ANNALES DE L'I. H. P.

E. SCHRÖDINGER

Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique

Annales de l'I. H. P., tome 2, nº 4 (1932), p. 269-310

<http://www.numdam.org/item?id=AIHP_1932__2_4_269_0>

© Gauthier-Villars, 1932, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P. » implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique

PAR

E. SCHRÖDINGER

I. — Introduction

J'ai l'intention d'exposer dans ces conférences diverses idées concernant la mécanique quantique et l'interprétation qu'on en donne généralement à l'heure actuelle ; je parlerai principalement de la théorie quantique relativiste du mouvement de l'électron. Autant que nous pouvons nous en rendre compte aujourd'hui, il semble à peu près sûr que la mécanique quantique de l'électron, sous sa forme idéale, *que nous ne possédons pas encore*, doit former un jour la base de toute la physique. A cet intérêt tout à fait général, s'ajoute, ici à Paris, un intérêt particulier : vous savez tous que les bases de la théorie moderne de l'électron ont été posées à Paris par votre célèbre compatriote Louis de BROGLIE.

Les recherches que je vais exposer ne forment nullement une théorie nette et complètement achevée (1). Le lien commun, un peu lâche d'ailleurs, qui les rattache les unes aux autres, la source commune dont elles dérivent, est le mécontentement que l'on éprouve quand on considère l'état présent de la théorie et surtout celui de l'*interprétation physique actuelle* de la mécanique quantique. Je voudrais

(1) Les mémoires originaux, qui forment la base de ces conférences, ont été publiés dans les *Sitzungsberichte der preussischen Akademie der Wissenschaften*, 1930, p. 418 ; 1931, pp. 63, 144, 238. Dans les pages qui vont suivre, quelques-uns des aspects des problèmes envisagés sont peut-être un peu mieux précisés ; on y trouvera également des résultats nouveaux (v. Notes I-III).

essayer d'attirer votre attention sur les grandes difficultés qui s'y présentent et dont la plus sévère et peut-être la plus inattendue concerne la réconciliation des conceptions de la théorie de la relativité restreinte d'une part, et de la mécanique quantique d'autre part.

Commençons par fixer les notions fondamentales pour être assurés que nous nous entendons bien entre nous. En mécanique quantique figurent deux espèces d'êtres mathématiques : les fonctions $\psi(x, t)$ et les opérateurs (linéaires et hermitiques) $\mathfrak{A}$. ψ est une fonction complexe des coordonnées d'un système physique et du temps t, qui est traité comme un paramètre ; ψ décrit *l'état* de ce système à un moment déterminé. Un opérateur est une loi permettant de former à partir d'une fonction quelconque ψ des coordonnées une autre fonction des mêmes arguments (1)

$$\psi \rightarrow \mathfrak{A}\psi.$$

La signification physique des opérateurs, est la suivante : à chaque quantité physique que l'on suppose *mesurable*, ou observable dans le système considéré, correspond un opérateur déterminé ; par exemple, à la coordonnée x correspond l'opérateur « multiplication par x », etc... Or, étant donné l'état ψ, si l'on répète un grand nombre de fois la mesure de la « quantité $\mathfrak{A}$ », on ne trouvera pas toujours la même valeur, mais la *moyenne* de ces mesures sera donnée par

$$\overline{\mathfrak{A}} = \int \psi^* \mathfrak{A} \psi \, dx,$$

où ψ^* est la conjuguée complexe de ψ ; $\int dx$ indique que l'intégration porte sur toutes les configurations du système. *Telle est l'interprétation généralement adoptée aujourd'hui.* Elle contient déjà l'indication que ψ fournit non seulement la valeur moyenne, mais encore toute la statistique de $\mathfrak{A}$, c'est-à-dire de chaque observable. Admettons qu'une fonction analytique $f(\mathfrak{A})$ — un *opérateur* — soit définie par une série de puissances :

$$f(\mathfrak{A}) = b_0 + b_1\mathfrak{A} + b_2\mathfrak{A}^2 + \cdots,$$

où $\mathfrak{A}^n$ signifie l'opération $\mathfrak{A}$ répété n fois.

(1) $\mathfrak{A}$ opère (en général) uniquement sur les coordonnées et non sur le temps, c'est-à-dire définit une loi reliant entre elles des fonctions des *coordonnées* et non des fonctions des coordonnées *et* du temps ; cependant l'opérateur $\mathfrak{A}$ peut dépendre du temps qui figure dans son expression, comme paramètre (voir ce qui suit).

En particulier, considérons une fonction comme celle indiquée sur la figure : nulle partout, sauf entre a_1 et a_2 où elle prend la valeur 1. Il est

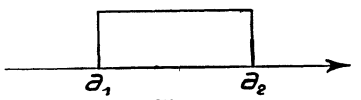

Fig. 1.

très facile d'approcher une telle fonction par une suite de fonctions analytiques. La valeur moyenne $\overline{f(\mathcal{A})}$ de la fonction $f(\mathcal{A})$ définie *par cette suite*, est évidemment la *probabilité* pour que $\mathcal{A}$ soit compris entre a_1 et a_2. On montre aussi très facilement que les seules valeurs possibles — les seules qui aient une probabilité différente de zéro — sont les valeurs propres de l'opérateur $\mathcal{A}$. Je ne peux pas insister sur ce point.

L'opérateur $\mathcal{H}$, qui correspond à *l'énergie* du système définit en même temps, la nature dynamique du système, c'est-à-dire la variation spontanée de ψ avec le temps par l'équation

$$\frac{h}{2\pi i}\frac{\partial\psi}{\partial t} = -\mathcal{H}\psi. \tag{1}$$

Il est facile de résoudre cette équation d'une manière tout à fait générale. La solution est

$$\psi(x, t) = e^{-\frac{\mathcal{H}t}{\varkappa}}\psi(x, 0), \qquad \varkappa = \frac{h}{2\pi i}.$$

$e^{\frac{\mathcal{H}t}{\varkappa}}$ est définie par une série de puissances bien connue. La vérification de cette solution est très aisée ; naturellement elle est encore assez implicite.

Il y a deux méthodes pour arriver à des énoncés plus explicites dans le problème posé par une « équation d'ondes » telle que (1). L'une, bien connue, est la recherche des valeurs propres et des fonctions propres de l'équation

$$\mathcal{H}\psi = E\psi,$$

ou, en d'autres termes, le développement de la solution en série de FOURIER relative au *temps*. C'est la méthode de la mécanique ondulatoire.

L'autre méthode est celle du *calcul opérationnel*. On cherche à éviter l'étude de la variation temporelle de ψ, et l'on se demande : n'y a-t-il pas à chaque instant un opérateur $\mathcal{A}(t)$, qui doit *dépendre* du paramètre t, tel que cet opérateur donne avec la fonction $\psi(x, o)$ la même valeur

moyenne et la même statistique que $\mathcal{A}$ donne avec $\psi(x, t)$? La réponse est *affirmative*. Et la variation temporelle de cet $\mathcal{A}(t)$ est, pour un opérateur quelconque, donnée par la même formule simple :

$$\varkappa \frac{d\mathcal{A}(t)}{dt} = \mathcal{H}\mathcal{A}(t) - \mathcal{A}(t)\mathcal{H}, \tag{2}$$

dont la solution générale est

$$\mathcal{A}(t) = e^{\frac{\mathcal{H}t}{\varkappa}} \mathcal{A}(0) e^{-\frac{\mathcal{H}t}{\varkappa}}.$$

L'avantage de cette méthode est que tout ce qu'on en déduira sera valable pour un état initial quelconque $\psi(x, o)$. On n'est pas obligé de le fixer à l'avance. En général on *supprime* l'argument t dans l'énoncé de la relation (2).

Signalons un cas particulier de grande importance, à savoir le cas, où $\mathcal{A}$ *commute* avec $\mathcal{H}$. La statistique d'une telle quantité physique est alors indépendante du temps. On peut dire qu'alors $\mathcal{A}$ est une « intégrale première » du système considéré.

II. — L'électron de Dirac.

Nous appliquerons ces faits bien connus à l'électron de DIRAC dans le cas de l'absence de champ. L'opérateur hamiltonien est dans ce cas le suivant :

$$\mathcal{H} = c(\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_4 mc),$$

Les p_k sont les opérateurs $\varkappa \frac{\partial}{\partial x_k}$, et x_1, x_2, x_3 représentent x, y, z. Les α_k, ($k = 1, 2, 3, 4$,) opèrent sur une variable différente de x_1, x_2, x_3, qu'on peut appeler ζ, et qui n'a que quatre valeurs possibles. Donc les α_k sont des matrices de 4×4 éléments. Mais tout ce qu'il est nécessaire de connaître de ces matrices sont les relations de commutation :

1) Ces matrices commutent naturellement avec chaque opérateur qui comme x_k, p_n, etc... n'opère pas sur ζ ; et

$$\alpha_k \alpha_l + \alpha_l \alpha_k = 2\delta_{kl}. \tag{2}$$

En appliquant le calcul opérationnel (équation (2)), on prouve facilement que les α_k ($k = 1, 2, 3$) ou plutôt les $c\alpha_k$ sont les opérateurs

qui correspondent aux composantes de la vitesse de l'électron. En effet, par la relation bien connue

$$p_k x_k - x_k p_k = \varkappa \cdot 1, \qquad (k = 1, 2, 3)$$

on trouve

$$\varkappa \frac{dx_k}{dt} = \mathcal{H} x_k - x_k \mathcal{H} = \varkappa c \alpha_k.$$

$$\frac{dx_k}{dt} = c\alpha_k.$$

Un des traits les plus intéressants de l'équation de DIRAC consiste précisément dans le fait que les notions de *moment* et de *vitesse* se *séparent*. L'opérateur de vitesse commute avec les coordonnées, elles sont donc observables simultanément, tandis que x_k et p_k se comportent comme en mécanique quantique ordinaire.

Ce qu'il y a de plus intéressant, c'est que même en l'absence de champ, les composantes de la vitesse *ne sont pas des intégrales premières*. Elles ne commutent pas avec $\mathcal{H}$. On trouve une valeur simple non pas pour leur *commutateur*, mais pour leur *anticommutateur* avec $\mathcal{H}$:

$$\mathcal{H}\alpha_k + \alpha_k \mathcal{H} = 2cp_k. \qquad (k = 1, 2, 3).$$

Donc

$$\varkappa \frac{d\alpha_k}{dt} = 2\mathcal{H}\alpha_k - 2cp_k = 2\mathcal{H}(\alpha_k - c\mathcal{H}^{-1}p_k) = 2\mathcal{H}\eta_k$$

en posant

$$\eta_k = \alpha_k - c\mathcal{H}^{-1}p_k.$$

On en tire, puisque $\mathcal{H}$ et p_k commutent avec $\mathcal{H}$,

$$\varkappa \frac{d\eta_k}{dt} = 2\mathcal{H}\eta_k,$$

qu'on peut intégrer :

$$\eta_k = \alpha_k - c\mathcal{H}^{-1}p_k = e^{\frac{2\mathcal{H}t}{\varkappa}} \eta_k^0.$$

Comme α_k correspond à $\frac{1}{c}\frac{dx_k}{dt}$, on obtient pour x_k, par une seconde intégration :

$$x_k = a_k + c^2\mathcal{H}^{-1}p_k t + \frac{c\varkappa}{2}\mathcal{H}^{-1}e^{\frac{2\mathcal{H}t}{\varkappa}} \eta_k^0.$$

Posons :

$$a_k + c^2\mathcal{H}^{-1}p_k t = \widehat{x}_k$$

$$\frac{cx}{2}\mathcal{H}^{-1}(\alpha_k - c\mathcal{H}^{-1}p_k) = \xi_k$$

Ou a :

$$x_k = \widehat{x}_k + \xi_k.$$

L'opérateur des coordonnées se décompose donc en deux parties. La première est une fonction linéaire du temps, comme on l'aurait attendu pour l'opérateur x_k tout entier, et même avec le facteur de t correct : p_k correspond à $\frac{mv_k}{\sqrt{1-\beta^2}}$, $\mathcal{H}^{-1}$ à $\frac{\sqrt{1-\beta^2}}{mc^2}$, par suite $c^2\mathcal{H}^{-1}p_k$ correspond à $\frac{1}{c^2}v_k \times c^2$, donc à v_k.

Mais il y a encore une autre partie, qui correspond à une sorte d'oscillation ou de « tremblement ». En calculant la valeur moyenne de x_k, on s'aperçoit qu'en général la valeur moyenne de la seconde partie *n'est pas nulle* : l'oscillation du « centre de gravité électrique » est parfaitement réelle.

Pourtant, il y a deux remarques à faire. La première est que l'*amplitude* de la vibration est toujours très petite. Posons

$$x_k = \widehat{x}_k + \xi_k$$

On trouve aisément

$$\xi_k^2 = \frac{h^2c^2}{16\pi^2}\mathcal{H}^{-2}(1 - \mathcal{H}^{-2}c^2p_k^2) = \text{constante}.$$

Mais

$$\mathcal{H}^2 = c^2(p_1^2 + p_2^2 + p_3^2 + m^2c^2).$$

Donc les valeurs propres de $\mathcal{H}^2$ sont toutes supérieures à m^2c^4, $|\mathcal{H}| > mc^2$. Donc celles de $\mathcal{H}^{-2}$ sont inférieures à $\frac{1}{m^2c^4}$ et l'on a :

$$\text{valeurs propres de } \xi_k^2 \leqslant \frac{h^2}{16\pi^2m^2c^2},$$

$$|\text{valeurs propres de } \xi_k| \leqslant \frac{h}{4\pi mc} \sim 10^{-11}\text{cm}.$$

La seconde remarque se rapporte au fait que le tremblement du centre de gravité s'annule dans un cas spécial, à savoir dans le cas où la fonction ψ est ce que nous appellerons plus tard « une fonction purement positive », c'est-à-dire qui ne contient que des fonctions propres de $\mathcal{H}$

appartenant à des valeurs propres positives. Cela deviendra bientôt plus clair. Pour le moment notons seulement le fait mathématique que ξ_k et η_k anticommutent avec $\mathcal{H}$:

$$\mathcal{H}\xi_k + \xi_k\mathcal{H} = 0$$
$$\mathcal{H}\eta_k + \eta_k\mathcal{H} = 0,$$

fait qui sera important dans la suite.

III. — Opérateurs pairs et impairs

Les opérateurs qui commutent avec $\mathcal{H}$ (intégrales en l'absence de champ) et les opérateurs qui anticommutent avec $\mathcal{H}$ (comme ξ_k, η_k) forment des cas spéciaux d'une classification importante que nous allons introduire maintenant : classification en opérateurs pairs et impairs.

Nous venons de prouver que les valeurs propres de $\mathcal{H}$ sont supérieures en valeur absolue à

$$|\mathcal{H}| \geqslant mc^2.$$

Donc elles sont ou $\geqslant mc^2$, ou $\leqslant -mc^2$. Il est facile de prouver que réciproquement une valeur quelconque, extérieure à l'intervalle $(-mc^2, +mc^2)$ est une valeur propre de $\mathcal{H}$, et qu'il lui correspond une fonction propre (une onde plane par exemple). On sait que c'est là la plus grosse difficulté de la théorie de DIRAC. En effet les valeurs propres négatives n'ont pas de signification physique ; on voudrait bien s'en débarrasser. Au moins il devrait être impossible qu'une fonction propre « positive » se transforme au cours du temps en donnant naissance à des fonctions « négatives », ou tout au moins cette variation ne devrait se produire qu'infiniment lentement, pour rendre suffisamment improbable l'*énorme* changement d'énergie $2mc^2$ que nous n'avons jamais observé.

On prévoit que dans cet ordre d'idées, il deviendra important de distinguer les opérateurs qui en opérant sur une fonction propre « positive » n'en déduisent qu'un ensemble de fontions propres « positives ». La méthode la plus simple de faire cette distinction est de représenter tous les opérateurs dans le système de matrices où $\mathcal{H}$ est diagonale. Considérons une telle représentation (v. Fig. 2). La croix centrale est vide, le champ d'éléments de matrice se décomposent en quatre régions.

Nous appellerons positive une fonction $\psi(x)$ qui, développée en série selon les fonctions propres de $\mathcal{H}$ (toujours l'énergie en l'absence de champ !) ne contient que des fonctions propres positives. Une fonction négative est définie de la même manière. Un opérateur dont les élé-

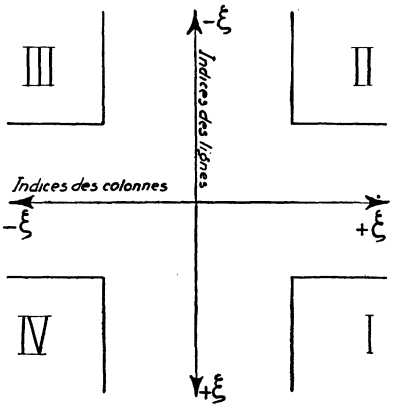

Fig. 2.

ments de matrice appartiennent uniquement aux régions I et III ne changera pas le caractère positif ou négatif d'une fonction positive ou négative. Nous l'appelerons opérateur « pair ». Pour les opérateurs « impairs » les éléments de matrice appartiennent aux régions II et IV ; appliqués à une fonction positive ces opérateurs la rendent négative et vice versa. Maintenant, il est facile de voir que les opérateurs qui commutent avec $\mathcal{H}$ sont pairs, ceux qui anticommutent avec $\mathcal{H}$ sont impairs. Par exemple, si pour l'équation

$$\mathcal{H}\xi_k + \xi_k\mathcal{H} = 0,$$

nous écrivons l'élément de matrice $(\rho\sigma)$ dans le schéma où $\mathcal{H}$ est diagonal, nous obtenons

$$(\mathcal{H}_\rho + \mathcal{H}_\sigma)(\xi_k)_{\rho\sigma} = 0$$

Donc, si $H_\rho \geqslant mc^2$, $H_\sigma \geqslant mc^2$, ou si $H_\rho \leqslant mc^2$, $H_\sigma \leqslant mc^2$, il s'ensuit que $(\xi_k)_{\rho\sigma}$ doit être nul, l'opérateur ξ_k est donc impair. Le même raisonnement s'applique aux η_k.

On prouve de la même façon qu'il suffit, pour qu'un opérateur soit pair (ou impair) que son deuxième, troisième, ou en général sur n-ème commutateur (ou anticommutateur) avec $\mathcal{H}$ soit nul. Il s'en suit que

la décomposition des coordonnées à laquelle nous avons été amenés plus haut :

$$x_k = \widehat{x}_k + \xi_k,$$

est précisément la décomposition en opérateurs pair et impair. En effet, nous avons prouvé que ξ_k est impair. $\widehat{x}_k$ de son côté est pair parce qu'il est fonction linéaire du temps en l'absence de champ. Donc son deuxième commutateur avec $\mathcal{H}$ s'évanouit, donc $\widehat{x}_k$ est pair. De même la décomposition

$$x_k = c\mathcal{H}^{-1}p_k + (x_k - c\mathcal{H}^{-1}p_k) = c\mathcal{H}^{-1}p_k + r_k,$$

représente une décomposition en pair et impair.

Nous ajouterons quelques propositions presque évidentes :

1) Un opérateur quelconque est décomposable d'une seule façon en parties paire et impaire.

2) Une fonction quelconque est décomposable d'une seule façon en parties positive et négative.

3) Une fonction positive quelconque est orthogonale à une fonction négative quelconque.

4) La moyenne d'un opérateur pair prise pour un état représenté par une fonction « pure » (c'est-à-dire positive ou négative), est nulle.

5) Les fonctions propres d'un opérateur pair peuvent toujours être choisies « pures » (Mais il se peut que des fonctions propres positives et négatives appartiennent à une seule et même valeur propre).

6) Un opérateur impair n'a jamais de fonctions propres pures sauf peut-être pour la valeur propre zéro.

7) Un produit de puissances d'un certain nombre d'opérateurs « purs » est pair ou impair suivant que le nombre d'opérateurs *impairs* qui y entrent (ou la somme de leurs exposants) est pair ou impair.

La proposition 4) démontre ce qui a été dit à la fin du dernier paragraphe sur le comportement du centre de gravité dans le cas où la fonction ψ est purement positive (ou négative).

IV. — **Cas d'un champ extérieur**

Nous utiliserons les résultats que nous avons obtenus pour répondre à la question suivante qui est d'un intérêt fondamental : l'état initial de l'électron étant donné par une fonction ψ *positive*, cette fonction restera-t-elle purement positive au cours du temps ?

En l'absence de champ, la réponse est affirmative. La variation de ψ est régie par l'équation de propagation

$$\varkappa \frac{\partial \psi}{\partial t} = - \mathcal{H}\psi$$

ou

$$d\psi = - \frac{\mathcal{H}}{\varkappa} \psi dt$$

L'opérateur $\mathcal{H}$ est pair par définition. Donc l'accroissement de ψ est positif à chaque instant et ψ reste positif.

Envisageons maintenant le cas d'un champ extérieur quelconque donné par les potentiels A_1, A_2, A_3, et V. La théorie de DIRAC tient compte de l'action d'un champ extérieur sur l'électron par les modifications suivantes dans l'opérateur $\mathcal{H}$:

1) Il faut y ajouter l'énergie potentielle eV.

2) Il faut remplacer p_k par $p_k + \frac{e}{c} A_k$.

Cette manière de tenir compte de l'existence d'un champ extérieur n'est nullement satisfaisante ; toute la conception de « champ extérieur » ne l'est pas. Ce n'est que quelque chose de provisoire, un « ersatz » pour une théorie que nous ne possédons pas encore. En vérité, les « champs extérieurs » proviennent d'autres électrons et de protons et l'on devrait donc traiter le problème comme un « problème de plusieurs corps » et cela en toute généralité, c'est-à-dire dans le cas où les corps sont si éloignés les uns des autres que la simple loi de COULOMB ne suffit plus à décrire leurs actions mutuelles. Nous ne savons pas encore attaquer ce problème ; du moins les méthodes qui ont été proposées jusqu'à présent, l'hyperquantification ou la quantification répétée, la quantification des champs, sont si compliquées et offrent de si grandes difficultés que je ne peux en parler aujourd'hui.

Employons donc la méthode de DIRAC qui, vous le savez, a conduit à de si bons résultats, par exemple en ce qui concerne structure fine de l'hydrogène, etc.... Soit $\mathcal{K}$ l'opérateur d'énergie dans un champ quelconque, c'est-à-dire l'opérateur $\mathcal{H}$ ayant subi les deux modifications indiquées plus haut. (La lettre $\mathcal{K}$ a été choisie pour rappeler le nom de KÉPLER) :

$$\mathcal{K} = c\,[\alpha_1 (p_1 + \frac{e}{c} A_1) + \cdots + \cdots + \alpha_4\, mc] + eV,$$

A_1, A_2, A_3, V étant des fonctions données de x_1, x_2, x_3, t. L'équation d'onde pour ψ est alors

$$d\psi = -\frac{\mathcal{K}}{\varkappa}\psi dt. \tag{4}$$

Préservera-t-elle le caractère positif de ψ, si ψ est positif au commencement ? Cela dépend de la parité de l'opérateur $\mathcal{K}$. Nous verrons bientôt que $\mathcal{K}$, quoique *grossièrement* pair, contient une petite partie impaire.

Avant d'aller plus loin, insistons bien sur le fait que dans les définitions de « positif » et « négatif », « pair » et « impair », c'est toujours l'opérateur $\mathcal{H}$ du mouvement *sans* champ de forces qui intervient, et non pas $\mathcal{K}$. En effet, ce qui nous intéresse c'est la probabilité d'apparition au cours du temps d'une énergie *interne* négative, énergie qui correspond classiquement à l'expression bien connue

$$\frac{mc^2}{\sqrt{1-\beta^2}},$$

et non pas la probabilité d'une énergie *totale* négative. En effet, pour éviter toute contradiction avec l'expérience, il est nécessaire d'exclure les états d'énergie *interne* négative : voilà le but que nous avons en vue.

Il faut donc étudier l'opérateur $\mathcal{K}$ et trouver sa décomposition en parties paire et impaire. Ceci est rendu facile par la proposition suivante, que nous démontrerons tout à l'heure. Étant donnée une décomposition des opérateurs $\mathcal{A}$, $\mathcal{A}'$, $\mathcal{A}''$, $\cdots$

$$\begin{aligned}\mathcal{A} &= p + i\\ \mathcal{A}' &= p' + i'\\ \mathcal{A}'' &= p'' + i''\\ &\cdots\cdots\cdots\cdots,\end{aligned}$$

la décomposition d'un opérateur $\mathcal{B}$, *fonction* de $\mathcal{A}$, $\mathcal{A}'$, $\mathcal{A}''$, $\cdots$ est donnée par

$$\begin{aligned}\mathcal{B} &= f(\mathcal{A}, \mathcal{A}', \mathcal{A}''\cdots)\\ &= \frac{1}{2}[f(p+i, p'+i', \cdots) + f(p-i, p'-i', \cdots)] +\\ &+ \frac{1}{2}[f(p+i, p'+i', \cdots) - f(p-i, p'-i', \cdots)]\end{aligned}$$

Cette proposition se démontre aisément. On commence par prendre pour fonction f le produit $\mathcal{A}\mathcal{A}'$ et on déduit que le théorème est valable pour un produit de puissances quelconque, ensuite pour une somme de produits de puissances, et enfin pour une fonction analytique quelconque.

Notre opérateur $\mathcal{K}$ est donné comme fonction des α_k, des p_k et des x_k. Nous avons obtenu au paragraphe précédent la décomposition des α_k et des x_k. Or, les p_k sont pairs (ils commutent avec $\mathcal{H}$), donc la décomposition de $\mathcal{K}$ peut s'effectuer au moyen du théorème précédent.

Examinons de plus près le cas de l'atome d'hydrogène. Nous avons

$$A_k = 0, \qquad eV = -\frac{e^2}{r}, \qquad r = \sqrt{x_1^2 + x_2^2 + x_3^2},$$

$$\mathcal{K} = \mathcal{H} - \frac{e^2}{r} = \mathfrak{L} + \mathfrak{J}$$

$$\mathfrak{L} = \mathcal{H} - \frac{e^2}{2}\left(r^{-1} + \frac{1}{\sqrt{(x_1 - 2\xi_1)^2 + \cdots}}\right)$$

$$\mathfrak{J} = \quad - \frac{e^2}{2}\left(r^{-1} - \frac{1}{\sqrt{(x_1 - 2\xi_1)^2 + \cdots}}\right)$$

$$x_k = \widehat{x}_k + \xi_k$$

$$\widehat{x}_k - \xi_k = x_k - 2\xi_k$$

Les opérateurs ξ_k sont un peu compliqués, il est vrai, puisqu'ils contiennent $\mathcal{H}^{-1}$ qui n'est pas un opérateur différentiel. Néanmoins tout est bien défini. L'extraction de la racine carré ne présente pas de difficultés, puisque la quantité sous le radical est un opérateur positif, c'est-à-dire ayant des valeurs propres positives, et il est à peu près évident qu'il faut prendre la racine carrée de manière que toutes ses valeurs propres soient positives (L'ambiguité qui se présente est en effet la même que celle qui apparaît dans la définition de l'opérateur r lui-même).

Nous avons posé la question suivante : L'équation (4) conserve-t-elle le caractère positif d'une fonction positive ? Évidemment non : $\mathcal{K}$ n'est pas pair. Sa partie impaire $\mathfrak{J}$, que nous avons isolée, va provoquer au cours du temps l'apparition de parties négatives dans la fonction d'ondes. Quoique $\mathfrak{J}$ soit assez petit comparé à $\mathfrak{L}$ (ainsi que nous le verrons tout à l'heure), la probabilité d'apparition d'un électron de masse négative devient *beaucoup* trop grande pour pouvoir être acceptée.

On pourrait s'en débarrasser en supprimant $\mathfrak{J}$ ou, en d'autres termes, en remplaçant l'opérateur $\mathfrak{K}$ par $\mathfrak{L}$, sa partie paire. Cela est-il permis ? L'opérateur $\mathfrak{J}$ est-il suffisamment *petit* comparé à $\mathfrak{L}$ ou à $\mathfrak{K}$ pour que la suppression de $\mathfrak{J}$ ne change pas sensiblement les valeurs propres, c'est-à-dire la valeur des termes de la structure fine de l'hydrogène ?

La petitesse de $\mathfrak{J}$ résulte de la petitesse des ξ_k, dont nous avons estimé l'ordre de grandeur à :

$$\{ \xi_k \} = \frac{h}{4\pi mc}.$$

(Nous utilisons le signe $\{ \}$ pour indiquer l'« ordre de grandeur »). Le rapport des ordres de grandeur de $\mathfrak{J}$ et de $\frac{e^2}{2r}$ (qui est précisément égal à la valeur du terme d'énergie) sera environ

$$2 \frac{\{ \xi_k \}}{\{ r \}}$$

Pour $\{ r \}$, on pourra prendre le « rayon de l'atome d'hydrogène » bien connu :

$$\{ r \} = \frac{h^2}{4\pi^2 me^2}.$$

Il en résulte

$$\frac{\{ \mathfrak{J} \}}{\{ \text{terme} \}} = \frac{2\pi e^2}{hc} = \alpha,$$

où α désigne, comme d'habitude, la constante de structure fine.

Vous savez que le rapport de la *structure fine* au terme est beaucoup plus petit, environ α^2. Par conséquent il semble que $\mathfrak{J}$ soit beaucoup trop grand pour que sa suppression soit permise. Il semble que la suppression de la partie impaire de $\mathfrak{J}$ changerait les valeurs propres de beaucoup plus que la séparation en structure fine, et par conséquent détruirait tout à fait l'accord avec l'expérience.

Mais il n'en est pas ainsi en réalité. Pour le voir, il est un peu plus commode de ne pas parler de la *suppression* de $\mathfrak{J}$, mais de partir de l'opérateur $\mathfrak{L}$ et de se demander quel changement subiront les valeurs propres de $\mathfrak{L}$ si on lui *ajoute* l'opérateur $\mathfrak{J}$. Vous savez que la per-

turbation d'une valeur propre se calcule *au premier ordre près* au moyen des intégrales (« matrices de perturbation »)

$$\int \psi_i^* \mathfrak{J} \psi_k \, dx,$$

où les ψ_k sont les fonctions propres du problème non perturbé, qui appartiennent à la valeur propre dont on veut calculer la perturbation. Mais, l'opérateur $\mathfrak{J}$ étant *impair, tous ces éléments de matrice disparaissent*, si les fonctions propres ψ_k qui appartiennent à cette valeur propre sont ou toutes positives, ou toutes négatives. Dans ce cas donc la perturbation des valeurs propres d'un opérateur *pair* par un petit opérateur impair est tout au plus *du second ordre*.

Avant de rechercher quel est l'ordre de grandeur qui entre en jeu dans notre cas, nous devons nous demander s'il est exact que les fonctions propres de $\mathfrak{L}$ appartenant à la même valeur propre sont, toutes positives, ou toutes négatives. Soit

$$\chi = \chi_+ + \chi_-$$

la décomposition d'une telle fonction propre avec la valeur propre P'. On a alors

$$\mathfrak{L}(\chi_+ + \chi_-) = \mathrm{P}' \cdot (\chi_+ + \chi_-)$$

et par décomposition en parties positive et négative :

$$\mathfrak{L}\chi_+ = \mathrm{P}'\chi_+ \qquad \mathfrak{L}\chi_- = \mathrm{P}'\chi_-.$$

Donc χ_+, χ_- sont elles-mêmes fonctions propres de $\mathfrak{L}$. On peut toujours prendre les fonctions propres de $\mathfrak{L}$ *pures*. Mais il se peut qu'à la même valeur propre appartiennent un certain nombre de fonctions positives et aussi quelques fonctions négatives.

Cette possibilité existe en général pour un opérateur pair, mais non pas pour $\mathfrak{L}$. Car il n'en est pas ainsi avec l'opérateur $\mathcal{H}$ par définition, et $\mathfrak{L}$ est suffisamment voisin de $\mathcal{H}$ pour qu'on puisse conclure que la même chose se produit avec $\mathfrak{L}$. Si l'on regarde $\mathfrak{L}$ comme le résultat d'une perturbation de $\mathcal{H}$, cette perturbation est d'un ordre de grandeur de *beaucoup* inférieur à $2mc^2$, — valeur qui sépare les deux catégories de valeurs propres de $\mathcal{H}$. Celles-ci ne peuvent donc pas être confondues par une telle perturbation.

Revenons à la recherche de l'ordre de grandeur de la perturbation produite par $\mathfrak{J}$. Nous avons trouvé que

$$\frac{\{\mathfrak{J}\}}{\{\text{terme}\}} = \alpha$$

et que la perturbation est du *second* ordre. Est-ce que cela veut dire qu'elle est de l'ordre α^2 par rapport au terme ? Heureusement non, puisqu'il faut la comparer à l'opérateur complet $\mathfrak{K}$ qui est de l'ordre de mc^2. Vous trouverez facilement que le *terme* n'est qu'une fraction α^2 de mc^2. Donc, comparé à $\mathfrak{K}$, le premier ordre est α^3, le second α^6, et ce second ordre comparé au terme n'est que α^4. Ceci est de beaucoup inférieur à la séparation en structure fine, et même de beaucoup inférieur à la largeur naturelle des raies, causée par l'amortissement du rayonnement et qui est de l'ordre de α^3.

On peut démontrer de même que dans l'effet STARK et dans l'effet ZEEMAN la perturbation est tout à fait insignifiante. Je n'insisterai pas là-dessus : il ne se présente pas de point de vue nouveau.

V. — **Le problème de la relativité**

Par un procédé bien déterminé, à savoir en supprimant la partie impaire de l'opérateur de DIRAC nous avons réussi à en déduire une équation d'ondes qui conserve le caractère positif de la fonction ψ et qui néanmoins fournit pour la structure fine de l'hydrogène, l'effet STARK et l'effet ZEEMAN des résultats en accord avec l'expérience.

Toutefois on ne doit pas estimer ce succès trop haut. En premier lieu on doit se rappeler que l'introduction de la notion de « champ extérieur » dans l'équation de DIRAC n'est elle-même au fond qu'un artifice qui nous permet d'éviter la résolution du problème des n corps, problème que nous ne savons pas encore traiter d'une manière qui tienne compte de la relativité et des potentiels *retardés*. En second lieu, il faut remarquer que l'*existence* d'un opérateur pair avec les valeurs propres correctes ne signifie rien : il est même évident qu'il existe une infinité de semblables opérateurs. Car, étant donnée une série quelconque de nombres réels

$$\varepsilon_1, \varepsilon_2, \varepsilon_3 \cdots$$

et un système complet de fonctions orthogonales quelconques

$$\omega_1, \omega_2, \omega_3, \cdots$$

on peut toujours — au moins formellement — écrire une fonction de deux groupes de variables

$$K(x,\xi) = \sum_k \varepsilon_k \, \omega^*_k(x)\omega_k(\xi),$$

qui, *regardée comme noyau intégral*, est un opérateur linéaire hermitique ayant les ε_k comme valeurs propres et les ω_k comme fonctions propres. Or, on peut choisir les ω de façon qu'ils soient tous « purs » et dans ce cas l'opérateur sera pair. Ce qu'il y a de satisfaisant dans notre procédé consiste uniquement dans le fait que nous avons réussi à former un tel opérateur par un procédé relativement *simple*.

Il y a un certain nombre de questions particulières qui devront être examinées. Par exemple, en introduisant des A_k qui correspondent à une onde lumineuse et en étudiant l'onde diffusée, il sera intéressant de voir si, pour le cas limite des hautes fréquences, on retrouve la formule classique de RAYLEIGH. I. WALLER a démontré qu'avec l'équation de DIRAC on la retrouve en effet, mais uniquement grâce à la possibilité de transitions entre les états d'énergie positive et négative. Dans le cas présent, *si l'on retrouve* la même formule (ce que j'espère et ce que je crois, — mais je ne l'ai pas encore démontré), le mécanisme mathématique doit être différent du précédent, car ces transitions sont strictement interdites après la « purification » que nous avons entreprise.

Mais il y a d'autres questions plus inquiétantes. On a regardé comme une des propriétés les plus importantes de l'équation de DIRAC, ou même peut-être *la plus importante* son invariance par rapport à la transformation relativiste de LORENTZ. Si on introduit pour les

$$x_1, x_2, x_3, t \rightarrow x'_1, x'_2, x'_3, t'$$

la transformation bien connue de LORENTZ et si en même temps on fait une certaine transformation linéaire entre les quatre composantes de ψ :

$$\psi_1, \psi_2, \psi_3, \psi_4 \rightarrow \psi'_1, \psi'_2, \psi'_3, \psi'_4,$$

l'équation à laquelle satisfont les ψ'_k est exactement celle à laquelle satisfaisaient les ψ_k.

Est-ce que cette invariance se conserve après la suppression de $\mathfrak{J}$?

La réponse est qu'elle ne se conserve pas exactement. Naturellement dans les cas qui nous intéressent le défaut d'invariance ne peut pas être grand, puisque le changement d'opérateur n'est pas grand. Mais il n'est pas possible de restituer à l'équation sa forme exacte.

Au premier abord, on est enclin à croire que ce résultat désespérant rend inutiles tous les résultats précédents. Mais je ne crois pas qu'il en soit ainsi. En effet, je suis convaincu qu'en mécanique quantique la question de l'invariance relativiste est beaucoup plus compliquée qu'on ne l'imaginait jusqu'à présent et j'ai l'intention dans un prochain paragraphe de reprendre le problème des rapports entre la mécanique quantique et la relativité restreinte d'un point de vue un peu plus général.

Pour le moment, on peut faire les remarques suivantes pour excuser en quelque sorte la non-invariance de notre nouvelle équation.

En premier lieu, on pourrait se demander quelle forme devrait avoir notre équation pour être invariante. A quelle sorte d'équation doit satisfaire une fonction ψ qui est « positive » pour toutes les valeurs du paramètre t et qui reste positive pour une transformation de LORENTZ quelconque ?

Cette question est susceptible d'une réponse précise, et même d'une réponse *unique*, assez frappante. Je vais seulement indiquer le résultat (1) : Si on impose à ψ les conditions d'être « positive » à tout instant et invariante pour une transformation de LORENTZ quelconque, cela implique inévitablement que ψ doit être une solution de

$$\varkappa \frac{\partial \psi}{\partial t} = - \mathcal{H}\psi$$

c'est-à-dire du problème en l'*absence* de champ. Donc la recherche d'une équation d'onde qui nous débarrasse des valeurs propres négatives et qui soit invariante, est inutile. L'équation précédente est la *seule* qui jouisse de ces propriétés et elle s'applique évidemment seulement au cas d'absence de champ.

Devra-t-on donc renoncer au postulat de la relativité ? Évidemment non. Mais la transformation de LORENTZ semble être une chose beaucoup plus compliquée en mécanique quantique qu'elle ne l'était en

(1) Voir Note I.

mécanique ordinaire. On s'aperçoit que dans la nouvelle théorie les opérateurs

$$x_k \qquad x_k - 2\xi_k$$

ou bien

$$\widehat{x}_k + \xi_k \qquad \widehat{x}_k - \xi_k$$

jouent un rôle tout à fait symétrique. Et cette symétrie est réelle. Je dois à M. v. NEUMANN la remarque qu'on peut indiquer une certaine transformation canonique qui change à la fois *tous* les opérateurs $p + i$ en $p - i$, c'est-à-dire qui change simplement le signe de tous les opérateurs impairs. Par conséquent toutes les relations qui existent entre les opérateurs $p + i$ (relations de commutation par exemple) existent également entre les $p - i$. Cela rend à peu près inévitable la conclusion, que dans la nouvelle théorie, ce ne sont pas les x_k, mais plutôt les opérateurs pairs $\widehat{x}_k$ qui correspondent aux coordonnées de l'électron. Par conséquent la transformation de LORENTZ devra être effectuée non pas sur les $\widehat{x}_k$ mais sur les x_k.

Mais alors la transformation de LORENTZ devient quelque chose de bien plus compliqué et dificile *non pas à formuler, mais à concevoir.* Lorsqu'on effectue des transformations sur les x_k, t, il est permis de fermer les yeux sur le fait que les x_k ne sont pas des nombres ordinaires, parce qu'ils commutent entre eux, de sorte que l'on peut choisir un système de matrices où ils sont diagonaux tous à la fois ; alors la transformation s'effectue sur les *valeurs propres* des x_k, c'est-à-dire sur les valeurs des coordonnées au sens élémentaire. C'est ce qu'on fait toujours tacitement. Il n'est pas possible d'appliquer cette manière de voir aux $\widehat{x}_k$, puisqu'ils ne commutent pas entre eux, de sorte qu'il n'existe pas de représentation dans laquelle tous les trois soient diagonaux. Il est difficile d'entrevoir comment la transformation des $\widehat{x}_k$ et de t devra être formulée, surtout dans une théorie qui traite le temps comme une variable ordinaire et non comme un opérateur.

Le fait que les vrais opérateurs-coordonnées de l'électron ne commutent pas entre eux me semble assez intéressant en lui-même, parce que cela veut dire qu'il n'est pas possible en général de mesurer exactement deux coordonnées ou toutes les trois coordonnées à la fois. Il existe des « relations d'incertitude » entre elles. L'incertitude corrélative de deux observables est principalement déterminée, vous le savez,

par l'ordre de grandeur de leur commutateur. Il n'est pas difficile de calculer les commutateurs des $\widehat{x_k}$ et cela est tout à fait intéressant, parce qu'ils sont intimement liés aux opérateurs du « spin » (1).

Le problème qui se pose me semble être de formuler la transformation de LORENTZ comme une équation opérationnelle entre les opérateurs non commutatifs $\widehat{x_k}$.

VI. — **Relativité et mécanique quantique**

Les difficultés que nous avons rencontrées en tâchant de tenir compte du point de vue relativiste dans la mécanique quantique me semblent d'autant plus intéressantes qu'elles sont tout à fait *imprévues*. Vous savez que la mécanique nouvelle, sous la forme de mécanique *ondulatoire* sous laquelle elle est appliquée presque universellement aujourd'hui, doit son origine aux célèbres recherches de M. L. DE BROGLIE, à son ingénieuse conception des ondes électroniques qui devaient accompagner le mouvement de l'électron. Les recerches de M. L. DE BROGLIE s'appuyaient sur la théorie de la relativité restreinte ; elles étaient pour ainsi dire imprégnées de relativité. Lorsqu'on les prit comme point de départ pour en tirer l'équation d'ondes et les problèmes de valeurs propres, on éprouva un peu de honte d'être obligé de *supprimer* d'abord le point de vue relativiste et on espéra que ce ne serait qu'une situation provisoire et de courte durée, et qu'il ne serait pas trop difficile d'introduire la relativité à nouveau dans les équations. Mais au lieu de diminuer, il semble bien que cette difficulté a crû d'une année à l'autre jusqu'à prendre aujourd'hui des proportions effrayantes.

Cependant la forme sous laquelle nous l'avons rencontré dans les discussions précédentes est compliquée et très spéciale. On pourrait objecter que l'équation de DIRAC dont nous sommes partis ne représente probablement pas la seule forme possible de l'équation relativiste de l'électron et que notre méthode de suppression de la partie impaire de l'opérateur hamiltonien n'est peut-être pas la seule méthode de modifier l'équation de DIRAC de façon à se débarrasser des énergies négatives. On pourrait supposer qu'il est nécessaire d'abandonner tout à fait cet ordre d'idées qui, malgré ses succès, est tout simple-

(1) Voir Note II.

ment condamné par les difficultés relativistes auxquelles il donne naissance.

C'est pourquoi je voudrais vous demander la permission de vous exposer quelques idées plus générales sur les rapports entre les conceptions de la théorie de la relativité et de la mécanique quantique. Je crois que d'un point de vue tout à fait général, en n'utilisant que des théorèmes fondamentaux de ces deux théories, on peut reconnaître que la difficulté de les concilier ne tient pas à la forme spéciale des équations adoptées, mais à la nature essentiellement différente des conceptions fondamentales des deux théories.

Depuis la grande découverte d'EINSTEIN, on s'était habitué à soumettre une théorie physique quelconque au postulat de la relativité restreinte, c'est-à-dire que l'on exigeait son invariance par rapport aux transformations de LORENTZ. On n'y rencontrait jamais de difficultés sérieuses. Au contraire, on était invariablement conduit à la généralisation de la théorie en question pour les corps en mouvement. Mais en essayant de soumettre la mécanique quantique au même postulat, on se heurte à des difficultés. Pourquoi ? Ce n'est pas si surprenant que cela en a l'air. Pour savoir ce que veut dire une transformation de LORENTZ dans un cas particulier, il faut avoir introduit deux systèmes de coordonnées de LORENTZ. Cette introduction est basée sur l'idée qu'il est possible en principe de mesurer les coordonnées d'un point matériel et le temps avec une exactitude illimitée et aussi souvent que l'on voudra, et d'en tirer des conclusions relatives à sa vitesse. Mais cela n'est pas permis en mécanique quantique. Donc la mécanique quantique n'est pas obligée à se soumettre au postulat de relativité avec une exactitude arbitrairement grande. Elle a le droit à son tour d'exiger d'examiner à quelle erreur près on peut définir un système de coordonnées. Ce n'est qu'à cette erreur près qu'elle est obligée de satisfaire au postulat relativiste.

Examinons-le de plus près. Parmi les opérations nécessaires pour établir un système de LORENTZ, se trouve l'opération qui consiste à régler les montres aux différents points du système au moyen de signaux lumineux ; c'est même peut-être la plus importante de toutes. Or, il est facile de voir que cette opération ne peut présenter qu'une exactitude limitée si les montres ne sont pas *infiniment lourdes*. La raison en est la suivante.

Imaginez un système de masse m (que j'appellerai « montre ») et deux

« événements » qui se passent dans ce système : par exemple deux positions de l'aiguille. Vous voulez mesurer le temps qui s'écoule entre ces deux évènements (« réglage de la montre »). Pour vous avertir du premier, la montre doit émettre un signal optique qui doit être d'autant plus court que vous exigez de savoir plus exactement l'instant où l'événement a eu lieu. Exigeons d'abord une précision de τ secondes.

La longueur spatiale du signal ne doit pas dépasser $c\tau$ centimètres. Mais alors on sait d'après les principes fondamentaux de l'optique que la lumière ne peut être monochromatique ; l'intervalle de fréquences qui entre en jeu sera au moins

$$\Delta\nu = \frac{1}{4\pi\tau}.$$

Maintenant, quand un système émet de la lumière de fréquence ν, il faut admettre qu'il recule et que la quantité de mouvement correspondante est

$$\frac{h\nu}{c}.$$

(EINSTEIN, dans sa célèbre démonstration de la formule de PLANCK en 1917 a démontré que ce résultat est une des conséquences les plus immédiates de la théorie des quanta ; elle est rendue presque inévitable par les faits expérimentaux connus sur le rayonnement noir. Si le système n'éprouvait pas ce recul, l'agitation thermique d'une molécule en équilibre avec le rayonnement noir ne serait pas la même que dans un gaz, donc un gaz ne serait pas en équilibre avec le rayonnement noir de même température.)

Le *non-monochromatisme* de la lumière émise comme signal amènera donc une incertitude sur le recul :

$$\Delta p = \frac{h}{4\pi c\tau}. \tag{1}$$

(Il est commode d'admettre que la partie principale du recul ait été compensée par un signal émis dans une direction opposée de manière qu'il ne subsiste que l'incertitude sur la quantité de mouvement indiquée par la formule (1).) Soit β l'incertitude correspondante sur la vitesse [1]. Nous avons

$$\frac{mc\beta}{\sqrt{1-\beta^2}} = \Delta p = \frac{h}{4\pi c\tau} \tag{2}$$

(1) Mesurée en unités c.

d'où l'on peut tirer β. Mais cette incertitude sur la *vitesse* amène un *deuxième* manque de précision dans la mesure du *temps* qui se passe entre les deux évènements. Car la montre est ralentie par le mouvement et il va de soi que ce qui nous intéresse, c'est le temps qui se *serait* écoulé si la montre était restée au repos ; en d'autres termes le temps dans le « système au repos » de la montre, que nous ne connaissons qu'avec une erreur β.

Donc le temps t que nous mesurons est trop grand, et doit être corrigé en multipliant t par un facteur compris entre $\sqrt{1-\beta^2}$ et 1. Donc la seconde erreur est

$$t(1-\sqrt{1-\beta^2}). \tag{3}$$

En tirant β de (2) et en substituant dans (3), on trouve

$$t\left(1-\frac{\tau}{\sqrt{\tau^2+\tau_0^2}}\right) \qquad \text{avec} \qquad \tau_0=\frac{h}{4\pi mc^2}.$$

L'erreur *totale* est donc

$$\Delta t=\tau+t\left(1-\frac{\tau}{\sqrt{\tau^2+\tau_0^2}}\right).$$

Elle est minima (t étant au moins égal à τ_0) pour :

$$\tau=\tau_0\sqrt{\left(\frac{t}{\tau_0}\right)^{\frac{2}{3}}-1}.$$

Elle a pour valeur

$$\Delta t=t\left\{1-\left[1-\left(\frac{\tau_0}{t}\right)^{\frac{2}{3}}\right]^{\frac{3}{2}}\right\}.$$

On démontre aisément que l'on a toujours $\Delta t \geqslant \tau_0$. Quand t devient égal à τ_0, on a $\tau=0$, $\Delta t=t$, l'erreur est de 100 %. Pour $t<\tau_0$, la valeur la plus favorable est $\tau=0$, l'erreur reste de 100 %.

Je crois que l'on peut en tirer la conclusion qu'une montre de masse m n'est capable d'être réglée qu'avec une précision de τ_0. Et même peut-être que les événements qui se passent dans un système de masse m ne peuvent être localisés dans le temps avec une précision supérieure à τ_0 (ou plutôt une erreur inférieure à τ_0) par un observateur extérieur. Toutefois c'est là une façon de parler un peu abrégée. Du point de vue de l'observateur extérieur la localisation peut s'effectuer avec toute la précision désirée. Mais, quant à ce qui se passe à l'intérieur du sys-

tème (ou en d'autres termes : quant au temps propre du système), cela n'aura aucune signification au delà de la précision τ_0. D'une manière toute analogue, on peut montrer que la mesure de la distance spatiale de deux points appartenant à un système matériel de masse totale m ne pourra être effectuée avec une erreur inférieure à

$$\frac{h}{4\pi mc}$$

environ. Car d'après la « loi d'incertitude » de HEISENBERG, si la position de l'un des deux points est observée avec une précision λ, cela implique une erreur sur la quantité de mouvement du système qui est d'au moins

$$\Delta p = \frac{h}{4\pi\lambda} = \frac{mc\beta}{\sqrt{1 - \beta^2}},$$

β étant l'incertitude sur la vitesse du système. Cela implique une incertitude relative à la contraction de LORENTZ, donc une incertitude additionnelle de

$$\lambda\left(1 - \frac{1}{\sqrt{1 - \beta^2}}\right),$$

sur la longueur l observée, si l'on veut en déduire la longueur vraie ou longueur au repos de la distance que l'on veut mesurer. On voit bien qu'en diminuant la première de ces deux erreurs, on augmente la dernière, et on peut entrevoir que l'erreur inévitable est de l'ordre de grandeur indiqué. Remarquez que si m est égal à la masse de l'électron, c'est précisément l'ordre de grandeur des opérateurs ξ_k qui représentent le changement des opérateurs-coordonnées dans notre équation « purifiée », comparée à l'équation de DIRAC. Donc le défaut d'invariance de notre équation au sens conventionnel se produit justement dans un domaine où la détermination des coordonnées elle-même, et par conséquent la conception de transformation de LORENTZ, devient problématique.

J'ai dit : « si m est égal à la masse électronique ». Naturellement si on admettait l'utilisation de montres et de règles de mesure aussi lourdes que l'on veut, alors le temps et le lieu de l'électron seront observables avec la précision arbitraire désirée. Mais je ne crois pas qu'il en soit ainsi, je ne crois pas que cela soit admissible (1).

(1) Voir Note III.

Il apparaît donc qu'au point de vue de la mécanique quantique, la théorie de la relativité se range au même niveau que la mécanique classique en ce sens qu'elle ne représente qu'une approximation relative au domaine macroscopique. On ne devra pas admettre tout simplement les formules de la relativité (par exemple les formules de LORENTZ) et les supposer valables sans changement dans le domaine intra-atomique. Elles devront être soumises à des modifications qui seront probablement analogues à celles qu'a subies la mécanique ordinaire pour se transformer en mécanique quantique. Il faudra « quantifier » la transformation de LORENTZ.

Quantification de la formule de LORENTZ : qu'est-ce que cela veut dire, qu'est-ce que cela *peut* signifier ? Je ne crois pas que cela puisse avoir une autre signification que celle-ci : il faudra vraisemblablement regarder les formules bien connues comme établies entre *les opérateurs* plutôt qu'entre *les nombres* (nombres-*c*) qui appartiennent aux coordonnées. En effet, en théorie classique, ces formules servent à calculer les valeurs des coordonnées et du temps pour un point (ou plutôt pour un événement ponctuel) dans *un* système de référence, si l'on *connaît* leurs valeurs dans un autre système. Mais en mécanique quantique, on ne connaît pas en général la *valeur* d'une quantité observable, mais seulement la probabilité pour qu'elle ait telle ou telle valeur. Ce que l'on demandera, c'est de pouvoir calculer les probabilités dans un système de référence, lorsqu'elles sont connues dans un autre système. Puisque la probabilité en mécanique quantique se calcule au moyen de l'*opérateur* qui appartient à l'observable en question, il est très naturel de penser que la transformation de LORENTZ prendra la forme d'équations entre les opérateurs qui représentent les *xyzt* dans les deux systèmes de référence.

Nous sommes donc amenés, d'un point de vue beaucoup plus général, au même problème que celui que nous avons rencontré en partant des équations de DIRAC. Là, ce qui nous donnait de la difficulté, c'était la non-commutabilité des coordonnées, qui devient très probable d'après la modification de la théorie de DIRAC que j'avais proposée. Mais même si on n'y croit pas ou si l'on en fait abstraction, il subsiste une autre difficulté pour la « quantification de la tranformation de LORENTZ ». Dans cette transformation, vous le savez bien, le *temps* et les coordonnées d'espace entrent d'une façon tout à fait symétrique ; ce sont des quantités de même espèce, constituant les

quatre composantes d'un vecteur d'Univers. On peut même dire que les équations de LORENTZ n'expriment rien d'autre que ce fait fondamental. Mais en mécanique quantique, au moins dans l'interprétation courante, il n'en est pas ainsi : le temps est toute autre chose que les coordonnées. Dans l'interprétation courante, tous les énoncés de la mécanique quantique se réduisent à la forme suivante : la probabilité pour qu'*à l'instant t* l'observable q ait telle ou telle valeur est telle ou telle. Jamais il ne s'agit de la probabilité pour que le *temps* ait telle ou telle valeur. Le temps n'est pas traité en observable, il n'y a pas d'opérateur qui servirait à trouver sa « statistique ». C'est un paramètre dont la valeur est supposée exactement connue : c'est en réalité le brave vieux temps de NEWTON et la mécanique quantique ne s'inquiète pas de l'existence de la brave vieille pendule dont elle aurait besoin pour connaître la valeur de ce paramètre t.

Je ne sais pas vous indiquer pour le moment comment on pourrait faire pour renoncer au dernier paramètre non statistique que la mécanique quantique a laissé survivre. Elle semble en avoir besoin puisque ses énoncés de probabilités ne sont certainement pas *invariables* et par conséquent doivent être exprimés en fonction de *quelque chose*. Peut-être pourrait-on entrevoir qu'ils seraient exprimés l'un en fonction de l'autre, mais ce n'est qu'une façon de parler assez vague.

Mais il me paraît tout à fait exempt de doute qu'on *devra* renoncer à cette notion beaucoup trop classique du temps, et non seulement à cause de la relativité. Cette notion de temps est un grave manque de conséquence dans la mécanique quantique (ou bien dans son interprétation courante), abstraction faite des postulats de relativité. Car effectivement la connaissance de la variable t est acquise de la même manière que celle de toute autre variable, en observant un certain système physique, à savoir une montre. t est donc une observable et doit être traité en observable ; le temps doit avoir en général une « statistique » et non une « valeur ». Le rôle exceptionnel du temps n'est donc pas justifié.

On pourrait objecter qu'en mécanique quantique il est toujours possible *qu'une* des observables (ou même la moitié des observables) ait des valeurs tout à fait déterminées (l'autre moitié étant alors tout à fait indéterminée). On pourrait dire que la montre idéale, dont nous avons besoin, est tout simplement un système où la variable « position de l'aiguille » a toujours une valeur exactement déterminée. Le rôle

exceptionnel du temps serait alors, sinon justifié, au moins permis.

Mais il n'en est pas ainsi. Je crois que l'on peut démontrer que l'état de « montre idéale » n'est pas possible pour un système réel. Je pourrais m'appuyer sur l'énoncé bien connu que le temps et l'énergie sont des variables canoniquement conjuguées. Mais je ne veux pas le faire, parce que dans cet énoncé (et dans cet énoncé seulement), on parle du temps *comme s'il* était traité en observable. Or, c'est là une hypocrisie : il ne l'est pas. Pour cela cet énoncé m'a toujours semblé un peu obscur. En outre, ce que nous voulons faire voir, c'est que le temps *doit* être traité en observable et non en paramètre. Ce serait donc une pétition de principe que de se servir de ce théorème, qui le suppose tacitement.

Pour examiner les propriétés d'une montre idéale *en marche*, souvenez-vous que d'après l'interprétation de la mécanique quantique, la probabilité d'une valeur quelconque d'une variable à l'instant t doit se calculer à partir de la fonction ψ de x et de t :

$$\psi(x, t).$$

En un mot, tout ce que l'on peut attendre d'une observation du système doit pouvoir se déduire de sa fonction d'ondes. Appliquons cela à notre montre idéale. Cela est parfaitement admissible, seulement il est nécessaire de supposer qu'il y en a deux : l'une pour y lire le temps, l'autre qui soit décrite par la fonction $\psi(x, t)$ et à laquelle s'appliquera notre raisonnement. Puis souvenez-vous qu'en développant la fonction ψ en série de FOURIER (ou en intégrale de FOURIER),

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-i\omega t} c(x, \omega) d\omega,$$

l'intégrale

$$\int dx \int_{\omega_0}^{\omega_1} | c(x, \omega) |^2 d\omega$$

représente la probabilité que l'on a de trouver pour l'énergie du système une valeur comprise entre

$$\frac{h\omega_0}{2\pi} \quad \text{et} \quad \frac{h\omega_1}{2\pi}.$$

Or, une montre idéale est un système qui doit satisfaire au postulat suivant : il y a des événements auxquels on doit s'attendre avec certitude à un instant donné et dont on est également certain qu'ils n'arri

veront à aucun autre instant. Par exemple, une certaine position de l'aiguille doit avoir la probabilité 1 à un certain instant et la probabilité zéro à tout autre instant, si la montre est absolument précise. Mais il vaut mieux commencer par supposer la précision limitée à τ secondes. Il s'ensuit qu'il doit y avoir une certaine qualité de la fonction ψ : celle qui confère à la probabilité de l'événement en question la valeur zéro — qualité que ψ doit posséder *toujours* sauf dans l'intervalle τ. Mais on voit aisément que cela est impossible si dans le développement de ψ ne figure, par exemple, qu'une seule valeur de ω, puisqu'aucune grandeur de probabilité n'est changée lorsqu'on multiplie ψ par un facteur de module 1, comme $e^{i\omega\tau}$. Et ce serait presque la même chose s'il ne figurait dans le développement de ψ qu'un intervalle

$$\Delta\omega \ll \frac{1}{\tau}.$$

Il faut tout au moins que l'intervalle soit

$$\Delta\omega \sim \frac{1}{\tau}.$$

Par conséquent si on voulait aller à la limite d'une montre absolument précise, $\Delta\omega$ et par suite l'intervalle d'incertitude sur l'énergie deviendrait infini.

Cela ne suffirait peut-être pas encore à rendre l'état de montre idéale physiquement impossible, car il se pourrait bien que malgré l'infinitude de $\Delta\omega$, la probabilité qu'une énergie dépasse une valeur ω_1 donnée, tende vers zéro quand ω_1 devient infini. Et voilà tout ce que l'on doit raisonnablement exiger. Mais jusqu'ici nous n'avons postulé la propriété de « montre idéale » que pour un seul instant. En l'appliquant à tous les instants, on peut en tirer la conclusion que la valeur de

$$\int |\varphi(x, \omega)|^2 dx$$

doit être indépendante de ω, de sorte que toutes les énergies deviennent également probables [1].

(1) Voir Note IV.

VII. — Une analogie entre la mécanique ondulatoire et quelques problèmes de probabilités en physique classique

Le sujet que je vais aborder maintenant n'est pas intimement lié aux questions dont il s'est agi dans les chapitres précédents. Tout d'abord vous aurez l'impression de choses qui ne sont pas du tout liées. Il s'agit d'un problème classique : problème de probabilités dans la théorie du mouvement brownien. Mais en fin de compte, il *ressortira* une analogie avec la mécanique ondulatoire, qui fut si frappante pour moi lorsque je l'eus trouvée, qu'il m'est difficile de la croire purement accidentelle.

A titre d'introduction, je voudrais vous citer une remarque que j'ai trouvée dans les « Gifford lectures » de A. S. EDDIGTON (Cambridge, 1928, p. 216 et sqq). EDDINGTON, en parlant de l'interprétation de la mécanique ondulatoire, fait dans une note au bas de la page la remarque suivante :

« The whole interpretation is very obscure, but it seems to depend on whether you are considering the probability *after you know what has happened* or the probability for the purposes of prediction. The $\psi\psi^*$ is obtained by introducing two symmetrical systems of ψ waves travelling in opposite directions in time ; one of these must presumably correspond to probable inference from what is known (or is stated) to have been the condition at a later time. »

Traduction en français :

« Toute l'interprétation est assez obscure, mais elle paraît être liée à la différence entre la probabilité *après que l'on sait ce qui s'est passé* ou la probabilité en vue de prédire ce qui se passera. La quantité $\psi\psi^*$ est obtenue en introduisant deux systèmes d'ondes ψ symétriques se propageant dans des directions de temps inverses ; l'un d'eux doit probablement correspondre à une conclusion probable, déduite de nos connaissances (ou nos hypothèses) concernant l'état du système à un instant postérieur ».

Vous savez sans doute que le mouvement brownien d'une particule qui n'est assujettie à aucune force en dehors des chocs moléculaires qui sont à l'origine du mouvement, est régi par l'équation de diffusion

$$\frac{\partial \omega}{\partial t} = D \cdot \Delta \omega. \tag{1}$$

D est la constante de diffusion et ω, ou plutôt

$$\omega(x, y, z, t)dx\,dy\,dz$$

est la probabilité de trouver la particule à l'instant t entre x, $x + dx$; y, $y + dy$; z, $z + dz$. On admet qu'à un instant donné, t_0, on connaissait les coordonnées de la particule, ou bien, plus généralement qu'à l'instant t_0 on connaissait la probabilité $\omega(x,y,z,t_0)$. La probabilité pour un instant t quelconque, postérieur à t_0 est la solution de (1) qui, pour t_0, prend la valeur donnée.

Vous savez aussi que SMOLUCHOWSKI et FOKKER, puis PLANCK ont généralisé ce problème pour des cas beaucoup plus compliqués, soit qu'il y ait des forces extérieures agissant sur la particule,soit qu'il ne s'agisse pas du tout d'une particule, mais d'un système tout à fait général assujetti à des influences d'une part régulières, comme par exemple la gravitation, d'autre part ayant un caractère irrégulier, sujettes au hasard, comme les chocs moléculaires ou, par exemple, les charges communiquées à un électromètre par un courant d'ionisation dû à la présence d'un corps radioactif.

Dans tous ces problèmes, c'est l'équation qu'on appelle « équation de FOKKER » qui régit la probabilité de trouver l'état du système compris entre des limites déterminées à l'instant t. Cette équation est une généralisation de l'équation de diffusion et elle a toujours la forme

$$\frac{\partial\omega}{\partial t} = \mathcal{H}\omega$$

où $\mathcal{H}$ est un opérateur différentiel qui est déterminé par la nature du système et les influences régulières et irrégulières auxquelles il est soumis ; ω est fonction de l'état du système et de t.

L'analogie superficielle qui existe entre cette théorie de probabilité classique et la mécanique ondulatoire, interprétée d'une manière statistique n'a probablement échappé à aucun physicien qui les connaît toutes les deux. La forme de l'équation des ondes pour un système quelconque est à peu près la même que celle de l'équation de FOKKER :

$$-\frac{h}{2\pi i}\frac{\partial\psi}{\partial t} = \mathcal{H}\psi.$$

Même la forme de l'opérateur $\mathcal{H}$ est la même dans les cas les plus simples, où il se réduit à l'opérateur de LAPLACE. Pourtant l'analogie ne peut être considérée que comme superficielle pour deux rai-

sons, et voilà pourquoi on n'en a pas beaucoup parlé jusqu'à présent. En premier lieu, le coefficient imaginaire a pour conséquence que l'équation des ondes, malgré son caractère parabolique *apparent*, est *essentiellement* hyperbolique, et par conséquent décrit des phénomènes essentiellement réversibles tandis que les phénomènes de diffusion, etc., sont essentiellement irréversibles. En second lieu, ψ n'est pas la probabilité, mais ce qu'on appelle l'amplitude de probabilité : les probabilités elles-mêmes sont des expressions bilinéaires en ψ et ψ^* ; dans le cas le plus simple cette probabilité est égale à :

$$\psi\psi^* dx.$$

Donc, malgrè l'analogie *des équations*, l'appareil mathématique du calcul des probabilités est assez différent dans les deux théories et cela au point que les traits les plus caractéristiques de la mécanique des quanta font complètement défaut en théorie classique.

Dans cette conférence, je veux attirer votre attention sur le fait qu'il existe des problèmes classiques qui présentent beaucoup plus d'analogie avec la mécanique ondulatoire et qui n'ont pas du tout été traités jusqu'à présent. Il s'agit de problèmes concernant exactement les mêmes systèmes avec les mêmes influences, régulières et irrégulières, la seule différence consistant uniquement dans la manière de poser la question. Pour simplifier parlons du cas le plus simple : mouvement brownien à une dimension, où l'équation de FOKKER se réduit à :

$$\frac{\partial \omega}{\partial t} = D \frac{\partial^2 \omega}{\partial x^2}; \quad \omega \text{ fonction de } x \text{ et de } t. \tag{2}$$

Supposez que quelqu'un (observateur A) ait observé l'état du système, c'est-à-dire la coordonnée x à l'instant t_0 *et* à l'instant t_1. Quelqu'un d'autre (observateur B) a observé x à l'instant t *entre* t_0 et t_1

$$t_0 < t < t_1.$$

L'observateur A veut déduire de *ses* observations la probabilité d'un résultat déterminé pour l'observation de B. C'est un problème tout à fait raisonnable, et qui peut encore être généralisé de la manière suivante : on peut admettre que l'observateur A n'ait pas observé *exactement* la coordonnée x aux instants t_0, t_1 mais qu'il ne connaisse que la *probabilité* pour ces deux instants :

$$\omega(x, t_0) = \omega_0(x) \; ; \quad \omega(x, t_1) = \omega_1(x)$$

Avec l'une *ou* l'autre de ces deux données, le problème serait d'un type bien connu. Avec la première donnée seule, ω serait celle des solutions de (2) qui se réduit à $\omega_0(x)$ pour t_0. Avec la seconde donnée seule, ω serait celle des solutions de

$$-\frac{\partial\omega}{\partial t} = D\frac{\partial^2\omega}{\partial x^2} \tag{3}$$

qui se réduit à ω_1 pour t_1. Mais si l'on possède à la fois les deux donées, ω ne sera ni l'une ni l'autre, car la solution de (2) est uniquement déterminée par sa valeur initiale, celle de (3) par sa valeur finale. Ni l'une ni l'autre n'aura la complaisance de se réduire à une valeur prescrite arbitrairement pour un *second* instant. Il s'agit donc d'un type de problème tout à fait nouveau. Je vous indiquerai la solution avant de la démontrer. ω est le *produit* d'une solution de (2) et d'une solution de l'équation adjointe (3). Les deux solutions devront être choisies de façon que leur produit satisfasse à la condition initiale et à la condition finale. Je n'ai pas pu réussir à prouver ni qu'il existe toujours de telles solutions, ni qu'elles sont uniques. Mais j'en suis tout à fait convaincu. Ce problème de conditions aux limites paraît être nouveau même en mathématiques.

La démonstration est un peu longue, j'espère qu'elle ne vous fatiguera pas trop. J'appellerai « probabilité intermédiaire » la fonction $\omega(x, t)$ que nous cherchons et probabilités finale et initiale, ou encore probabilités terminales les quantités ω_0 et ω_1. Commençons par traiter le cas particulier signalé auparavant, ou les probabilités terminales sont toutes les deux des *certitudes*. La particule étant observée à l'instant t_0 au point x_0, et à l'instant t_1 au point x_1, quelle est la probabilité de la trouver à l'instant t entre x et $x + dx$? Laissons partir de x_0 à l'instant t_0 un très grand nombre N de particules (se déplaçant indépendamment). Nous ne nous intéressons qu'à celles d'entre elles qui se trouvent près de x_1 à l'instant t_1, soit entre x_1 et $x_1 + dx_1$. Leur nombre est donné par une formule bien connue, la solution fondamentale de l'équation de diffusion :

$$\begin{aligned} n_1 &= N\frac{1}{\sqrt{4\pi D(t_1 - t_0)}}\, e^{-\frac{(x_1 - x_0)^2}{4D(t_1 - t_0)}}\, dx_1 \\ &= N \cdot g(x_1 - x_0, t_1 - t_0)dx_1, \end{aligned} \tag{4}$$

pour abréger. D'autre part la fraction des N particules qui satisfait à la double condition de : 1) se trouver dans l'intervalle $(x, x + dx)$

à l'instant t et 2) se trouver dans l'intervalle $(x_1, x_1 + dx_1)$ à l'instant t_1, est donnée par

$$n = \mathrm{N} \cdot g(x - x_0, t - t_0)dx \cdot g(x_1 - x, t_1 - t)dx_1,$$

puisque ces deux probabilités sont évidemment indépendantes. Or, la probabilité que nous cherchons est évidemment $\frac{n}{n_1}$. Donc nous trouvons dans ce cas spécial :

$$\omega(x, t) = \frac{g(x - x_0, t - t_0)g(x_1 - x, t_1 - t)}{g(x_1 - x_0, t_1 - t_0)}. \tag{5}$$

Il est facile de voir que ω est le produit d'une solution de (2) et d'une solution de (3), en accord avec le résultat général que nous avons indiqué.

Pour traiter le cas général avec des probabilités terminales ω_0, ω_1, il faut encore se figurer un très grand nombre N de particules, mais qui ne partent pas toutes du même endroit. Il faut en faire partir le nombre

$$\mathrm{N}\omega_0(x_0)dx_0 \tag{A}$$

de la « cellule » $(x_0, x_0 + dx_0)$. et répéter cette expérience *un grand nombre de fois,* car dans la plupart des cas, elle n'aura pas le résultat désiré à savoir, amener

$$\mathrm{N}\omega_1(x_1)dx_1 \tag{B}$$

particules en $(x_1, x_1, + dx_1)$ à l'instant t_1. Pourtant la probabilité que *cette dernière* répartition se réalise à l'instant t_1 n'est pas zéro, et ce qu'il faut faire (seulement en pensée, heureusement !) c'est de répéter l'expérience un nombre suffisant de fois jusqu'à ce que même la petite proportion d'expériences qui amènent le résultat désiré, forme un très grand nombre. Alors cette petite fraction représente précisément le matériel statistique qui convient à la question posée.

Maintenant, tout ce qu'il nous faut savoir, c'est combien de particules (A) se trouvent parmi les particules (B). Si nous connaissions ce nombre pour une valeur quelconque de x_0 et x_1, il suffirait de le multiplier par l'expression (5) et par dx pour trouver la proportion de ces particules (provenant de x_0 et arrivant en x_1,) qui se trouvent entre x et $x + dx$ à l'instant t. En les ajoutant pour toutes les valeurs de x_0 et de x_1, c'est-à-dire en intégrant par rapport à x_0 et x_1, on trouvera le nombre que nous cherchons à déterminer.

Donc le problème principal est de trouver le nombre des (A) qui font partie des (B). Pour cela, il ne suffit pas de fixer son attention sur *un* groupe (A) et *un* groupe (B) : il faut les envisager tous ensemble. Divisons l'échelle des x en cellules égales, que nous prendrons, pour simplifier les notations, comme unité de longueur. Soit a_k le nombre de particules provenant de la $k^{ième}$ cellule et b_l le nombre des particules arrivant à la $l^{ième}$ cellule. Soit c_{kl} le nombre des particules provenant de la $k^{ième}$ cellule *et arrivant* à la $l^{ième}$. Par g_{kl} nous désignerons la probabilité *à priori* pour qu'une particule qui provient de la $k^{ième}$ cellule arrive à la $l^{ième}$ (g_{kl} n'est donc autre chose que l'expression (4) abrégée). Les c_{kl} satisferont aux relations :

$$\sum_l c_{kl} = a_k, \qquad \sum_k c_{kl} = b_l \tag{6}$$

et à aucune autre relation : un système quelconque de c_{kl} qui est en accord avec les relations précédentes est *possible*, et arrivera quelquefois à se réaliser. Donc les nombres c_{kl} qui sont justement ceux dont nous avons besoin, ne sont pas du tout fixés univoquement : il faudrait en effet rechercher la *statistique* des c_{kl}. Je ne l'ai pas fait. Il me semble assez évident que dans le cas limite où le nombre N est très grand, ne se réaliseront que des systèmes de nombres c_{kl} qui sont très voisins du système le plus probable, c'est-à-dire du système c_{kl} pour lequel l'événement improbable que nous admettons comme constaté (à savoir les nombres a_k et b_l) acquiert au moins la plus grande probabilité (plus grande que pour tout autre système c_{kl}).

Si nous connaissions pour chaque particule individuelle aussi bien la cellule d'origine que la cellule de destination, la probabilité pour que l'événement observé, c'est-à-dire la transformation de la répartition a_k en b_l se réalise de *cette* manière particulière, serait

$$\prod_k \prod_l g_{kl}^{c_{kl}}.$$

Mais ce n'est pas encore la probabilité entière pour que cet événement se réalise au moyen d'un système déterminé de c_{kl}. Car on pourra permuter les a_k particules provenant de la $k^{ième}$ cellule

$$\frac{a_k!}{\prod_l c_{kl}!}$$

fois, et de même pour tous les k, donc

$$\prod_k \frac{a_k!}{\prod_l c_{kl}!} = \frac{\prod_k a_k!}{\prod_k \prod_l c_{kl}!}$$

fois. Donc la probabilité entière est

$$\prod_k a_k! \prod_k \prod_l \frac{g_{kl}^{c_{kl}}}{c_{kl}!}.$$

Il faut déterminer le maximum de cette expression quand on fait varier les c_{kl} en tenant compte des relations (6). Comme $\prod_k a_k!$ ne varie pas, il suffira de trouver le maximum du produit double ou, ce qui est plus commode, de son logarithme. Donc

$$\delta \sum_k \sum_l \{ c_{kl} \log g_{kl} - c_{kl} (\log c_{kl} - 1) \}$$
$$+ \delta \sum_k \lambda_k \sum_l c_{kl} + \delta \sum_l \mu_l \sum_k c_{kl} = 0$$

ou

$$\sum_k \sum_l \delta c_{kl} \{ \log g_{kl} - \log c_{kl} + \lambda_k + \mu_l \} = 0.$$

Donc

$$c_{kl} = e^{\lambda_k + \mu_l} g_{kl}.$$

Posons

$$\psi_k = e^{\lambda_k} \qquad \varphi_l = e^{\mu_l}.$$

de façon que

(7) $$c_{kl} = \psi_k \varphi_l g_{kl}.$$

Les ψ_k, φ_l seront déterminés par

(8) $$\varphi_l \sum_k \psi_k g_{kl} = b_l, \qquad \psi_k \sum_l \varphi_l g_{kl} = a_k.$$

La solution est contenue dans (7) et (8), il faut seulement les traduire en langage du continu. Nous écrirons

$$\begin{array}{lll} c(x_0, x_1)dx_0dx_1 & \text{pour} & c_{kl} \\ N\omega_0(x_0)dx_0 & \text{»} & a_k \\ N\omega_1(x_1)dx_1 & \text{»} & b_l \\ \sqrt{N}\psi(x_0)dx_0 & \text{»} & \psi_k \\ \sqrt{N}\varphi(x_1)dx_1 & \text{»} & \varphi_l. \end{array}$$

En effet il est évident qu'à ψ_k, φ_l doivent correspondre des fonctions de x_0 et de x_1 ; la racine carrée $\sqrt{N}$ n'est introduite que pour simplifier les notations. Nous trouvons alors

$$(9) \qquad \begin{aligned} \psi(x_0) \int \varphi(x_1) g(x_1 - x_0, t_1 - t_0)dx_1 &= \omega_0(x_0) \\ \varphi(x_1) \int \psi(x_0) g(x_1 - x_0, t_1 - t_0)dx_0 &= \omega_1(x_1) \end{aligned}$$

et

$$c(x_0, x_1)dx_0dx_1 = Ng(x_1 - x_0, t_1 - t_0)\psi(x_0)\varphi(x_1)dx_0dx_1.$$

Voilà le nombre de particules qui proviennent de x_0 et arrivent en x_1. Comme nous l'avons indiqué, il faut le multiplier par l'expression (5) et par dx (et diviser par N), puis intégrer par rapport à x_0 et x_1 pour trouver la « probabilité intermédiaire ». Celle-ci a donc la valeur suivante :

$$\omega(x, t) = \int g(x - x_0, t - t_0)\psi(x_0)dx_0 \cdot \int g(x_1 - x, t_1 - t)\varphi(x_1)dx_1.$$

On voit aisément que le premier facteur est solution de (2), le deuxième solution de (3).

φ et ψ sont déterminés par les équations intégrales (9) qui me semblent être assez compliquées quoiqu'elles n'expriment rien d'autre que les conditions aux limites qui sont imposées au produit des deux solutions.

On peut ajouter à ce que nous venons de démontrer quelques remarques qui sont tout à fait intéressantes, en dehors de toute analogie avec la mécanique ondulatoire. En premier lieu, on démontre aisément que $\omega(x,t)$ satisfait à la condition que son intégrale donne une valeur constante, à savoir 1 (si ω_0 et ω_1 ont été normées de la même façon). Ensuite on peut montrer que le « centre de gravité », c'est-à-dire $\int x\omega(x, t)dx$ se

meut à vitesse constante à partir de sa position initiale jusqu'à sa position finale. Dans le cas simple où les probabilités aux limites sont des certitudes, le maximum de probabilité est lui aussi fonction linéaire du temps : il s'agit à chaque instant d'une répartition de GAUSS qui s'élargit jusqu'à ce que la moitié du temps se soit écoulée, puis qui se contracte vers le point final.

D'autres remarques assez intéressantes tiennent à ce que le rôle des deux fonctions ω_0, ω_1 est absolument symétrique ; on peut dire qu'aucune des deux directions du temps n'est privilégiée. En échangeant ω_0 et ω_1, on trouve exactement la même évolution de probabilité, mais en sens inverse, de t_1 vers t_0.

Si les probabilités terminales ω_1, ω_2 sont données de telle manière que ω_1 soit exactement la répartition à l'instant t_1 qui découlerait automatiquement (c'est-à-dire d'après l'équation de diffusion) de la répartition ω_0, donnée à t_0, alors la solution de nos deux équations intégrales est presque évidente : on doit prendre

$$\varphi \equiv 1, \qquad \psi \equiv \omega_0$$

et l'évolution de la probabilité a lieu comme dans un problème de diffusion ordinaire.

Mais le cas « inverse » pour ainsi dire, est tout aussi simple ; si ω_0 et ω_1 sont donnés de manière que ω_1 résulterait de ω_0 dans le temps $t_1 - t_0$ d'après l'équation *adjointe*, ou, ce qui revient au même, si ω_0 résultait de ω_1 par l'équation de diffusion ordinaire, alors la solution est également presque évidente : il faut prendre

$$\varphi \equiv \omega_1, \qquad \psi \equiv 1$$

et tout se passera comme d'après l'équation « inversée ».

On peut en déduire une conclusion assez curieuse sur la manière dont se produisent les fluctuations thermodynamiques considérables qui arrivent quelquefois (quoique très rarement) dans un système en équilibre. Imaginez que vous observez un système de particules en diffusion, qui soient en équilibre thermodynamique. Admettons qu'à un instant donné t_0 vous les ayez trouvées en répartition à peu près uniforme et qu'à $t_1 > t_0$ vous ayez trouvé un écart spontané et *considérable* par rapport à cette uniformité. On vous demande de quelle manière cet écart s'est produit. Quelle en est la manière la plus probable ?

La réponse est la suivante : le plus probable, c'est qu'il se soit produit par un renversement complet des lois de diffusion, de sorte qu'à chaque instant, le courant de diffusion ait été dirigé dans la direction du *gradient* (et non de la « chute ») de la concentration et à une intensité qui correspond exactement, au signe près, à la valeur ordinaire de la constante de diffusion D.

Je crois même que le degré de certitude avec lequel on peut indiquer cette probabilité est exactement le même que dans le cas direct. En outre, je crois que tout cela est applicable non seulement au cas de la diffusion, mais à toutes les autres lois irréversibles de la physique.

Il ne me reste plus qu'à attirer votre attention sur la grande analogie que présentent les probabilités envisagées en mécanique quantique avec ce que nous avons appelé *probabilités intermédiaires* dans les problèmes du type précédent.

Cette analogie est beaucoup plus marquée pour les problèmes de ce type que pour les anciens problèmes que SMOLUCHOWSKI, FOKKER et PLANCK avaient traités. Dans les nouveaux problèmes, non seulement l'équation de FOKKER ressemble à l'équation fondamentale de la mécanique quantique, mais, tout comme dans cette dernière, la probabilité intermédiaire est donnée par le *produit* de deux solutions de *deux* équations qui ne diffèrent entre elles que par le signe de la variable temps. C'est précisément grâce à cette symétrie par rapport au temps que l'évolution de la probabilité intermédiaire est *réversible*, de la même manière que celle de la probabilité quantique. Cela veut dire qu'en changeant le signe du temps dans l'expression d'une probabilité intermédiaire, qui évolue en accord avec les équations générales du problème, on arrive à une expression qui définit également une évolution compatible avec ces équations et qui, par conséquent, peut se réaliser en choisissant convenablement les conditions aux limites. Il n'en est pas du tout ainsi dans les problèmes habituels de mouvement brownien, lesquels se réduisent mathématiquement à des problèmes de diffusion ou de conduction de chaleur et définissent une évolution essentiellement irréversible.

Néanmoins, une différence considérable subsiste, qu'il faut signaler : l'évolution de la probabilité intermédiaire *n'est pas ondulatoire*. Cela tient au fait que dans les équations (2) et (3) la constante D est essentiellement réelle, tandis que dans l'équation des ondes de la mécanique

quantique intervient d'une façon assez mystérieuse la racine de l'unité négative $\sqrt{-1}$. Mathématiquement, cela implique :

1° Le caractère ondulatoire de l'évolution de la fonction ψ ;

2° Son caractère essentiellement complexe.

La fonction complexe ψ correspond à *deux* fonctions réelles, de sorte qu'il suffit de définir les conditions aux limites en se donnant la valeur de ψ à *un seul* instant déterminé ; c'est la façon de voir généralement adoptée en mécanique quantique. Est-elle la seule admissible ? Dans notre problème, cela reviendrait à regarder comme données les valeurs de ψ et de φ à un instant déterminé (au lieu des valeurs de leur produit, à deux instants différents), chose inadmissible et absolument dénuée de sens.

Doit-on interpréter la remarque d'Eddington, citée plus haut, comme signalant la nécessité de modifier cette manière de voir en mécanique ondulatoire et prendre comme conditions aux limites les valeurs d'une seule probabilité réelle à deux instants différents ?

NOTES MATHÉMATIQUES

NOTE I

La démonstration est assez facile en faisant usage de nos conceptions de pair, impair, etc. Je vais l'indiquer puisqu'elle n'a pas encore été publiée.

Il faut savoir de quelle manière une transformation de LORENTZ agit sur ψ. Premièrement, on transforme les coordonnées, x_1, x_2, x_3, t de la manière bien connue ; puis il faut transformer la variable ζ, c'est-à-dire appliquer une certaine transformation linéaire à ψ_1, ψ_2. ψ_3, ψ_4, dont les coefficients dépendent de ceux de la transformation de LORENTZ.

Pour le but que nous avons en vue, il suffit d'utiliser une transformation infinitésimale. On peut l'écrire en opérateur s'appliquant à ψ :

$$\mathcal{L} = 1 + \varepsilon_{kl}\left(x_k \frac{\partial}{\partial x_l} - x_l \frac{\partial}{\partial x_k}\right); \qquad x_4 = ict$$

(sommer par rapport à k et à l).

Mais il faut, comme je l'ai indiqué, appliquer encore un autre opérateur opérant seulement sur ζ, à savoir le suivant :

$$\Lambda = 1 + \frac{\varepsilon_{kl}}{2}\gamma_k\gamma_l \qquad \begin{aligned} \gamma_1 &= -i\alpha_4\alpha_1 \\ \gamma_2 &= -i\alpha_4\alpha_2 \\ \gamma_3 &= -i\alpha_4\alpha_3 \\ \gamma_4 &= \alpha_4. \end{aligned}$$

Prenons le cas particulier :

$$k = 1, \qquad l = 4,$$

qui nous suffira, c'est-à-dire supposons que seulement ε_{14} (que nous appellerons ε) diffère de zéro. Nous aurons :

$$\gamma_k\gamma_l = -i\alpha_4\alpha_1\alpha_4 = +i\alpha_1$$

$$\Lambda\mathfrak{L} = 1 + \frac{i\varepsilon}{2}\alpha_1 + \varepsilon\left(x_1\frac{\partial}{\partial(ict)} - ict\frac{\partial}{\partial x_1}\right).$$

Cet opérateur doit *laisser* ψ positive. Donc sa partie impaire, appliquée à ψ, doit donner zéro.

Dans tout ceci, la multiplication par t, de même que la différentiation $\frac{\partial}{\partial t}$ doivent être regardées comme des opérateurs pairs, puisque : 1) la multiplication par un paramètre ne change pas le signe ; 2) ψ est supposé *identiquement* positif en t. Le dernier terme est donc pair ; les parties impaires du second et du troisième terme sont évidemment (en supprimant $i\varepsilon$) :

$$\frac{1}{2}\eta_1 - \frac{1}{c}\xi_1\frac{\partial}{\partial t};$$

mais

$$\xi_1 = -\frac{c\varkappa}{2}\eta_1\mathcal{H}^{-1}.$$

Donc, $\left(\text{en supprimant le facteur } \frac{1}{2}\right)$:

$$\left(\eta_1 + \varkappa\eta_1\mathcal{H}^{-1}\frac{\partial}{\partial t}\right)\psi = 0,$$

$$\eta_1\mathcal{H}^{-1}\left(\mathcal{H} + \varkappa\frac{\partial}{\partial t}\right)\psi = 0,$$

$$-\mathcal{H}^{-1}\eta_1\left(\mathcal{H} + \varkappa\frac{\partial}{\partial t}\right)\psi = 0;$$

ou, en appliquant l'opérateur $\mathcal{H}$:

$$\eta_1\left(\mathcal{H} + \varkappa\frac{\partial}{\partial t}\right)\psi = 0.$$

η_1 n'a pas de valeur propre nulle (au moins si l'énergie reste finie). Donc il faut que

$$\left(\mathcal{H} + \varkappa\frac{\partial}{\partial t}\right)\psi = 0 \qquad \text{C. Q. F. D.}$$

NOTE II

Les opérateurs du spin, dont il n'a pas encore été question dans ces conférences, sont les *produits* des α_k, de façon que, par exemple

$$s_1 = -i\alpha_2\alpha_3 = +i\alpha_3\alpha_2.$$

Déterminons la *partie paire* des s_k, que nous indiquerons par $\widehat{s}_k$. La décomposition de α_k étant

$$\alpha_k = c\mathcal{H}^{-1}p_k + \eta_k,$$

on trouvera $\widehat{s}_k$ en ajoutant au produit des parties paires des α_2 et α_3 le produit de leurs parties impaires.

$$\begin{aligned}\widehat{s}_1 &= -i(c^2\mathcal{H}^{-2}p_2p_3 + \eta_2\eta_3)\\ &= +i(c^2\mathcal{H}^{-2}p_2p_3 + \eta_3\eta_2)\end{aligned}$$

En ajoutant ces deux équations :

$$\widehat{s}_1 = \frac{i}{2}(\eta_3\eta_2 - \eta_2\eta_3).$$

La partie paire du spin est le commutateur des η_k, (de même que le spin lui-même est le commutateur des α_k). De cette équation on tire également :

1) que le spin est pair *pour la plus grande part* ;

2) que sa partie paire est une intégrale du mouvement en l'absence de champ ; en effet les $\widehat{s}_k$ commutent avec $\mathcal{H}$ parce que les η_k *anticommutent.*

Comme

$$\xi_k = \frac{c\varkappa}{2}\mathcal{H}^{-1}\eta_k,$$

les commutateurs des η_k sont intimement liés aux commutateurs des ξ_k. On trouve aisément (attention au signe !)

$$\xi_2\xi_3 - \xi_3\xi_2 = -\frac{ic^2\varkappa^2}{2}\mathcal{H}^{-2}\widehat{s}_1.$$

Les commutateurs des ξ_k sont les mêmes que ceux des x_k, avec le signe opposé. Pour le voir, écrivons que les x_k commutent :

$$x_kx_l - x_lx_k = 0\,;$$

remplaçons-les par leurs décompositions :

$$\left(\widehat{x}_k + \xi_k\right)\left(\widehat{x}_l + \xi_l\right) - \left(\widehat{x}_l + \xi_l\right)\left(\widehat{x}_k + \xi_k\right) = 0,$$

et égalons à zéro la partie *paire* de cette équation. Par conséquence :

$$\widehat{x_2}\widehat{x_3} - \widehat{x_3}\widehat{x_2} = + \frac{ic^2\varkappa^2}{2}\mathcal{H}^{-2}\widehat{s_1} = - \frac{ic^2\varkappa^2}{8\pi^2}\mathcal{H}^{-2}\widehat{s_1}.$$

Comme $\mathcal{H}^{-2}$ est en général de l'ordre de grandeur de $m^{-2}c^{-4}$, les commutateurs sont de l'ordre de

$$\frac{1}{2}\left(\frac{h}{4\pi mc}\right)^2.$$

C'est peut-être l'ordre de grandeur auquel il fallait s'attendre puisque la commutation n'a été troublée que par l'addition d'opérateurs dont l'ordre de grandeur est $\frac{h}{4\pi mc}$.

NOTE III

Ce qui subsiste en tout cas, c'est l'incertitude de HEISENBERG. Si l'on désire une précision λ sur les coordonnées, cela implique une incertitude sur la vitesse que l'on peut estimer grossièrement à

$$\Delta v = \frac{h}{4\pi m\lambda},$$

où m est la masse électronique, en supprimant le facteur $\sqrt{1-\beta^2}$, l'ordre de grandeur seul étant intéressant. Rien n'empêche, il est vrai, de faire λ aussi petit que l'on voudra. Mais pour que cette précision ait une signification physique, il faut en même temps déterminer le *temps* avec une précision de τ où

$$\tau \cdot \Delta v \leqslant \lambda.$$

Maintenant ne parlons plus de la *masse* de la montre qui sert à mesurer le temps, parlons de sa *grandeur*. Je vais démontrer qu'elle doit être extrêmement petite.

Soit l cette grandeur ; c'est évidemment aussi l'*incertitude* sur le *lieu* pour lequel elle indique le temps. En appliquant son indication de temps à un mobile de vitesse Δv, cela impliquera une incertitude

$$\frac{l\Delta v}{c^2} \leqslant \tau$$

qui devrait être au plus égale à τ, donc

$$\frac{l(\Delta v)^2}{c^2} \leqslant \lambda$$

$$\frac{l}{\lambda^2}\left(\frac{h}{4\pi mc}\right)^2 \leqslant \lambda$$

$$l \leqslant \frac{\lambda^3}{\left(\frac{h}{4\pi mc}\right)^2}$$

Si l'on veut augmenter la précision au delà de $\frac{h}{4\pi mc}$, il faut avoir une montre *plus petite* que cela, c'est-à-dire de 10^{-11} cm. (Et en même temps très lourde).

NOTE IV

Formellement, la marche de la démonstration est la même que pour la proposition analogue concernant des observables qui sont canoniquement conjuguées.

Il doit exister un opérateur hermitique $\mathcal{A}$ (« aiguille ») tel que $\psi(x, t)$ soit identiquement fonction propre de $\mathcal{A}$ avec la valeur propre t :

$$\mathcal{A}\psi(x, t) - t\psi(x, t) = 0.$$

En remplaçant ψ par l'intégrale de Fourier on trouve

$$\int e^{-i\omega t}(\mathcal{A}c - tc)d\omega = 0.$$

L'intégration par parties du second terme donne

$$\int e^{-i\omega t}\left(\mathcal{A}c - i\frac{\partial c}{\partial \omega}\right)d\omega = 0$$

donc

$$\mathcal{A}c - i\frac{\partial c}{\partial \omega} = 0$$

et aussi

$$\mathcal{A}^*c^* + i\frac{\partial c^*}{\partial \omega} = 0.$$

En multipliant respectivement par c^* et c et ajoutant :

$$\frac{\partial}{\partial \omega}|c|^2 = i(c^*\mathrm{A}c - c\mathrm{A}^*c^*)$$

L'intégrale $\int dx$ du second membre est nulle par suite des propriétés des opérateurs hermitiques (ou self-adjoints). L'intégrale $\int |c|^2 dx$ est donc indépendante de ω, comme nous l'avions annoncé.

Comme *toutes* les valeurs de l'énergie sont également probables, il s'ensuit qu'à un intervalle limité d'énergie appartient une probabilité nulle. Donc, il est infiniment improbable de trouver pour l'énergie de notre montre une valeur *finie*. Donc elle est physiquement impossible !

(Conférences faites à l'Institut Henri-Poincaré en mai 1931).

Manuscrit reçu le 25 mai 1931