

Sur la théorie de la mécanique quantique

P. A. M. Dirac

1. Introduction et résumé.

La nouvelle mécanique de l'atome introduite par Heisenberg¹ peut être basée sur la supposition que les variables qui décrivent un système dynamique n'obéissent pas à la loi commutative de la multiplication, mais vérifient à la place certaines conditions quantiques. On peut établir une théorie en ne sachant rien des variables dynamiques si ce n'est les lois algébriques auxquelles elles sont soumises, et on peut montrer qu'elles peuvent être représentées par des matrices à chaque fois qu'un ensemble de variables uniformisantes existe pour le système dynamique.² On peut montrer pourtant (voir § 3), qu'il n'y a pas d'ensemble de variables uniformisantes pour un système contenant plus d'un électron, de telle façon que la théorie ne peut pas progresser beaucoup en suivant ce genre de voie.

Un nouveau développement de la théorie a été donné récemment par Schrödinger.³ En partant de l'idée qu'un système atomique ne peut pas être représenté par une trajectoire, i.e., par un point bougeant dans l'espace des coordonnées, mais doit être représenté par une onde dans l'espace, Schrödinger obtient d'un principe de variation une équation différentielle que la fonction d'onde doit satisfaire. Cette équation différentielle s'avère être très liée à l'équation de l'Hamiltonien qui spécifie le système, notamment, si

$$H(q_r, p_r) - W = 0$$

est l'équation de l'Hamiltonien du système, quand les q_r , p_r sont les variables canoniques, alors l'équation d'onde pour ψ est

$$(1) \quad \left\{ H \left(q_r, ih \frac{\partial}{\partial q_r} \right) - W \right\} \psi = 0$$

où h est $(2\pi)^{-1}$ fois la constante de Planck habituelle. Chaque moment p_r dans H est remplacé par l'opérateur $ih\partial/\partial q_r$, et est supposé opérer sur tout ce qui apparaît sur son côté droit dans le terme dans lequel il apparaît. Schrödinger prend la valeur du paramètre W pour laquelle il existe un ψ satisfaisant (1) qui est continue, à valeur unique et bornée sur la totalité du q -espace pour représenter les niveaux d'énergie du système, et montre que quand la solution générale de (1) est connue, les matrices représentant les p_r et les q_r peuvent facilement être obtenues, respectant toutes les conditions qu'elles doivent satisfaire selon la mécanique des matrices d'Heisenberg, et consistantes avec les niveaux d'énergie précédemment trouvés. L'équivalence mathématique des théories est ainsi établie.

Dans le présent article, la théorie de Schrödinger est considérée dans le § 2 d'un point de vue légèrement plus général, dans lequel le temps t et son moment conjugué $-W$ sont traités depuis le début sur le même pied que les autres variables. Une méthode plus générale, nécessitant seulement

P. A. M. Dirac, Lycée St. John, Cambridge. (Communiqué par R. H. Fowler, F.R.S., Reçu le 26 Août 1926.)
traduction Denise Vella-Chemla, janvier 2021.

1. Voir plusieurs articles de Born, Heisenberg et Jordan, Zeits. f. Phys., à partir du vol. 33.

2. Roy. Soc. Proc., A, vol. 110, p. 561 (1926).

3. Voir les articles variés dans les Ann. d. Phys., en commençant par le vol. 79, p. 361 (1926).

de l'algèbre symbolique élémentaire, d'obtenir des représentations matricielles des variables dynamiques est fournie.

Au § 3, on considère le problème d'un système contenant plusieurs particules semblables, tel qu'un atome avec plusieurs électrons. Si les positions de deux des électrons sont interchangées, le nouvel état de l'atome est physiquement indistinguable de l'état original. Dans un tel cas, on s'attendrait à ne pouvoir représenter par des matrices que les fonctions symétriques des coordonnées de tous les électrons. On trouve que cela permet d'obtenir deux solutions du problème satisfaisant toutes les conditions nécessaires, et la théorie est incapable de décider laquelle des deux est correcte. Une des solutions amène au principe de Pauli selon lequel pas plus d'un électron ne peut être sur une orbite donnée, et l'autre, quand elle est appliquée au problème analogue du gaz parfait, amène à la mécanique statistique d'Einstein-Bose.

L'effet d'une perturbation arbitraire d'un système atomique est étudié au § 5 à l'aide d'une nouvelle supposition. La théorie est appliquée à l'absorption et à l'émission stimulée de radiation par un atome. Une généralisation de la description du phénomène par les coefficients B d'Einstein est obtenue, dans laquelle les phases jouent leur propre rôle. La méthode ne peut pas être appliquée à l'émission spontanée.

2. Théorie générale.

Selon le nouveau point de vue introduit par Schrödinger, nous ne laissons plus non spécifiée la nature des variables dynamiques qui décrivent un système atomique, mais nous comptons les q et les t comme des variables mathématiques ordinaires (ceci étant permis par le fait qu'elles permutent l'une avec l'autre) et nous appelons p et W les opérateurs différentiels

$$(2) \quad p_r = -ih \frac{\partial}{\partial q_r}, \quad -W = -ih \frac{\partial}{\partial t}.$$

À chaque fois qu'un p_r ou W intervient dans un terme d'équation, il doit être considéré comme représentant l'opérateur différentiel correspondant agissant sur tout ce qu'il y a à sa droite dans le terme en question. Ainsi, en effectuant cela, on peut réduire n'importe quelle fonction de p , q , W et t à une fonction des q et t seulement.

Les relations (2) nécessitent que soient effectuées deux modifications évidentes dans l'algèbre gouvernant les variables dynamiques. D'abord, seules les fonctions rationnelles entières des p et W ont un sens, et ensuite, on peut multiplier une équation par un facteur (intégral selon les p et W) du côté gauche, mais on ne peut pas, en général, multiplier par un facteur du côté droit. Ainsi, si on donne l'équation $a = b$, on peut inférer d'elle que $Xa = Xb$, où X est quelconque, mais on ne peut pas en déduire que $aX = bX$.

Il y a, pourtant, certaines équations $a = b$ pour lesquelles il est vrai que $aX = bX$ pour tout X , et nous appelons ces équations des identités. Les conditions quantiques

$$q_r p_s - p_s q_r = ih \delta_{rs}, \quad p_r p_s - p_s p_r = 0,$$

avec les relations similaires dans lesquelles interviennent $-W$ et t , sont des identités, comme il peut facilement être vérifié (et cela a été vérifié par Schrödinger) que les relations

$$(q_r p_s - p_s q_r)X = ih\delta_{rs}X,$$

etc., sont vérifiées pour tout X . Ces relations constituent la principale justification des hypothèses (2).

Si $a = b$ est une identité, nous pouvons déduire, puisque $aX = bX$ et $Xa = Xb$, que

$$aX - Xa = bX - Xb,$$

ou

$$[a, X] = [b, X].$$

Par conséquent, nous pouvons égaler le crochet de Poisson de n'importe quel côté d'une identité avec une quantité arbitraire, et ainsi notre identité quantique est l'analogie d'une identité de la théorie classique. Nous supposons que l'équation générale $xy - yx = ih[x, y]$ et que les équations du mouvement d'un système dynamique sont des identités.

Un système dynamique est spécifié par une équation hamiltonienne entre les variables

$$(3) \quad H(q_r, p_r, t) - W = 0,$$

ou plus généralement

$$(4) \quad F(q_r, p_r, t, W) = 0,$$

et les équations du mouvement sont

$$dx/ds = [x, F],$$

où x est n'importe quelle fonction des variables dynamiques, et s est une variable qui dépend de la forme dans laquelle (4) est écrite, et en particulier, vaut juste t si (4) est écrite sous la forme (3). Dans la nouvelle théorie, on considère l'équation

$$(4) \quad F\psi = 0,$$

qui, si l'on choisit ψ comme étant une fonction des variables q et t seulement, est une équation différentielle ordinaire pour ψ . De la solution générale de l'équation différentielle, les matrices qui constituent la solution du problème mécanique peuvent être trouvées très facilement.

Puisque (5) est linéaire en ψ , sa solution générale est de la forme

$$(6) \quad \psi = \sum c_n \psi_n,$$

où les c_n sont des constantes arbitraires et les ψ_n constituent un ensemble de solutions indépendantes, qu'on peut appeler fonctions propres. Seules les solutions qui sont continues, à valeur unique, et bornées sur la totalité du domaine des q et t sont reconnues par la théorie. Plutôt qu'un ensemble discret de fonctions propres ψ_n , on peut avoir un ensemble continu $\psi(\alpha)$, dépendant d'un paramètre α , et vérifiant l'équation différentielle pour toutes les valeurs de α dans un certain domaine, auquel cas, la somme dans (6) doit être remplacée par une intégrale $\int c_\alpha \psi(\alpha) d\alpha$ ⁴, ou à la fois un ensemble discret et un ensemble continu peuvent advenir ensemble. Pour la définition, pourtant, nous écrivons seulement la somme discrète explicitement dans le travail ci-dessous.

4. La solution générale peut contenir des quantités, comme ψ_a et $\partial\psi_a/\partial a$, qui vérifient l'équation différentielle (5), mais qui ne peuvent strictement être mises sous la forme $\int c_a \psi_a da$, bien qu'on puisse les voir comme les limites de séries de quantités qui sont de cette forme.

Nous allons maintenant montrer que toute constante d'intégration du système dynamique (soit la première intégrale soit la seconde intégrale) peut être représentée par une matrice dont les éléments sont des constantes, une colonne et une ligne de la matrice correspondant à chaque fonction propre ψ_n . Soit a une constante d'intégration d'un tel système, i.e., une fonction des variables dynamiques telle que $[a, F] = 0$ identiquement. On a la relation

$$Fa = aF,$$

que nous pouvons multiplier par ψ_n du côté droit, puisque c'est une identité. Nous obtenons ainsi

$$Fa\psi_n = aF\psi_n = 0,$$

puisque $F\psi_n = 0$ (bien que non identiquement). Par conséquent $a\psi_n$ est une solution de l'équation différentielle (5), de telle façon qu'elle peut être développée sous la forme (6), i.e.,

$$a\psi_n = \sum_m \psi_m a_{mn},$$

où les a_{mn} sont des constantes. Nous prenons comme quantités a_{mn} les éléments de la matrice qui représente a . La règle de multiplication matricielle est vérifiée puisque, si b est une autre constante d'intégration du système, on a

$$ab\psi_n = a \sum_m \psi_m b_{mn} = \sum_{mk} \psi_k a_{km} b_{mn},$$

et également

$$ab\psi_n = \sum_k \psi_k (ab)_{kn},$$

de telle façon que

$$(ab)_{kn} = \sum_m a_{km} b_{mn}.$$

Comme exemple de constante d'intégration du système dynamique, on peut prendre la valeur $x(t_0)$ qu'une fonction arbitraire x des p, q, W et t prend à un instant donné $t = t_0$. La matrice qui représente $x(t_0)$ consistera en éléments dont chacun est une fonction de t_0 . En écrivant t pour t_0 , nous voyons qu'une fonction arbitraire des variables dynamiques $x(t)$, ou simplement x , peut être représentée par une matrice dont les éléments sont des fonctions de t seulement.

La représentation matricielle que nous avons obtenue n'est pas unique, puisque tout ensemble de fonctions propres indépendantes ψ_n convient. Pour obtenir les matrices de la mécanique quantique originale de Heisenberg, nous devons choisir les ψ_n d'une façon particulière. Nous pouvons toujours, par une transformation linéaire, obtenir un ensemble de ψ_n qui rend diagonale la matrice représentant toute constante d'intégration du système dynamique donnée. Supposons maintenant que l'Hamiltonien F ne contienne pas explicitement le temps, de telle façon que W est une constante du système, et est l'énergie, et choisissons ψ_n de façon à rendre diagonale la matrice représentant W , i.e., de telle façon à rendre

$$(7) \quad W\psi_n = W_n\psi_n,$$

où W_n est une constante numérique. Dénotons par x n'importe quelle fonction des variables dynamiques *dans lesquelles le temps n'intervient pas explicitement*, et posons

$$x\psi_n = \sum_m x_{mn}\psi_m,$$

où les x_{mn} sont des fonctions du temps uniquement. Nous allons montrer maintenant que les x_{mn} ne sont pas de la forme

$$(8) \quad x_{mn} = a_{mn} e^{i(W_m - W_n)t/h},$$

où les a_{mn} sont des constantes, comme dans la théorie d'Heisenberg. On a

$$(9) \quad \begin{aligned} Wx\psi_n &= \sum_m Wx_{mn}\psi_m \\ &= \sum_m (Wx_{mn} - x_{mn}W)\psi_m + \sum_m x_{mn}W\psi_m \\ &= \sum_m ih\dot{x}_{mn}\psi_m + \sum_m x_{mn}W_m\psi_m. \end{aligned}$$

Également, puisque x ne contient pas t explicitement,

$$(10) \quad \begin{aligned} Wx\psi_n &= xW\psi_n = xW_n\psi_n = W_nx\psi_n \\ &= W_n \sum_m x_{mn}\psi_m. \end{aligned}$$

En égalant les coefficients de ψ_m dans (9) et (10), on obtient

$$ih\dot{x}_{mn} = x_{mn}(W_n - W_m),$$

ce qui montre que x_{mn} est de la forme (8).

Nous avons ainsi montré que si les ψ_n sont choisis de cette manière, les matrices satisfont toutes les conditions de la mécanique matricielle de Heisenberg, exceptée la condition que les matrices qui représentent les quantités doivent être hermitiennes (i.e., doivent avoir leurs mn et nm qui sont des quantités imaginaires conjuguées). Il ne semble pas y avoir une preuve générale que c'est effectivement le cas, car la preuve devrait utiliser le fait que les ψ_n sont bornés. Il est facile de démontrer le cas particulier dans lequel la matrice représentant W est hermitienne, i.e., lorsque les W_n sont réels, puisque de (7), les ψ_n doivent être de la forme

$$\psi_n = u_n e^{-iW_n t/h},$$

où u_n est indépendant de t , et si W_n contient une partie imaginaire, ψ_n ne devrait pas rester borné lorsque t devient infini. En général, les matrices représentant des quantités réelles devraient être hermitiennes seulement si les constantes numériques arbitraires par lesquelles les ψ_n peuvent être multipliés sont choisies d'une façon particulière.

On peut voir une fonction propre ψ_n comme étant associée à des valeurs numériques définies pour certaines constantes d'intégration du système. Ainsi, si nous trouvons des constantes d'intégration a, b, \dots telles que

$$(11) \quad a\psi_n = a_n\psi_n, \quad b\psi_n = b_n\psi_n, \dots$$

où a_n, b_n, \dots sont des constantes numériques, nous pouvons dire que les ψ_n représentent un état du système dans lequel a, b, \dots prennent les valeurs numériques a_n, b_n, \dots (Noter que a, b, \dots doivent commuter pour que (11) soit possible.). De cette manière, nous pouvons avoir des fonctions propres représentant des états stationnaires d'un système atomique avec des valeurs définies pour l'énergie, le moment angulaire, et d'autres constantes d'intégration.

On devrait noter que le choix du temps t qui apparaît comme variable dans les éléments des matrices représentant des quantités variables est assez arbitraire, et que toute fonction de t et des q qui croît devrait le faire régulièrement. Pour déterminer précisément la radiation émise par le système dans la direction de l'axe des x , on devrait utiliser $(t - x/c)$ plutôt que t^5 . Il est probable que la représentation d'une constante d'intégration du système par une matrice d'éléments constants est

5. Roy. Soc. Proc., A, vol. 111, p. 405 (1926).

plus fondamentale que la représentation d'une quantité variable par une matrice dont les éléments sont des fonctions d'une variable telle que t ou $(t-x/c)$. Il semblerait qu'il soit possible de construire une théorie électromagnétique dans laquelle les potentiels du champ en un point donné x_0, y_0, z_0, t_0 dans l'espace-temps sont représentés par des matrices d'éléments constants qui sont des fonctions de x_0, y_0, z_0, t_0 .

3. Systèmes contenant plusieurs particules similaires

Dans la mécanique matricielle d'Heisenberg, on suppose que les éléments des matrices qui représentent les variables dynamiques déterminent les fréquences et les intensités des composants de la radiation émise. La théorie permet par conséquent de calculer juste les quantités qui sont importantes d'un point de vue physique, et ne donne pas d'information à propos des quantités comme les fréquences orbitales que l'on ne pourra jamais espérer mesurer expérimentalement. Nous devrions nous attendre à ce que cette caractéristique très satisfaisante perdue dans les développements ultérieurs de la théorie.

Considérons maintenant un système qui contient deux particules similaires ou davantage, disons, pour prendre un exemple précis, un atome avec deux électrons. Dénotons par (mn) cet état de l'atome dans lequel un électron est sur une orbite étiquetée par m , et l'autre sur une orbite étiquetée par n .

La question se pose de savoir si les deux états (mn) et (nm) , qui sont physiquement indistingables puisqu'ils diffèrent seulement par l'échange des deux électrons doivent être comptés comme deux états différents ou comme un seul état, i.e., donnent-ils naissance à deux colonnes et lignes dans les matrices ou seulement à une? Si la première alternative est correcte, alors la théorie devrait permettre de calculer les intensités dues aux deux transitions $(mn) \rightarrow (m'n')$ et $(mn) \rightarrow (n'm')$ séparément, puisque l'amplitude correspondant à chacun d'elles devrait être donnée par un élément défini dans la matrice représentant la polarisation totale. Les deux transitions sont, pourtant, physiquement indistingables, et seule la somme des intensités pour les deux ensemble devrait pouvoir être déterminée expérimentalement. Par conséquent, pour garder cette caractéristique essentielle de la théorie qui devrait permettre le calcul seulement pour des quantités observables, on doit adopter la seconde alternative qui est que (mn) et (nm) ne comptent que pour un état seulement.

Cette alternative, pourtant, amène également des difficultés. La symétrie entre les deux électrons nécessite que l'amplitude associée à la transition $(mn) \rightarrow (m'n')$ de x_1 une coordonnée de l'un des électrons, devrait égaler l'amplitude associée à la transition $(nm) \rightarrow (n'm')$ de x_2 , la coordonnée correspondante de l'autre électron, i.e.,

$$(12) \quad x_1(mn ; m'n') = x_2(nm ; n'm').$$

Si nous comptons maintenant (mn) et (nm) comme définissant tous deux les mêmes colonne et ligne des matrices, et similairement pour $(m'n')$ et $(n'm')$, l'équation (12) montre que chaque élément de la matrice x_1 égale l'élément correspondant de la matrice x_2 , de telle sorte que nous devrions avoir l'équation matricielle

$$x_1 = x_2.$$

Cette relation est évidemment impossible, puisque, entre autres choses, elle est inconsistante avec

les conditions quantiques. Nous devons inférer que les fonctions assymétriques des coordonnées (et des moments) des deux électrons ne peuvent être représentées par des matrices. Les fonctions symétriques, telles que la polarisation totale de l'atome, peuvent être considérées comme étant représentées par des matrices sans inconsistance, et ces matrices sont par elles-mêmes suffisantes pour déterminer toutes les propriétés physiques du système.

Une conséquence de ces considérations est que la théorie d'uniformisation des variables introduite par l'auteur ne peut plus s'appliquer. Cela est dû au fait que, correspondant à toute transition $(mn) \rightarrow (m'n')$, il devrait y avoir un terme $e^{i(aw)}$ dans les développements de Fourier, et nous devrions obliger à ce qu'il n'y ait là qu'un seul état, $(m''n'')$, disons, tel que le même terme $e^{i(aW)}$ correspond à la transition $(m'n') \rightarrow (m'',n'')$ et $e^{2i(aW)}$ correspond à $(mn) \rightarrow (m''n'')$. Si les m et les n sont des nombres quantiques, et que nous prenons le cas d'un nombre quantique par électron pour garantir la définition, nous devrions avoir

$$m'' - m' = m' - m, \quad n'' - n' = n' - n.$$

Puisque, pourtant, l'état $(m'n')$ pourrait aussi bien être appelé l'état $(n'm')$, on pourrait aussi bien prendre

$$m'' - n' = n' - m, \quad n'' - m' = m' - n,$$

qui donnerait un état différent $(m''n'')$. Il n'y a donc pas d'état unique $(m''n'')$ que la théorie des variables uniformisantes nécessite.

Si nous négligeons l'interaction entre deux électrons, alors nous pouvons obtenir les fonctions propres de l'atome tout entier simplement en multipliant les fonctions propres pour un électron quand il existe tout seul dans l'atome par les fonctions propres de l'autre électron tout seul, et en prenant la même variable de temps pour chacun⁶. Par conséquent, si $\psi_n(x, y, z, t)$ est la fonction propre pour un seul électron sur l'orbite n , alors la fonction propre pour l'atome entier dans l'état (mn) est

$$\psi_m(x_1, y_1, z_1, t)\psi_n(x_2, y_2, z_2, t) = \psi_m(1)\psi_n(2),$$

disons, où x_1, y_1, z_1 et x_2, y_2, z_2 sont les coordonnées des deux électrons et $\psi(r)$ signifie $\psi(x_r, y_r, z_r, t)$. La fonction propre $\psi_m(2)\psi_n(1)$, pourtant, correspond aussi au même état de l'atome si nous comptons les états (mn) et (nm) comme identiques. Mais deux fonctions propres indépendantes doivent donner naissance à deux lignes et colonnes dans les matrices. Si nous devons avoir seulement une ligne et une colonne dans les matrices correspondant à la fois à (mn) et (nm) , nous devons trouver un ensemble de fonctions propres de la forme

$$\psi_{mn} = a_{mn}\psi_m(1)\psi_n(2) + b_{mn}\psi_m(2)\psi_n(1),$$

où les a_{mn} et les b_{mn} sont constants, dont l'ensemble doit contenir seulement une ψ_{mn} correspondant à la fois à (mn) et (nm) , et doit être suffisant pour permettre d'obtenir la matrice représentant la fonction symétrique A des deux électrons. Cela signifie que ψ_{mn} doit être choisi de telle façon que A compte le nombre de fois où chaque ψ_{mn} peut être développé en fonction des ψ_{mn} choisis dans la forme

$$(13) \quad A\psi_{mn} = \sum_{m'n'} \psi_{m'n'} A_{m'n'.mn},$$

6. La même variable temps t doit être prise pour chacun des deux en raison du fait que nous pouvons écrire l'équation de l'Hamiltonien pour le système complet : $H(1) + H(2) - W = 0$, où $H(1)$ et $H(2)$ sont les Hamiltoniens pour les deux électrons séparément, de telle façon qu'il y a un temps commun conjugué à moins l'énergie W .

dans laquelle les $A_{m'n'.mn}$ sont des constantes ou des fonctions du temps seulement.

Il y a deux manières de choisir l'ensemble des ψ_{mn} pour satisfaire les conditions. Nous pouvons soit prendre $a_{mn} = b_{mn}$, ce qui fait de chaque ψ_{mn} une fonction symétrique des deux électrons, de telle façon que le côté gauche de l'égalité (13) sera symétrique et que seules les fonctions propres symétriques seront nécessaires pour le développement, ou bien nous pouvons prendre $a_{mn} = -b_{mn}$, ce qui rend ψ_{mn} anti-symétrique, de telle façon que le côté gauche de (13) sera anti-symétrique et que seules les fonctions anti-symétriques seront requises pour son développement. Par conséquent, les fonctions propres symétriques seules donnent une solution complète du problème. La théorie à présent est incapable de décider quelle solution est la solution correcte. Nous sommes capables d'obtenir des solutions du problème qui utilisent des fonctions propres en nombre moindre que le nombre total possible de fonctions propres, au détriment d'être capable de représenter seulement les fonctions symétriques des deux électrons par des matrices.

Ces résultats peuvent évidemment être développés pour n'importe quel nombre d'électrons. Pour r électrons n'interagissant pas, de coordonnées $x_1, y_1, z_1, \dots, x_r, y_r, z_r$, les fonctions propres symétriques sont

$$(14) \quad \sum_{a_1, \dots, a_r} \psi_{n_1}(\alpha_1) \psi_{n_2}(\alpha_2) \dots \psi_{n_r}(\alpha_r),$$

où $\alpha_1, \alpha_2, \dots, \alpha_r$ est n'importe quelle permutation des entiers $1, 2, \dots, r$, alors que celles qui sont anti-symétriques peuvent être écrites sous la forme du déterminant suivant

$$(15) \quad \begin{vmatrix} \psi_{n_1}(1), & \psi_{n_1}(2), & \dots & \psi_{n_1}(r) \\ \psi_{n_2}(1), & \psi_{n_2}(2), & \dots & \psi_{n_2}(r) \\ \dots & \dots & \dots & \dots \\ \psi_{n_r}(1), & \psi_{n_r}(2), & \dots & \psi_{n_r}(r) \end{vmatrix}$$

S'il y a interaction entre les électrons, il y aura à nouveau des fonctions propres symétriques et anti-symétriques, bien qu'elles ne puissent plus être mises dans ces formes simples. Dans tous les cas, les symétriques seules ou les anti-symétriques seules donnent une solution complète du problème.

Une fonction propre anti-symétrique s'évanouit identiquement quand deux des électrons sont sur la même orbite. Cela signifie que dans la solution du problème avec des fonctions propres anti-symétriques, il peut ne pas y avoir d'état stationnaire avec deux électrons ou plus sur la même orbite, ce qui est juste le principe d'exclusion de Pauli.⁷ La solution avec des fonctions propres symétriques, d'un autre côté, permet à n'importe quel nombre d'électrons d'être sur la même orbite, de telle façon que cette solution ne peut être une solution correcte pour le problème des électrons dans un atome⁸.

4. Théorie du gaz parfait.

Les résultats de la section précédente s'appliquent à n'importe quel système contenant différentes particules semblables, en particulier à un assemblage de molécules de gaz. Il y aura deux solutions du problème, dans l'une desquelles les fonctions propres sont des fonctions symétriques des coor-

7. Pauli, Zeits. f. Phys., vol. 31, p. 765 (1925).

8. Prof. Born m'a informé du fait que Heisenberg a obtenu indépendamment des résultats équivalents à ceux-ci (Ajouté à la preuve) - voir Heisenberg, 6 Zeit. fur Phys., vol. 38, p. 411 (1926).

données de toutes les molécules, et dans l'autre, ce sont des fonctions anti-symétriques.

L'équation d'onde pour une molécule simple de masse au repos m se déplaçant librement dans l'espace est

$$\{p_x^2 + p_y^2 + p_z^2 - W^2/c^2 + m^2c^2\}\psi = 0$$

$$\left\{ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{\partial^2}{c^2 \partial t^2} - \frac{m^2c^2}{h^2} \right\} \psi = 0,$$

et sa solution est de la forme

$$(16) \quad \psi_{\alpha_1\alpha_2\alpha_3} = \exp.i(\alpha_1x + \alpha_2y + \alpha_3z - Et)/h,$$

où $\alpha_1, \alpha_2, \alpha_3$ et E sont des constantes vérifiant

$$\alpha_1^2 + \alpha_2^2 + \alpha_3^2 - E^2/c^2 + m^2c^2 = 0.$$

Les fonctions propres (16) représentent un atome ayant des composantes de moment $\alpha_1, \alpha_2, \alpha_3$ et l'énergie E .

Nous devons maintenant obtenir une restriction sur les fonctions propres possibles due à la présence des murs limites. On suppose habituellement que la fonction propre ou la fonction d'onde associée à une molécule, s'évanouit sur la frontière, mais on devrait s'attendre à être capable de déduire cela, si c'est vrai, de la théorie générale. Nous supposons comme une généralisation naturelle des méthodes de la section précédente qu'il doit y avoir seulement des fonctions propres juste suffisantes pour que l'on puisse représenter par une matrice toute fonction des coordonnées qui a un sens physique. Supposons pour la définition que chaque molécule soit confinée entre deux limites en $x = 0$ et $x = 2\pi$. Alors seules les fonctions de x qui sont définies seulement pour $0 < x < 2\pi$ ont un sens physique et doivent pouvoir être représentées par des matrices (cela nécessitera moins de fonctions propres que si toute fonction de x devait pouvoir être représentée par une matrice.). Ces fonctions $f(x)$ peuvent toujours être développées en séries de Fourier de la forme

$$(17) \quad f(x) = \sum_n a_n e^{inx},$$

où les a_n sont des constantes et les n sont des entiers. Si nous choisissons dans ces fonctions propres (16) celles pour lesquelles α_1/h est un entier, alors $f(x)$ compte le nombre de fois où une fonction propre donnée peut être développée en une série parmi les fonctions propres dont les coefficients sont des fonctions de t seulement, et par conséquent $f(x)$ peut être représenté par une matrice. Par conséquent ces fonctions propres choisies sont suffisantes, et peuvent facilement être vues comme étant juste suffisantes, pour la représentation matricielle de toute fonction de x de la forme (17). Au lieu de choisir ces fonctions propres avec des valeurs entières pour α_1/h , nous pourrions aussi bien prendre celles avec α_1/h égal à un demi entier impair ou plus généralement, avec $\alpha_1/h = n + \varepsilon$, où n est un entier et ε est n'importe quel nombre réel. La théorie est incapable de décider quels sont les choix corrects. Pour les problèmes statistiques, pourtant, ils mènent tous aux mêmes résultats.

Quand y et z sont aussi bornés par $0 < y < 2\pi, 0 < z < 2\pi$, on trouve comme nombre d'ondes associé aux molécules dont les énergies sont comprises entre E et $E + dE$ la valeur

$$\frac{4\pi}{c^3 h^3} (E^2 - m^2c^4)^{\frac{1}{2}} E dE.$$

Cette valeur est en accord avec la supposition habituelle que la fonction d'onde s'évanouit sur la

frontière. Elle se réduit, quand on néglige la mécanique relativiste, à l'expression familière

$$(18) \quad \frac{2\pi}{h^3} (2m)^{\frac{1}{2}} E_1^{\frac{1}{2}} dE_1,$$

où $E_1 = E - mc^2$ est l'énergie cinétique. Pour un volume arbitraire de gaz V , l'expression doit être multipliée par $V/(2\pi)^3$.

Pour passer aux fonctions propres pour l'assemblage de molécules, entre lesquelles on suppose qu'il n'y a pas d'interaction, nous multiplions les fonctions propres pour les molécules séparées, et alors nous prenons soit les fonctions propres symétriques, de la forme (14), soit les anti-symétriques, de la forme (15). Nous devons maintenant faire la nouvelle supposition que tous les états stationnaires de l'assemblage (chacun représenté par une fonction propre) ont la même probabilité a priori. Si maintenant nous adoptons la solution du problème dans lequel interviennent des fonctions propres symétriques, nous pouvons trouver que toutes les valeurs pour le nombre de molécules associées à n'importe quelle onde ont la même probabilité a priori, ce qui donne juste la mécanique statistique d'Einstein-Bose⁹. D'un autre côté, nous pourrions obtenir une mécanique statistique différente, si nous adoptons la solution avec les fonctions propres anti-symétriques, de même que nous pourrions alors avoir soit 0 soit 1 molécule associée à chaque onde. La solution avec les fonctions d'ondes symétriques doit être la solution correcte lorsqu'on l'applique à des quanta de lumière, puisque il est connu que la mécanique statistique d'Einstein-Bose mène à la loi de Planck de la radiation du corps noir. La solution avec les fonctions propres anti-symétriques, pourtant, est probablement celle qui est correcte pour des molécules de gaz, puisqu'on sait qu'elle est correcte pour les électrons dans un atome, et que l'on s'attendrait à ce que les molécules ressemblent à des électrons davantage qu'à des quanta de lumière.

Nous allons maintenant traiter, selon les principes bien connus, l'équation d'état du gaz en supposant que la solution avec les fonctions d'ondes anti-symétriques est celle qui est correcte, de telle façon que pas plus d'une molécule ne peut être associée à chaque onde. Divisons les ondes en un certain nombre d'ensembles de façon à ce que les ondes dans chaque ensemble soient associées avec des molécules ayant sensiblement la même énergie. Dénotons par A_s le nombre d'ondes dans le s -ième ensemble, et par E_s l'énergie cinétique d'une molécule associée à l'un d'entre eux. Alors la probabilité qu'une distribution (ou que le nombre de fonctions propres antisymétriques correspondant aux distributions) dans laquelle N_s molécules sont associées avec des ondes dans le s -ième ensemble est

$$W = \prod_s \frac{A_s!}{N_s!(A_s - N_s)!},$$

ce qui donne pour l'entropie

$$S = k \log W = k \sum_s \{A_s(\log A_s - 1) - N_s(\log N_s - 1) - A_s - N_s\} [\log(A_s - N_s) - 1].$$

Cela doit être maximum, de telle façon que

$$\begin{aligned} 0 = \delta S &= k \sum_s \{-\log N_s + \log(A_s - N_s)\} \delta N_s \\ &= k \sum_s \log(A_s/N_s - 1) \cdot \delta N_s, \end{aligned}$$

pour toutes les variations δN_s qui laissent le nombre total de molécules $N = \sum_s N_s$ et l'énergie

9. Bose, Zeits. f. Phys., vol. 26, p. 178 (1924); Einstein, Sitzungsber. d. Preuss. Ac., p. 261 (1924) et p. 3 (1925).

totale $E = \sum_s E_s N_s$ inchangés, de telle façon que

$$\sum_s \delta N_s = 0, \quad \sum_s E_s \delta N_s = 0.$$

Nous obtenons donc

$$\log(A_s/N_s - 1) = \alpha + \beta E_s,$$

où α et β sont des constantes, ce qui donne

$$(19) \quad N_s = \frac{A_s}{e^{\alpha + \beta E_s}} + 1.$$

En faisant varier l'énergie totale E et en mettant $\delta E/\delta S = T$, la température, nous trouvons facilement que $\beta = 1/kT$, de telle façon que (19) devient

$$N_s = \frac{A_s}{e^{\alpha + E_s/kT} + 1}.$$

Cette formule donne la distribution en énergie des molécules. Dans la théorie d'Einstein-Bose, la formule correspondante est

$$N_s = \frac{A_s}{e^{\alpha + E_s/kT} - 1}.$$

Si le s -ième ensemble d'ondes contient les ondes associées aux molécules dont les énergies sont comprises entre E_s et $E_s + dE_s$, nous obtenons de (18) [où E_s signifie maintenant E_1 de l'équation (18)],

$$A_s = 2(2m)^{\frac{1}{2}} E_s^{\frac{1}{2}} dE_s / (2\pi h)^3,$$

où V est le volume du gaz. Cela donne

$$N = \sum N_s = \frac{2\pi V (2m)^{\frac{1}{2}}}{(2\pi h)^3} \int_0^\infty \frac{E_s^{\frac{1}{2}} dE_s}{e^{\alpha + E_s/kT} + 1}$$

et

$$E = \sum E_s N_s = \frac{2\pi V (2m)^{\frac{1}{2}}}{(2\pi h)^3} \int_0^\infty \frac{E_s^{\frac{3}{2}} dE_s}{e^{\alpha + E_s/kT} + 1}.$$

En éliminant α de ces deux équations et en utilisant la formule $PV = \xi E$, où P est la pression, ce qui est vérifié pour n'importe quelle mécanique statistique, l'équation d'état peut être obtenue.

Le phénomène de saturation de la théorie d'Einstein-Bose n'advient pas dans la théorie présente. On peut montrer que la chaleur spécifique tend facilement vers 0 lorsque $T \rightarrow 0$, plutôt que de d'abord croître jusqu'à ce que le point de saturation soit atteint et ensuite décroître, comme dans la théorie d'Einstein-Bose.

5. Théorie des perturbations arbitraires.

Dans cette section, nous considérerons le problème d'un système atomique sujet à une perturbation de l'extérieur (e.g., un champ électromagnétique incident) qui peut varier avec le temps d'une manière arbitraire. Soit l'équation d'onde du système non perturbé

$$(20) \quad (H - W)\psi = 0$$

dans laquelle H est une fonction des p et des q seulement. Sa solution générale est de la forme

$$(21) \quad \psi = \sum_n c_n \psi_n,$$

où les c_n sont des constantes. Nous supposons que les ψ_n doivent être choisis de telle façon que l'un d'entre eux est associé à chaque état stationnaire de l'atome, et doit être multiplié par les constantes propres pour que les matrices qui représentent des quantités réelles soient hermitiennes.

Maintenant supposons qu'une perturbation est appliquée, débutant au temps $t = 0$. L'équation d'onde pour le système perturbé sera de la forme

$$(22) \quad (H - W + A)\psi = 0,$$

où A est une fonction des p , des q et de t et est réelle. On montrera qu'on peut obtenir une solution de cette équation de la forme

$$(23) \quad \psi = \sum_n a_n \psi_n,$$

où les a_n sont des fonctions de t seulement, qui peut prendre des valeurs arbitraires c_n à l'instant $t = 0$. Nous considérerons que la solution générale (21) de l'équation (20) représente un assemblage d'atomes perturbés dans lequel $|c_n|^2$ est le nombre d'atomes dans le n -ième état, et nous assumerons que (23) représente de la même manière un assemblage des atomes perturbés, $|a_n(t)|^2$ étant le nombre de ceux dans le n -ième état à n'importe quel instant t . On prend $|a_n|^2$ au lieu de n'importe quelle autre fonction de a_n parce que, comme cela sera montré ultérieurement, cela permet au nombre total d'atomes de rester constant.

La condition que ψ , défini par l'équation (23), doit satisfaire l'équation (22) est

$$(24) \quad \begin{aligned} 0 &= \sum_n (H - W + A)a_n \psi_n \\ &= \sum_n a_n (H - W + A)\psi_n - i\hbar \sum_n \dot{a}_n \psi_n, \end{aligned}$$

puisque H et A commutent avec a_n ¹⁰ alors que $W a_n = a_n W = i\hbar \dot{a}_n$ identiquement. Supposons que $A\psi_n$ soit développé sous la forme

$$A\psi_n = \sum_n A_{mn} \psi_m,$$

où les coefficients A_{mn} sont des fonctions de t seulement, et satisfont $A_{mn}^* = A_{nm}$, où l'* dénote l'imaginaire conjugué. L'équation (24) devient maintenant, puisque $(H - W)\psi_n = 0$.

En sortant le coefficient de ψ_m , on trouve

$$(25) \quad i\hbar \dot{a}_m = \sum_n a_n A_{mn},$$

qui est une simple équation différentielle montrant comment les a_m varient avec le temps.

En prenant les imaginaires conjugués, on trouve

$$-i\hbar \dot{a}_m^* = \sum_n a_n^* A_{mn}^* = \sum_n a_n^* A_{nm}.$$

Par conséquent, si $N_m = a_m a_m^*$ est le nombre d'atomes dans le m -ième état, nous avons

$$\begin{aligned} i\hbar \dot{N}_m &= i\hbar (\dot{a}_m a_m^* + \dot{a}_m^* a_m) \\ &= \sum_n (\dot{a}_n A_{mn} a_m^* - a_n A_{nm} \dot{a}_m). \end{aligned}$$

10. L'assertion a commute avec b signifie $ab = ba$ identiquement.

Cela donne

$$ih \sum_m N_m = \sum_{nm} (a_m^* A_{mn} a_n - a_n^* A_{nm} a_m) = 0,$$

comme il le faut.

Si la perturbation consiste en une radiation électromagnétique incidente se déplaçant dans la direction de l'axe des x et polarisée selon un plan avec le vecteur électrique dans la direction de l'axe des y , le terme de perturbation A dans l'Hamiltonien est, en négligeant la mécanique relativiste, $\kappa/c \cdot \dot{\eta}$ ¹¹, où η est la polarisation totale dans la direction de l'axe des y et O, κ, O, O sont les composants du potentiel de la radiation incidente. On peut développer $\eta\psi_n$ et $\dot{\eta}\psi_n$ sous la forme

$$\begin{aligned} \eta\psi_n &= \sum_m \eta_{mn} e^{i(W_m - W_n)t/h\psi_m}, \\ \psi_n &= \sum_m \dot{\eta}_{mn} e^{i(W_m - W_n)t/h\psi_m}, \end{aligned}$$

où les η_{mn} et les $\dot{\eta}_{mn}$ sont constants, et $\dot{\eta}_{mn} = i(W_m - W_n)/h \cdot \eta_{mn}$. Notre A_{mn} précédent est maintenant $\kappa/c \cdot e^{i(W_m - W_n)t/h}$, et l'équation (25) devient

$$(26) \quad ihc \dot{a}_m = \sum_n a_n \kappa \dot{\eta}_{mn} e^{i(W_m - W_n)t/h}.$$

Nous pouvons intégrer cette équation du premier ordre en κ en remplaçant les a_n du côté droit par leurs valeurs c_n à l'instant $t = 0$. Cela donne

$$(27) \quad a_m = c_m + 1/ihc \cdot \sum_n c_n \dot{\eta}_{mn} \int_0^t \kappa(s) e^{i(W_m - W_n)s/h} ds.$$

Pour obtenir une seconde approximation, nous écrivons pour les a_n du côté droit de (26) leurs valeurs fournies par (27). Nous trouvons alors pour les valeurs des a_m à l'instant T ,

$$(28) \quad a_m = c_m + 1/ihc \cdot \sum_n c_n \dot{\eta}_{mn} \int_0^T \kappa(t) e^{i(W_m - W_n)t/h} dt$$

disons, où c'_m et c''_m dénotent les termes du premier et second ordre respectivement.

Cela donne pour le nombre d'atomes dans l'état m à l'instant T

$$N_m = a_m a_m^* = c_m c_m^* + c'_m c_m^* + c_m c_m'^* + c'_m c_m'^* + c'_m c_m'^* + c''_m c_m^* + c_m c_m''^*.$$

Si nous souhaitons obtenir des effets qui sont indépendants des phases initiales des atomes, nous devons substituer $c_m \exp.i\gamma m$ pour c_m et calculer la moyenne sur toutes les valeurs des y_m de 0 à 2π . Cela fait s'évanouir les termes du premier ordre en N_m , notamment, $c'_m c_m^*$, et $c_m c_m'^*$, alors que les termes du second ordre donnent

$$\begin{aligned} &1/h^2 c^2 \cdot \sum_n c_n c_n^* \dot{\eta}_{mn} \dot{\eta}_{mn}^* \int_0^T \kappa(t) e^{i(W_m - W_n)t/h} dt \cdot \int_0^T \kappa(t) e^{-i(W_m - W_n)t/h} dt \\ &- 1/h^2 c^2 \cdot \sum_n c_m c_m^* \dot{\eta}_{nm} \dot{\eta}_{mn}^* \int_0^T \kappa(t) e^{i(W_m - W_n)t/h} dt \cdot \int_0^T \kappa(t) e^{-i(W_n - W_m)s/h} ds \\ &- 1/h^2 c^2 \cdot \sum_n c_m c_m^* \dot{\eta}_{nm} \dot{\eta}_{mn}^* \int_0^T \kappa(t) e^{i(W_m - W_n)t/h} dt \cdot \int_0^T \kappa(t) e^{-i(W_m - W_n)s/h} ds \end{aligned}$$

11. Nous avons négligé un terme dans lequel intervient κ^2 . Cette approximation est légitime, même si nous évaluerons ultérieurement le nombre de transitions qui adviennent à l'instant T comme étant de l'ordre de κ^2 , sous la condition que T est grand comparativement aux périodes de l'atome.

qui se réduit à

$$(29) \quad 1/h^2 c^2 \cdot \sum_n \{ |c_n|^2 - |c_m|^2 \} |\dot{\eta}_{nm}^2|^2 \left| \int_0^T \kappa(t) e^{i(W_m - W_n)t/h} dt \right|^2.$$

Cela donne ΔN_m , l'accroissement du nombre d'atomes dans l'état m de l'instant $t = 0$ à l'instant $t = T$. Le terme dans la sommation qui a le suffixe n doit être vu comme dû aux transitions entre l'état m et l'état n .

Si l'on résoud la radiation de l'instant $t = 0$ à l'instant $t = T$ en ses composantes harmoniques, on trouve pour l'intensité de la fréquence ν par unité de fréquence du domaine la valeur

$$I_\nu = 2\pi\nu^2 c^{-1} \left| \int_0^T \kappa(t) e^{2\pi i \nu t} dt \right|^2.$$

Par conséquent, le terme dans l'expression (29) pour ΔN_m dû aux transitions entre l'état m et l'état n peut s'écrire

$$1/2^2 \nu^2 c \cdot \{ |c_n|^2 - |c_m|^2 \} |I_\nu,$$

où

$$2\pi\nu = (W_m - W_n)/h,$$

ou

$$2\pi/h^2 c \cdot \{ |c_n|^2 - |c_m|^2 \} |\nu_{nm}^2 I_\nu.$$

Si l'on effectue la moyenne selon toutes les directions et états de polarisation de la radiation incidente, cela devient

$$2/3 h^2 c \cdot \{ |c_n|^2 - |c_m|^2 \} |P_{nm}^2 I_\nu,$$

où

$$|P_{nm}|^2 = |\xi_{nm}|^2 + |\eta_{nm}|^2 + |\zeta_{nm}|^2,$$

ξ , η et ζ étant les trois composants de la polarisation totale. Par conséquent, on peut dire que la radiation a causé $2\pi/3 h^2 c \cdot |c_n|^2 |P_{nm}|^2 I_\nu$, transitions de l'état n à l'état m , et $2\pi/3 h^2 c \cdot |c_m|^2 |P_{nm}|^2 I_\nu$, transitions de l'état m à l'état n , le coefficient de probabilité pour chacun des processus étant

$$B_{n \rightarrow m} = B_{m \rightarrow n} = 2\pi/3 h^2 c \cdot |P_{nm}|^2$$

en accord avec la théorie d'Einstein.

La présente théorie prend par conséquent en compte l'absorption et l'émission stimulée de radiation et elle montre que les éléments des matrices représentant la polarisation totale déterminent les probabilités de transition. On ne peut prendre en compte l'émission spontanée sans une théorie plus élaborée faisant intervenir les positions des différents atomes et l'interférence de leurs émissions individuelles, car les effets dépendront du fait que les atomes sont distribués au hasard, ou bien arrangés dans un réseau cristallin, ou tous confinés dans un volume de petite taille comparée à la longueur d'onde. La dernière alternative mentionnée, qui n'est d'aucun intérêt pratique, semble être théoriquement la plus simple.

Il faut observer que nous avons obtenu les résultats simples d'Einstein seulement parce que nous avons calculé la moyenne sur toutes les phases initiales des atomes. L'argument suivant montre, pourtant, que les phases initiales sont d'une importance physique réelle, et qu'en conséquence,

les coefficients d'Einstein ne sont pas adéquats pour décrire le phénomène excepté dans des cas particuliers. Si initialement tous les atomes sont dans l'état normal, alors on voit facilement que l'expression (29) pour ΔN_m est respectée sans le processus de calcul de la moyenne, de telle façon que dans ce cas, les coefficients d'Einstein sont adéquats. Si maintenant nous considérons le cas où les atomes sont initialement dans un état excité, nous pouvons supposer qu'ils ont été amenés à cet état par une radiation incidente sur les atomes avant l'instant $t = 0$. L'effet de la radiation incidente suivante doit alors dépendre des relations de phase entre elle et la radiation incidente précédente, puisqu'une manière correcte de traiter le problème serait de résoudre les deux radiations incidentes dans une seule intégrale de Fourier. Si nous ne souhaitons pas que la radiation qui a eu lieu la première apparaisse explicitement dans le calcul, nous devons supposer qu'elle imprime certaines phases à l'atome qu'elle excite, et que ces phases sont importantes pour déterminer l'effet de la radiation suivante. Il ne devrait donc pas être permis de calculer la moyenne de ces phases, mais on devrait travailler directement à partir de l'équation (28).

