t\langle a\mathcal{E}(u),U_tL_-(b\mathcal{E}(v))\rangle + \bar{u}_tv_t\langle a\mathcal{E}(u),U_tL_o(b\mathcal{E}(v))\rangle + \bar{u}_t\langle a\mathcal{E}(u),U_tL_+(b\mathcal{E}(v))\rangle + \langle a\mathcal{E}(u),U_tL(b\mathcal{E}(v))\rangle .$$

(Cf. §I, (15),(18),(22)). On a écrit partout $L_\varepsilon(b\mathcal{E}(v))$ au lieu de $L_\varepsilon\otimes I(\ )$.

Application : Définissons un opérateur J_t sur $\mathcal{B}$ par

$$\langle a,J_tb\rangle = \langle a\mathcal{E}(u),U_t(b\mathcal{E}(v))\rangle \qquad (u,v \text{ fixés})$$

Il résulte sans peine des majorations faites dans la démonstration d'existence que J_t est borné. On a

$$(15)\qquad \frac{d}{dt}J_t = J_t(v_tL_-+\bar{u}_tv_tL_o+\bar{u}_tL_++L)$$

qui est une équation d'évolution non homogène (sauf si u=v=0, cas où l'on a simplement $J_t=e^{tL}$).

La formule suivante est plus longue, c'est pourquoi on a posé $a\mathcal{E}(u)=\alpha$, $b\mathcal{E}(v)=\beta$. On adopte la même convention que plus haut (L_ε pour $L_\varepsilon\otimes I$). Pour la justification, voir §I, (16), (19),...(25).

$$(16)\quad \frac{d}{dt}\langle U_t(a\mathcal{E}(u)),U_t(b\mathcal{E}(v))\rangle =$$
$$v_t[\langle U_tL_+\alpha,U_t\beta\rangle + \langle U_t\alpha,U_tL_-\beta\rangle + \langle U_tL_+\alpha,U_tL_o\beta\rangle] +$$
$$+\ \bar{u}_tv_t[\langle U_tL_o\alpha,U_t\beta\rangle + \langle U_t\alpha,U_tL_o\beta\rangle + \langle U_tL_o\alpha,U_tL_o\beta\rangle] +$$
$$+\ \bar{u}_t[\langle U_tL_-\alpha,U_t\beta\rangle + \langle U_t\alpha,U_tL_+\beta\rangle + \langle U_tL_o\alpha,U_tL_+\beta\rangle] +$$
$$+\ [\langle U_tL\alpha,U_t\beta\rangle + \langle U_t\alpha,U_tL\beta\rangle + \langle U_tL_+\alpha,U_tL_+\beta\rangle] .$$

[Variante : si X est un opérateur borné sur $\mathcal{B}$, on peut calculer de même $\frac{d}{dt}\langle U_t\alpha,(X\otimes I)U_t\beta\rangle$: abrégeant comme plus haut $X\otimes I$ en X, le seul changement est le remplacement, partout au second membre, de U_t par XU_t à droite de la virgule].

Application : Comme plus haut, fixons u et v et définissons un opérateur borné K_t sur $\mathcal{B}$ par

$$\langle a,K_tb\rangle = \langle U_t(a\mathcal{E}(u)),U_t(b\mathcal{E}(v))\rangle$$

Alors on a en récrivant (16)

$$\frac{d}{dt}\langle a,K_tb\rangle = v_t[\langle L_+a,Kb\rangle + \langle a,KL_-b\rangle + \langle L_+a,KL_ob\rangle] + \bar{u}_tv_t[....$$

d'où pour K_t une équation d'évolution linéaire non homogène

$$(17)\quad \frac{d}{dt}K_t = v_t(L_+^*K_t+K_tL_-+L_+^*K_tL_o) + \bar{u}_t v_t(L_o^*K_t+K_tL_o+L_o^*K_tL_o)$$
$$+ \bar{u}_t(L_-^*K_t+K_tL_++L_o^*K_tL_+) + (L^*K_t+K_tL+L_+^*K_tL_+) .$$

[Variante. On aurait la même équation d'évolution (17) si l'on avait défini K_t par $\langle a,K_tb\rangle = \langle U_t(a\mathcal{E}(u)),XU_t(b\mathcal{E}(v))\rangle$.]

4. Application aux évolutions unitaires.

Dans cette section, nous allons déterminer les conditions, nécessaires et suffisantes, pour que les opérateurs U_t solutions de l'é.d.s. (10) soient prolongeables en des opérateurs unitaires sur $\mathfrak{G}\otimes\mathfrak{F}$.

Conditions nécessaires. Nous supposons les U_t isométriques. Le côté gauche de (16) est donc nul, et comme u,v sont arbitraires, nous obtenons les quatre équations

$$(18)\qquad L_+^*+L_-+L_+^*L_o=0 \ , \quad L_o^*+L_o+L_o^*L_o=0 \ , \ L_-^*+L_++L_o^*L_+=0 \ , \ L^*+L+L_+^*L_+=0 \ .$$

Nous considérons ensuite les opérateurs U_t^*, également isométriques : ils satisfont eux aussi à une é.d.s., et à des conditions analogues, parmi lesquelles nous retiendrons

$$(18')\qquad L_o^*+L_o+L_oL_o^*=0 \ .$$

On en déduit

$$(19_1)\qquad W = I+L_o \text{ est unitaire}$$

Après quoi la première et la troisième équation (18) sont équivalentes, et nous donnent

$$(19_2)\qquad L_+=-WL_-^*, \ L_-=-L_+^*W$$

Enfin, la dernière équation (18) peut s'écrire

$$(19_3)\qquad L = iH - \tfrac{1}{2}L_+^*L_+ = iH - \tfrac{1}{2}L_-L_-^*$$

où H est autoadjoint borné.

Conditions suffisantes. Inversement, donnons nous W,H,L_- (par exemple) et déterminons L_o par (19_1), L_+ par (19_2) et L par (19_3). Montrons que les U_t sont prolongeables en unitaires.

La famille d'opérateurs $K_t=\langle\mathcal{E}(u),\mathcal{E}(v)\rangle I$ satisfait à l'équation d'évolution (17), d'après ces hypothèses. Comme cette équation a une solution unique, pour laquelle $K_0=\langle\mathcal{E}(u),\mathcal{E}(v)\rangle I$, on a $\langle U_t(a\mathcal{E}(u)),U_t(b\mathcal{E}(v))\rangle = \langle a\mathcal{E}(u),b\mathcal{E}(v)\rangle$, et les U_t sont prolongeables en opérateurs isométriques - en particulier bornés, et l'on peut parler de leurs adjoints U_t^*, qui satisfont alors à une é.d.s. analogue. En raisonnant de même sur celle-ci, on voit que les U_t^* sont aussi isométriques, donc les U_t sont unitaires.

Redescente sur l'espace initial $\mathfrak{G}$. Etant donné un opérateur borné H

sur $\mathcal{B}\otimes\mathfrak{h}$, définissons un opérateur $\tilde{H} = \mathbf{E}_{\mathbb{1}}[H]$ sur $\mathcal{B}$ en posant

$$<a,\tilde{H}b> = < a\otimes\mathbb{1}, H(b\otimes\mathbb{1}) >$$

L'application $\mathbf{E}_{\mathbb{1}}$ est ainsi notée, parce qu'elle est analogue à une espérance conditionnelle par rapport à la position initiale, en probabilités classiques. En particulier, nous nous intéresserons aux opérateurs

$$P_t = \mathbf{E}_{\mathbb{1}}[U_t] \text{ sur } \mathcal{B} \quad , \text{ et } \quad \rho_t : X \longmapsto \mathbf{E}_{\mathbb{1}}[U_t^* X U_t] \text{ sur } \mathfrak{L}(\mathcal{B}) \ .$$

D'après (15) et (17)(variante), nous pouvons écrire

$$\frac{d}{dt}P_t = P_t L \qquad (L = iH - \tfrac{1}{2}L_-L_-^*)$$

$$\frac{d}{dt}\rho_t X = G(\rho_t X) \text{ où } G(Y) = L^*Y + YL + L_+^* Y L_+$$

$$= -i[H,Y] + (L_+^* Y L_+ - \tfrac{1}{2}L_+^* L_+ Y - \tfrac{1}{2} Y L_+^* L_+)$$

de sorte que ces familles d'opérateurs sont des semi-groupes à générateur borné, donc uniformément continus. Le premier est un semi-groupe de contractions de $\mathcal{B}$, le second un semi-groupe de noyaux sous-markoviens (cf. exposé I, fin du §III) sur l'espace des "mesures" $\mathfrak{M}(\mathcal{B})$: un semi-groupe d'opérateurs sur $\mathfrak{L}(\mathcal{B})$, positifs (même complètement positifs) et diminuant la trace. La théorie des é.d.s. a permis une nouvelle approche du problème des dilatations unitaires de tels semi-groupes d'opérateurs. Nous espérons en parler dans un exposé ultérieur.

§ III. OPERATEURS DEFINIS PAR DES NOYAUX.

1. Nous reprenons la théorie de Maassen, des opérateurs définis par des noyaux : notre but est de parvenir à un véritable calcul stochastique, pour une classe d'opérateurs stable par composition. Nous allons traiter de façon plus complète les opérateurs introduits de manière formelle au § III, exposé IV.

Nous commençons par introduire une classe restreinte de vecteurs-test. Identifiant le vecteur $\int_\rho f(H)dX_H \in \Phi$ à la fonction $f=(f_n)$ sur $L^2(\rho)$, nous dirons que f est un vecteur-test (au sens restreint, ou au sens de Maassen) si

1) f est à support borné : il existe $T<\infty$ tel que $f(H)=0$ pour $H \notin [0,T]$.
2) Il existe $M<\infty$ tel que $|f(H)| \leqq M^{|H|}$ p.p..

On a alors $\|f_n\|_2 \leqq M^n T^n/n!$, donc f est un vecteur-test au sens de l'exposé IV, III.2 . Les conditions précédentes sont satisfaites, d'autre part, par les vecteurs cohérents $\mathcal{E}(u)$ si u est bornée à support dans $[0,T]$, et par les éléments $f_n(s_1,\ldots,s_n)$ des chaos de Wiener, où f_n est bornée à support borné.

L'espace des vecteurs-test de Maassen est stable par transformation de Fourier-Wiener. Il n'est pas intrinsèque sur l'espace de Fock $\Phi(L^2(\mathbb{R}_+))$,

en ce sens que les conditions ne sont pas invariantes sous l'action des opérateurs $\Phi(U)$ associés à un opérateur unitaire U sur $L^2(\mathbb{R}_+)$.

Dans la suite du paragraphe, le mot vecteur-test est pris au sens restreint, sauf mention du contraire.

Nous définissons maintenant les noyaux. Nous avons vu dans l'exposé IV qu'il y en a deux classes : des noyaux K(A,B) permettant de représenter les opérateurs $a^{\pm}$ et une classe naturelle qu'ils engendrent par intégrales stochastiques, et des noyaux K(A,B,C) permettant d'introduire, de plus, les opérateurs $N_t=a^o_t$. Dans les deux cas, les conditions imposées seront les mêmes : outre la mesurabilité sur Ω^2 resp. Ω^3 :

1) Une condition de support borné : K=0 si AUB, ou AUBUC , n'est pas contenu dans [0,T].

2) Une condition de majoration : $|K(A,B,C)| \leqq M^{|A|+|B|+|C|}$ (ou[1] la forme analogue, plus simple, pour K(A,B)).

Nous faisons opérer les noyaux sur les vecteurs-test par la formule

$$(1) \qquad Kf(H) = \int_\Omega \Sigma_{A\subset H,\ B\subset H,\ A\cap B=\emptyset} K(A,B,C) f(C\cup H\setminus A)\,dC \quad (2)$$

Les noyaux à deux arguments sont les noyaux K(A,B,C) nuls pour $B\neq\emptyset$: ils opèrent par la formule plus simple

$$(1') \qquad Kf(H) = \int_\Omega \Sigma_{A\subset H} K(A,B) f(B\cup H\setminus A)\,dB \ .$$

On peut imposer K(A,B,C)=0 si A,B,C ne sont pas disjoints

Nous avons un premier résultat, très simple :

THEOREME. Si f est une fonction-test, Kf est une fonction-test.

Dém. Nous pouvons supposer que les constantes T et M figurant dans les définitions de K et f sont les mêmes. En (1), seuls les termes pour lesquels A,B,C,CUH\A sont contenus dans [0,T] donnent une contribution non nulle : comme A et H\A sont dans [0,T], on voit que Kf est à support dans [0,T] aussi. Nous avons ensuite, en décomposant suivant les valeurs de |C|=n, et en remarquant que C est p.s. disjoint de H fixé

$$|Kf(H)| \leqq \Sigma_{A,B}\ \Sigma_n\ M^{|A|+|B|+n} M^{n+|H|-|A|}\ \frac{T^n}{n!}$$

ce dernier facteur étant l'intégrale de dC sous les conditions |C|=n, C⊂[0,T]. Les facteurs $M^{|A|}$ disparaissent, $M^{|H|}$ sort, et l'on est ramené à

$$\Sigma_{B\subset H\setminus A} M^{|B|} = (1+M)^{|H\setminus A|}\ ,\quad \Sigma_{A\subset H} (1+M)^{|H\setminus A|} = (2+M)^{|H|}$$

$$|Kf(H)| \leqq (M(2+M))^{|H|}\ \Sigma_n\ M^{2n} T^n/n!$$

qui est une inégalité du type cherché.

---

1. Bien entendu, en jouant sur M cela équivaut à une condition du type $O(M^{|A|+|B|+|C|})$. 2. La convergence absolue de cette intégrale résulte de majorations analogues à celles que l'on fait plus bas.

2. Quelques exemples de noyaux.

a) Les deux formules (35), à la fin du §I de l'exp. IV, nous permettent d'écrire les noyaux correspondants aux opérateurs $a_h^{\pm}$ (h bornée à support borné). Ce sont des noyaux à deux arguments

(2)
$$K_h^-(A,B) = 0 \text{ si } |A|\neq 0 \text{ ou } |B|\neq 1 \; ; \; K_h^-(\emptyset,\{t\}) = \bar{h}(t)$$
$$K_h^+(A,B) = 0 \text{ si } |A|\neq 1 \text{ ou } |B|\neq 0 \; ; \; K_h^+(\{t\},\emptyset) = h(t) \; .$$

b) Dans l'exp. IV, §II, formule (31), nous avons écrit la formule de multiplication de Wiener

(3)
$$g\times f(H) = \int_{\rho} \Sigma_{A\subset H}\, g(A\cup B) f(B\cup H\setminus A)\, dB$$

qui montre que le noyau de l'opérateur M_g de multiplication de Wiener par une fonction-test g est simplement $M_g(A,B)=g(A\cup B)$.

c) Cela va nous permettre de déterminer les noyaux de certains opérateurs de Weyl, $W_{ih} = e^{iQh}$ (h réelle, bornée à support borné). En effet, ceci est un opérateur de multiplication de Wiener par $e^{i\tilde{h}} = \mathcal{E}(ih)e^{-\|h\|^2/2}$, et le vecteur cohérent $\mathcal{E}(ih)$ se lit comme la fonction-test

$$g(A) = \prod_{s\in A} (ih(s))$$

Ainsi le noyau de l'opérateur de Weyl W_k pour k=ih (h réelle) est

(4)
$$e^{-\|h\|^2/2} \prod_{s\in A\cup B} (ih(s)) = e^{-\|k\|^2/2} \prod_{s\in A} k(s) \prod_{s\in B} (-\bar{k}(s))$$

Nous allons en déduire le noyau de l'opérateur de Weyl $W_h=e^{-iP_h}$ pour h réelle, grâce à la remarque suivante. Soit $\mathcal{F}$ la transformation de Fourier-Wiener (multiplication par i^n sur le n-ième chaos). Si un opérateur X est représenté par un noyau K(A,B), l'opérateur $\mathcal{F}^{-1}X\mathcal{F}$ admet le noyau $i^{|B|-|A|}K(A,B)$ et l'opérateur $\mathcal{F}X\mathcal{F}^{-1}$ le noyau $i^{|A|-|B|}K(A,B)$. Ici on a $-P_h=\mathcal{F}Q_h\mathcal{F}^{-1}$, et la même relation pour les exponentielles. Par conséquent, le noyau de W_h pour h réelle est

(4')
$$e^{-\|h\|^2/2}\, i^{|A|-|B|} \prod_{s\in A\cup B} (ih(s)) = e^{-\|h\|^2/2} \prod_{s\in A} h(s) \prod_{s\in B} (-\bar{h}(s)) \; .$$

Maassen montre que la formule valable en toute généralité pour le noyau de W_f (f complexe, bornée à support borné) est

(5)
$$e^{-\|f\|^2/2} \prod_{s\in A} f(s) \prod_{s\in B} (-\bar{f}(s)) \; .$$

Cette formule peut en principe se déduire des deux cas particuliers que nous venons d'indiquer, et de la formule de composition des noyaux, que nous donnons plus loin (Maassen procède différemment).

d) Nous passons à l'opérateur a_h^o , qui va justifier l'introduction des noyaux à trois arguments.

Nous prenons

(6) $K(A,B,C)=0$ si $|A|\neq 0$ ou $|B|\neq 1$ ou $|C|\neq 0$, $K(\emptyset,\{t\},\emptyset)=h(t)$

et alors

$$Kf(H) = \int_{\mathcal{P}_0} \sum_{t\in H} K(\emptyset,\{t\},C)f(C\cup H)dC = (\Sigma_{t\in H}\, h(t))f(H)$$

qui représente bien l'opérateur $a_h^\circ=N_h=\lambda(m_h)$ (exposé IV, §II.4, et §I, formule (18) : rappelons que $\lambda(X)$ est la "seconde quantification différentielle" d'un opérateur X sur $L^2(\mathbb{R}_+)$, et m_h l'opérateur de multiplication par h).

(Autres exemples au n°5)

3. Composés et adjoints

Les résultats de cette section (indiqués sans démonstration par Maassen) reposent sur la formule (33) de l'exposé IV, §II.7, que nous recopions sous la forme qui nous convient ici :

$$(7)\qquad \int_{\mathcal{P}} \Big(\sum_{A+A'=U} \phi(A,A')\Big)dU = \int_{\mathcal{P}\times\mathcal{P}} \phi(A,A')dAdA' \ .$$

Nous commençons par le résultat le plus simple, qui concerne les adjoints : si l'on pose $K^*(A,B)=\overline{K(B,A)}$, on a pour tout couple (f,g) de vecteurs-test

$$(9)\qquad < g,Kf > = < K^*g,f >$$

ce qui revient à vérifier que

$$(10)\qquad \int \overline{g}(H) \sum_{A+A'=H} K(A,B)f(A'\cup B)\,dHdB = \int f(\lambda) \sum_{\alpha+\alpha'=\lambda} K(\mu,\alpha)\overline{g}(\alpha'\cup\mu)d\lambda d\mu$$

Nous transformons la première et la seconde expression respectivement en

$$\int \overline{g}(A\cup A')K(A,B)f(A'\cup B)dAdA'dB \text{ et } \int f(\alpha\cup\alpha')K(\mu,\alpha)\overline{g}(\alpha'\cup\mu)d\alpha d\alpha' d\mu$$

grâce à (7). Pour vérifier l'égalité il suffit de poser $\mu=A$, $\alpha=B$, $\alpha'=A'$.

On a un résultat analogue pour les noyaux à trois arguments, en posant $K^*(A,B,C)=\overline{K(C,B,A)}$. Cette fois ci, on a

$$<g,Kf> = \int g(H) \sum_{A+A'+A''=H} K(A,A',C)f(C\cup A'\cup A'')dHdC$$

$$= \int g(A\cup A'\cup A'')K(A,A',C)f(C\cup A'\cup A'')dAdA'dA''dC$$

par une double application de (7), et on continue comme ci-dessus.

Nous passons à la composition des noyaux : désignant par la même lettre le noyau et l'opérateur associé, le résultat de Maassen nous dit (dans le cas des noyaux à deux arguments) que si K,L sont deux noyaux, l'opérateur composé LK est associé au noyau (nous laissons le lecteur vérifier que c'en est bien un)

$$(11)\qquad M(A,B) = \int \sum_{\substack{U+U'=A\\ V+V'=B}} L(U,V\cup C)K(U'\cup C,V')dC$$

Un calcul sans mystère donne en effet

$$Mf(H) = \int \Sigma_{R+R'+R''=H,\, S+S'=T} L(R,S\cup C)K(R'\cup C,S')f(T\cup R'')dCdT$$

qui se transforme, grâce à (7), en

$$(\square) \qquad \int \Sigma_{R+R'+R''=H} L(R,S\cup C)K(R'\cup C,S')f(S\cup S'\cup R'')dSdS'dC$$

D'autre part, un autre calcul sans mystère donne

$$LKf(H) = \int \Sigma_{\alpha+\alpha'+\alpha''=H,\ \beta+\beta'=\varepsilon} L(\alpha,\varepsilon)K(\alpha'\cup\beta,\gamma)f(\alpha''\cup\beta'\cup\gamma)d\varepsilon d\gamma$$

qui devient de la même façon

$$(\square\square) \qquad \int \Sigma_{\alpha+\alpha'+\alpha''=H} L(\alpha,\beta\cup\beta')K(\alpha'\cup\beta,\gamma)f(\alpha''\cup\beta'\cup\gamma)d\beta d\beta' d\gamma$$

Pour identifier ($\square$) et ($\square\square$), il suffit de poser

$$\alpha=R \quad \alpha'=R' \quad \alpha''=R'' \quad \beta=C \quad \beta'=S \quad \gamma=S' \ .$$

La formule de composition des noyaux à trois arguments est nettement plus compliquée : en posant M=LK comme dans (11), on a

$$(12) \qquad M(A,B,C) = \int \sum_{\substack{\alpha+\alpha'+\alpha''=A\\ \beta+\beta'+\beta''=B\\ \gamma+\gamma'+\gamma''=C}} L(\alpha,\alpha'\cup\beta\cup\beta',\gamma\cup\gamma'\cup X)K(X\cup\alpha'\cup\alpha'',\beta'\cup\beta''\cup\gamma',\gamma'')dX$$

J'ai bien vérifié que cette formule donne correctement le composé, et que M est un noyau ; je n'ai pas vérifié que l'opération ainsi définie entre noyaux est associative (ce qui ne résulte pas de l'associativité de la composition : un opérateur ne détermine pas uniquement son noyau).

4. Intégrales stochastiques

Nous abordons la partie la plus surprenante du travail de Maassen : la manière dont se calcule le noyau d'une intégrale stochastique d'opérateurs

$$I_T^\varepsilon(K) = \int_0^T K_s da_s^\varepsilon \qquad (\varepsilon = -,\circ,+)$$

pour une famille adaptée (K_s) d'opérateurs associés à des noyaux (nous identifions désormais un opérateur à son noyau). Dire que la famille est adaptée voudra dire, sur les noyaux correspondants, que

$$K_s(A,B,C)=0 \text{ si } A\cup C \not\subset [0,s[\ , \ K_s(A,B,C)=K_s(A,B\cap[0,s[,C)$$

(nous écrivons [0,s[plutôt que [0,s], parce que cela correspond à la prévisibilité classique, qui est une hypothèse naturelle). Remarquons de manière heuristique que le <<noyau>> de l'<<opérateur>> da_s^ε est donné par

$$da_s^-(\emptyset,\emptyset,\{s\})=1 \quad , \quad da_s^\circ(\emptyset,\{s\},\emptyset)=1 \quad , \quad da_s^+(\{s\},\emptyset,\emptyset)=1$$

et $da_s^\varepsilon(A,B,C)=0$ dans tous les autres cas. Ainsi le noyau de $K_s da_s^\varepsilon = dI_s^\varepsilon$ s'obtient ainsi ($\vee A$ désignant le plus grand élément de l'ensemble A, et A- l'ensemble $A\setminus\{\vee A\}$

$$dI_s^+(A,B,C) = K_s(A-,B,C) \text{ si } \vee A=s \ , \ 0 \text{ sinon}$$
$$dI_s^\circ(A,B,C) = K_s(A,B-,C) \text{ si } \vee B=s \ , \ 0 \text{ sinon}$$
$$dI_s^-(A,B,C) = \ldots$$

Toujours de manière heuristique, on est amené aux formules suivantes pour les noyaux des intégrales stochastiques I^ε_∞ (on élimine T de la notation : il suffit de remplacer K_s par 0 pour s>T)

(13) $I^\varepsilon(A,B,C) = K_{VA}(A-,B,C)$, $K_{VB}(A,B-,C)$, $K_{VC}(A,B,C-)$ $(\varepsilon=+,\circ,-)$.

Noter une formule analogue pour $g = \int f_s dB_s$ (i.s.ordinaire): $g(A) = \delta_{VA}(A-)$.

Nous allons vérifier que ces formules sont correctes (par une méthode différente de celle de Maassen, dont on parlera plus loin). Tout d'abord, si les K_t sont des noyaux, pour lesquels la condition de majoration a lieu uniformément en t sur tout intervalle [0,T], il est facile de vérifier que I^ε_T est un noyau. Prenant $\varepsilon=+$ par exemple, nous allons vérifier la formule

$$<\mathcal{E}(u),I^\varepsilon\mathcal{E}(v)> = \int <\mathcal{E}(u),K_t\mathcal{E}(v)>\bar{u}_t dt$$

qui caractérise l'intégrale stochastique (§I, (18)). Nous avons comme ci-dessus supposé que $K_t=0$ pour t>T, de sorte que tout se passe sur [0,T].

a) Nous partons d'une formule déjà utilisée au n°3 (entre (10) et (11)) mais avec des notations différentes ; L est un noyau quelconque

$$(14) \qquad <f,Lg> = \int \bar{f}(A\cup U\cup V)L(U,V,W)g(V\cup W\cup A)dAdUdVdW$$

b) Nous remplaçons f,g par $\mathcal{E}(u),\mathcal{E}(v)$: ainsi $f(A)=\prod_{s\in A} u(s)$ - notons en passant que $\int f(A)dA = \exp(\int f(s)ds)$, comme on le vérifie en explicitant la mesure dA . Comme dans (14) les quatre ensembles A,U,V,W sont p.s. disjoints, l'intégrale en A sort, et il reste

$$(15) <\mathcal{E}(u),L\mathcal{E}(v)> = \Big(\int \prod_{s\in A} \bar{u}v(s)dA\Big)\Big(\int \prod_{s\in U\cup V} \bar{u}(s) \prod_{s\in V\cup W} v(s)L(U,V,W)dUdVdW\Big)$$

le premier facteur valant $\exp(\int \bar{u}vds)=<\mathcal{E}(u),\mathcal{E}(v)>$.

c) Nous remplaçons L par $I^+(U,V,W)=K_{VU}(U-,V,W)$. Nous appliquons le théorème de Fubini, et fixons V,W en posant $k_t(U)=K_t(U,V,W)$, $i^+(U)= I^+(U,V,W)=k_{VU}(U-)$. La formule à vérifier est alors

$$\int \prod_{s\in U} \bar{u}(s)i^+(U)dU = \int \bar{u}(t)dt \int \prod_{s\in A} \bar{u}(s)k_t(A)dA$$

qui se ramène à la formule évidente

$$\int_{s_1<\ldots<s_n<t} \bar{u}(s_1)\ldots\bar{u}(s_n)\bar{u}(t)f(s_1,\ldots,s_n,t)ds_1\ldots ds_n dt$$
$$= \int \bar{u}(t)dt \int_{s_1<\ldots<s_n} \bar{u}(s_1)\ldots\bar{u}(s_n)f(s_1,\ldots,s_n,t)ds_1\ldots ds_n$$

pour une fonction $f(s_1,\ldots,s_n,t)$ nulle si $\{s_1,\ldots,s_n\} \not\subset [0,t[$. Les autres formules se vérifient de même.

COMMENTAIRE. On voit l'assouplissement que la théorie des noyaux apporte au calcul stochastique de Hudson-Parthasarathy : les opérateurs ne sont plus définis seulement sur les vecteurs cohérents, mais sur un domaine

commun stable (les fonctions test) ; ils sont composables et admettent des adjoints ; sous des conditions de norme très raisonnables, les intégrales stochastiques d'opérateurs donnés par des noyaux sont encore données par des noyaux. Par exemple, tous les opérateurs I_t^α du paragraphe précédent, formule (11), sont des noyaux (dont la norme doit être facile à estimer). Dans ces conditions, les << formules d'Ito >> peuvent être mises sous leur forme habituelle de différentielle stochastique d'un composé d'opérateurs, et non plus sous la forme un peu maladroite des paragraphes précédents.

Il resterait cependant beaucoup de détails à écrire, et nous préférons nous arrêter.

REMARQUE. Avec les notations précédentes, on a aussi

$$\int k_t(U)dUdt = \int i^+(U)dU$$

et de même, bien sûr, pour $|k_t(U)|^2$ et $|i^+(U)|^2$. Cela signifie (remarque Maassen) que si l'on munit l'espace des noyaux de la norme $\|K\|_2$ (norme dans $L^2(dAdBdC)$, qui bien sûr n'est pas la norme de Hilbert-Schmidt de l'opérateur associé), et l'espace des processus adaptés de noyaux de la norme $(\int \|K_t\|_2^2 dt)^{1/2}$, l'intégration par rapport à da^+, et de même da^-, da°, est isométrique comme dans le cas classique.

5. Application aux fermions.

Rappelons la formule de multiplication de Clifford de deux fonctions

$$g*f(H) = \int_\rho (\Sigma_{A+B=H}\, g(A\cup M)f(B\cup M))(-1)^{n(A\cup M, B\cup M)}dM \tag{16}$$

Il ne semble pas que l'opérateur $g*$ (où g est une fonction-test) soit donné par un noyau - en tout cas, ce n'est pas évident, contrairement à la multiplication de Wiener. Mais nous allons voir qu'il en est bien ainsi lorsque g appartient au premier chaos . Nous retrouverons ainsi, par la méthode de Maassen, les résultats de Hudson-Parthasarathy sur la représentation des opérateurs de création et d'annihilation de fermions comme intégrales stochastiques par rapport aux opérateurs correspondants pour les bosons (voir leur preprint : Unification of Fermion and Boson Stochastic Calculus, Juin 1985)[1]

a) La méthode de H-P repose sur l'emploi de l'opérateur $J_t=(-1)^{N_t}$, qui est aussi la seconde quantification (exposé IV, §1, (15))

1. J'ai appris tout récemment l'existence d'une série de travaux de P. Garbaczewski sur l'unification des bosons et des fermions. Le plus proche de notre point de vue est : Representations of the CAR generated by representations of the CCR Fock space. Comm.M.Phys. 43, 1975, 131-136.

$\Phi(M_t)$ de l'opérateur M_t de multiplication par la fonction

$$m_t(s)=-1 \text{ pour } s<t,\ m_t(s)=+1 \text{ pour } s\geq t\ .$$

Notre premier travail va consister à montrer que cet opérateur J_t est donné par un noyau de la forme (t omis pour alléger)

(17) $\qquad J(\emptyset,B,\emptyset) = j(B)\ ,\quad J(A,B,C)=0 \text{ si } A\cup C\neq\emptyset\ .$

Comment opère un noyau de la forme (17) ? On a d'après (1)

(18) $\qquad Jf(H) = f(H)\phi(H)\ ,\ \text{avec}\ \phi(H)=\Sigma_{B\subset H}\, j(B)$

Mais inversement, la fonction ϕ étant donnée, on peut calculer j par la formule d'inversion de Moebius

$$j(B) = \Sigma_{C\subset B}\,(-1)^{|B-C|}\phi(C)$$

Le cas qui nous intéresse est celui où $\phi(C)=(-1)^{|C\cap[0,t[|}$, de sorte que $j(B)=(-2)^{|B\cap[0,t[|}$: l'ordre de croissance est bien celui qui convient à un noyau ; $|B\cap[0,t[|$ s'écrit plus agréablement $n(t,B)$.

Nous rétablissons l'indice t, et calculons les intégrales stochastiques

(19) $\qquad \beta_g^{\pm} = \int g(s)J_s da_s^{\pm}\quad (\ g \text{ bornée à support borné }\ )$

Appliquant (13), nous trouvons pour ces opérateurs les noyaux

$$\beta_g^{+}(\{s\},B,\emptyset) = \beta_g^{-}(\emptyset,B,\{s\}) = g(s)(-2)^{n(s,B)}$$

et 0 dans les autres cas. Par conséquent

$$\beta_g^{+}f(H) = \Sigma_{s\in H}\,(-1)^{n(s,H)}g(s)f(H\setminus\{s\})$$

$$\beta_g^{-}f(H) = \int ds\,(-1)^{n(s,H)}g(s)f(H\cup\{s\})$$

Si l'on compare cela à la formule (16), dans laquelle on prend $g(A)=0$ si $|A|\neq 1$, $g(\{s\})=g(s)$, on voit que l'<u>opérateur de multiplication de Clifford par un élément</u> g <u>du premier chaos est égal à</u> $\beta_g^{+}+\beta_g^{-}$. Pour établir que β_g^{+} et β_g^{-} sont les opérateurs de création et d'annihilation fermioniques, il faut encore travailler un peu, mais nous nous arrêterons là.

FIN PROVISOIRE DES NOTES