

CAUCHOIS (Yvette), Professeur à la Faculté des Sciences de Paris. — **Les spectres des rayons X et la structure électronique de la matière**, avec une Préface de M. E. MOTT, Professeur à l'Université de Bristol. In-8 (16-25) de VIII-104 pages, avec 43 figures et 3 planches; 1948.

COMPARAT (P.), Ingénieur E. G. L., Docteur ès Sciences, **ECOCHARD (F.)**, Agrégé de l'Université, **LAFOUCRIERE (J.)**, Assistant à la Faculté des Sciences de Lyon, et **THIBAUD (J.)**, Professeur à la Faculté des Sciences de Lyon. — **Uranium 235 et Plutonium 239**. In-8 (16-25), de 112 pages, avec figures; 1948.

DESTOUCHES (J.-L.). — **Corpuscules et systèmes de Corpuscules**. TOME I : *Notions fondamentales*. Préface de M. L. BRILLOUIN, Professeur au Collège de France. In-8 (16-25), de XIV-346 pages, avec 7 figures; 1941.

FREYMANN (René), Maître de Conférences à la Faculté des Sciences de Rennes. — **Spectre infra-rouge et structure moléculaire**. *Etude des forces intramoléculaires et intermoléculaires par la perturbation des fréquences infrarouges ou Raman*. In-8 (16-25) de VI-98 pages, avec 50 figures; 1947.

GALLAIS, Professeur à la Faculté des Sciences de Toulouse, et **DAUDEL**, Secrétaire général du Centre de Chimie théorique en France. — **Mécanique ondulatoire et Chimie**. Deux Conférences faites à la Faculté des Sciences de Toulouse les 26 février et 5 mars 1947. Préface de M. G. DUPOUY, Doyen de la Faculté des Sciences de Toulouse. In-8 (16-25), de 62 pages, avec 33 figures; 1947.

THIBAUD (Jean), Directeur de l'Institut de Physique atomique, Professeur à la Faculté des Sciences de Lyon, **CARTAN (Louis)**, Laboratoire de Physique des rayons X, Paris, et **COMPARAT (Paul)**, Institut de Physique atomique, Lyon. — **Quelques techniques actuelles en Physique nucléaire**. *Méthode de la trochoïde. Electrons positifs. Spectrographie de masse. Isotopes. Compteurs de particules à amplification linéaire. Compteurs de Geiger et Muller*. In-8 (16-25) de VI-276 pages, avec 154 figures et 12 planches; 1928.

DE PHYSIQUE GÉNÉRALE ET MATHÉMATIQUE
ŒUVRES RÉUNIES PAR M. DE LAUNAY

III.

ÉLÉMENTS

DE

THÉORIE DES QUANTITÉS

ET DE

MÉCANIQUE ONDULATOIRE

PAR

Louis de BROGLIE

SECRÉTAIRE PERPÉTUEL DE L'ACADÉMIE DES SCIENCES
PROFESSEUR A LA SORBONNE



PARIS

GAUTHIER-VILLARS, IMPRIMEUR

15, RUE DES LONGÈRES, 15
MONTMARTRE, PARIS

III.

ÉLÉMENTS

DE

THÉORIE DES QUANTA

ET DE

MÉCANIQUE ONDULATOIRE

OUVRAGES DU MÊME AUTEUR

Mécanique ondulatoire des systèmes de corpuscules. In-8 (16-25), 164 pages; 1950.

Problèmes de propagation des ondes électromagnétiques. In-8 (16-25), vi-114 pages, 14 figures; nouvelle édition, 1950.

Théorie générale des particules à spin. Méthode de fusion
In-8 (16-25), 210 pages; 1943.

Mécanique ondulatoire du photon et théorie quantique des champs. In-8 (16-25), vi-208 pages; 1949.

La théorie des particules à spin 1/2 (Électrons de Dirac). In-8 (16-25), 164 pages; 1951.

OUVRAGE DE MM. M. ET L. DE BROGLIE

Introduction à la Physique des rayons X et des rayons gamma.
In-8 (16-25), 201 pages, 27 figures et 11 planches hors texte.

TRAITÉ DE PHYSIQUE THÉORIQUE ET DE PHYSIQUE MATHÉMATIQUE

OUVRAGES RÉUNIS PAR JEAN-LOUIS DESTOUCHES

150.145
III.

ÉLÉMENTS

DE

THÉORIE DES QUANTA

ET DE

MÉCANIQUE ONDULATOIRE

PAR

Louis de BROGLIE

SECRÉTAIRE PERPÉTUEL DE L'ACADÉMIE DES SCIENCES
PROFESSEUR A LA SORBONNE



PARIS

GAUTHIER-VILLARS, IMPRIMEUR-ÉDITEUR

LIBRAIRE DU BUREAU DES LONGITUDES, DE L'ÉCOLE POLYTECHNIQUE

Quai des Grands-Augustins, 55

—
1953

Copyright by Gauthier-Villars, 1953.

**Tous droits de traduction, de reproduction et d'adaptation
réservés pour tous pays.**

PRÉFACE.

Ce volume reproduit l'essentiel du cours que j'ai professé à l'École Normale Supérieure depuis 1934. Il contient l'exposé des principales questions qui figurent dans la partie fixe du programme du nouveau certificat de Théories physiques de la Faculté des Sciences de Paris. Seules ont été laissées de côté quelques questions, telles que la Relativité généralisée, la théorie quantique de la Chimie, la théorie quantique des solides et des liquides et l'étude des phénomènes nucléaires sur lesquelles le programme demande seulement aux candidats de posséder quelques notions, ainsi que les preuves expérimentales de l'existence des atomes, des molécules et des électrons qui sont aujourd'hui connues de tous les étudiants en sciences.

Nous n'avons pas fait une étude détaillée approfondie des théories de la nouvelle Physique quantique car, le présent Ouvrage étant destiné aux candidats au certificat de Théories physiques et à ceux qui veulent s'initier à la Mécanique ondulatoire, il doit partir des connaissances que tout étudiant qui a suivi les cours de Physique générale doit posséder et montrer comment et pourquoi se sont introduites les idées fondamentales de la Physique contemporaine. C'est seulement ainsi que l'on peut être amené à bien comprendre ces idées très subtiles qui sont souvent encore mal interprétées. C'est seulement ensuite que l'étudiant avancé pourra entreprendre avec fruit l'étude plus approfondie des prolongements actuels de la Physique quantique et, en particulier, des théories encore très mouvantes et très mal assurées qui s'efforcent de rendre

compte de la nature des champs, de la structure des particules de l'échelle atomique et des interactions à l'intérieur des noyaux.

Nous avons fait suivre chaque chapitre d'une courte bibliographie où nous avons fait figurer principalement des Ouvrages écrits ou traduits en français. Ces bibliographies sont évidemment très incomplètes : elles n'ont d'ailleurs pour but que de permettre au lecteur d'approfondir certains points qui n'ont pu qu'être effleurés dans le texte.

LOUIS DE BROGLIE.

PROGRAMME
DU CERTIFICAT D'ÉTUDES SUPÉRIEURES
DE THÉORIES PHYSIQUES

PARTIE FIXE.

- Électromagnétisme, équations de Maxwell; théorie de Lorentz.
- Théorie de la Relativité restreinte; notions sur les théories de la Relativité généralisée.
- Notions de thermodynamique. Mécanique statistique classique.
- Structure discontinue de la matière, molécules, atomes, électrons.
- Base de la théorie des quanta; corps noir, loi de Planck; émission des raies spectrales, atome de Bohr, théorie de Sommerfeld.
- Principe de correspondance.
- Structure de la lumière, ondes et photons.
- Mécanique ondulatoire; principes fondamentaux, relations d'incertitude, vérifications expérimentales, optique électronique.
- Mécanique ondulatoire des systèmes, théorèmes fondamentaux; méthodes d'intégration, notions sur les méthodes d'approximation et de perturbation.
- Spin, principe de Pauli, symétrie et antisymétrie.
- Applications à la théorie des atomes et des molécules. Atome d'hélium, molécule d'hydrogène; notions sur la théorie quantique de la Chimie.
- Statistique quantique, théorie des gaz, des solides, des métaux, notions sur les liquides.
- Notions sur la Mécanique ondulatoire relativiste; électrons de Dirac, nucléons, photon.
- Théorie des phénomènes nucléaires.

VIII PROGRAMME DU CERTIFICAT D'ÉTUDES SUPÉRIEURES DE THÉORIES PHYSIQUES.

PARTIE MOBILE.

Cours des Professeurs et Maîtres de conférences de Théories physiques et de Relativité et Quanta pendant le semestre qui a précédé la session d'examen.

ÉPREUVE PRATIQUE.

Calculs analytiques et numériques se rattachant au programme ci-dessus.

ÉLÉMENTS

DE

THÉORIE DES QUANTA

ET DE

MÉCANIQUE ONDULATOIRE

CHAPITRE I.

RÉSUMÉ DE LA THÉORIE DE MAXWELL ET DE LA THÉORIE DES ÉLECTRONS.

1. Grandes lignes de la Théorie de Maxwell. — La théorie électromagnétique élaborée par le savant anglais John Clerk Maxwell vers 1870 repose sur un groupe d'équations fondamentales à l'aide desquelles Maxwell a représenté l'ensemble des faits expérimentaux macroscopiques concernant l'Électricité et le Magnétisme qui étaient connus de son temps.

Pour décrire le champ électromagnétique dans les phénomènes à grande échelle, il est nécessaire d'introduire quatre vecteurs définis en chaque point de l'espace. Ce sont, d'une part, les vecteurs champ électrique \mathbf{h} et champ magnétique \mathbf{H} ; d'autre part, les vecteurs induction électrique \mathbf{b} et induction magnétique \mathbf{B} . Dans le vide et dans les milieux qui, au point de vue électromagnétique, se comportent à peu près comme le vide (l'air par exemple), il est permis de confondre chaque induction avec le champ correspondant en posant $\mathbf{b} = \mathbf{h}$ et $\mathbf{B} = \mathbf{H}$. Mais dans la plupart des milieux matériels, il y a lieu de distinguer les inductions des champs. Tous les traités de Physique exposent la manière de définir les champs et les inductions au moyen d'une sorte d'expérience idéale en distinguant la forme allongée ou aplatie de la cavité creusée dans un corps matériel et à l'intérieur de laquelle on mesure l'action subie par une charge électrique ou par un pôle magnétique. La différence

entre l'induction et le champ provient de la manière dont le milieu réagit *en se polarisant* sous l'influence du champ qui lui est imposé.

Pour décrire les charges et les courants, il faut introduire la densité de charge électrique δ en chaque point d'un milieu matériel et la densité de courant électrique \mathbf{i} qui représente en chaque point la charge qui passe par unité de temps à travers une surface unité placée perpendiculairement à la direction du mouvement local de l'électricité.

Les équations de Maxwell expriment les relations qui lient les champs, les inductions, les charges et les courants. Elles ont la forme suivante :

$$(1) \quad \text{rot } \mathbf{h} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t},$$

$$(2) \quad \text{rot } \mathbf{H} = \frac{1}{c} \frac{\partial \mathbf{b}}{\partial t} + 4\pi \frac{\mathbf{i}}{c},$$

$$(3) \quad \text{div } \mathbf{B} = 0,$$

$$(4) \quad \text{div } \mathbf{b} = 4\pi\delta,$$

c est ici le rapport entre les unités de charge dans les systèmes d'unités électromagnétiques et électrostatiques. Les équations (1) à (4) sont écrites dans le système d'unités mixtes de Lorentz qui exprime le champ et l'induction électriques ainsi que les densités de charge et de courant en u. e. s., le champ et l'induction magnétiques en u. e. m.

L'équation (1) traduit la loi de l'induction de Faraday. L'équation (4) est fournie par la loi de Coulomb et le théorème de Gauss. L'équation (2) exprime la liaison entre le champ magnétique et les courants : les faits expérimentaux connus du temps de Maxwell ne permettaient d'écrire que la relation de circuitation d'Ampère

$$\text{rot } \mathbf{H} = 4\pi \frac{\mathbf{i}}{c}$$

qui traduit l'ensemble des lois de Biot et Savart et de Laplace; l'idée géniale de Maxwell fut d'ajouter le terme $\frac{1}{c} \frac{\partial \mathbf{b}}{\partial t}$ au second membre de cette équation, terme représentant une nouvelle sorte de courant, le *courant de déplacement*, qui existe dans les milieux diélectriques et même dans le vide et grâce auquel il y a toujours, même dans un circuit coupé par un condensateur (circuit ouvert), conservation du courant total. Ce fut bien là une idée géniale, car c'est elle qui a permis le développement de la théorie électromagnétique de la lumière dont la découverte par Hertz des ondes qui portent son nom est venue, vingt ans plus tard, apporter une éclatante confirmation.

Enfin l'équation (3) exprime qu'il n'y a pas de magnétisme vrai, qu'il est impossible d'isoler un pôle d'aimant.

Les équations de Maxwell ne permettent de définir l'évolution du champ électromagnétique que si l'on admet certaines relations entre les inductions et les champs. Ces relations dépendent naturellement des propriétés des milieux envisagés. Dans un grand nombre de milieux isotropes, on peut se contenter des relations vectorielles linéaires suivantes :

$$(5) \quad \mathbf{b} = k \mathbf{h}, \quad \mathbf{B} = \mu \mathbf{H},$$

k et μ étant des constantes, c'est-à-dire des quantités indépendantes des champs, que l'on nomme respectivement « la constante diélectrique » et la « perméabilité magnétique ». Bien entendu, dans les milieux non homogènes, k et μ sont des fonctions du point considéré : ces quantités peuvent même être variables avec le temps si le milieu n'est pas dans un état stationnaire. Pour le vide et approximativement pour certains milieux comme l'air, il est permis de poser les constantes k et μ égales à 1.

Les équations de Maxwell ont comme conséquence, quelles que soient les relations entre les inductions et les champs, que l'on a

$$\frac{\partial \delta}{\partial t} + \operatorname{div} \mathbf{i} = 0.$$

Cette équation, du type de l'équation hydrodynamique de continuité, exprime la conservation de l'électricité. C'est pour obtenir cette relation de conservation que Maxwell a introduit le terme « courant de déplacement » dans ces équations.

Il est également utile de rappeler que le courant dans les conducteurs est relié au champ électrique par une relation qui, dans beaucoup de cas, peut se mettre sous la forme simple

$$\mathbf{i} = \sigma \mathbf{h} \quad (\text{loi d'Ohm}),$$

σ étant la conductivité du milieu envisagé.

Terminons ce rapide coup d'œil sur la théorie de Maxwell par des considérations énergétiques.

Soit un volume \mathcal{V} d'un seul tenant limité par une surface S . Évaluons le travail effectué par le champ électromagnétique pendant un temps dt sur les charges contenues dans \mathcal{V} : on voit aisément qu'il est égal à $\int_{\mathcal{V}} (\mathbf{i} \cdot \mathbf{h}) d\tau dt$. Si alors on calcule la valeur de ce travail à l'aide des

équations de Maxwell en admettant les relations linéaires (5) entre champs et inductions, on voit que tout se passe comme si, dans le champ électromagnétique, l'énergie se trouvait localisée avec la densité spatiale

$$(6) \quad w = \frac{1}{8\pi} [kh^2 + \mu \mathbf{H}^2]$$

et de plus que le vecteur

$$(7) \quad \mathbf{S} = \frac{c}{4\pi} [\mathbf{h} \times \mathbf{H}]$$

dit « vecteur radiant de Poynting » représente en chaque point le flux de l'énergie électromagnétique par unité de surface. Max Abraham et Henri Poincaré ont montré que le vecteur $\frac{1}{c^2} \mathbf{S}$ représente la quantité de mouvement du champ électromagnétique par unité de volume.

2. Théorie électromagnétique de la Lumière. — Par définition, la constante c figurant dans les équations de Maxwell est égale au rapport de l'unité u. e. m. à l'unité u. e. s. de charge électrique. L'une des conséquences remarquables de la théorie de Maxwell est que c doit aussi être égale à la vitesse de propagation de la lumière dans le vide.

Maxwell a eu, en effet, l'idée admirable que l'on pouvait regarder la lumière comme une perturbation électromagnétique et il a ainsi réuni en une vaste synthèse les domaines jusqu'alors totalement séparés de l'Optique et de l'Électromagnétisme. Pour voir comment les ondes électromagnétiques se propagent dans le vide ou dans les milieux électromagnétiquement équivalents au vide, nous remarquerons que si l'on y fait $\mathbf{i} = \delta = 0$, $\mathbf{b} = \mathbf{h}$ et $\mathbf{B} = \mathbf{H}$, les équations de Maxwell admettent des solutions de la forme

$$\begin{aligned} \mathbf{h} &= \mathbf{h}_0 \cos 2\pi\nu \left[t - \frac{\alpha x + \beta y + \gamma z}{c} \right], \\ \mathbf{H} &= \mathbf{H}_0 \cos 2\pi\nu \left[t - \frac{\alpha x + \beta y + \gamma z}{c} \right], \end{aligned}$$

avec

$$|\mathbf{h}_0| = |\mathbf{H}_0|, \quad \alpha^2 + \beta^2 + \gamma^2 = 1.$$

Les vecteurs \mathbf{h} et \mathbf{H} sont d'ailleurs perpendiculaires entre eux et perpendiculaires à la direction de cosinus directeurs α, β, γ . Une telle solution représente une onde électromagnétique plane et monochromatique se propageant dans la direction de cosinus directeurs α, β, γ . D'après Maxwell, la lumière est constituée par des ondes de ce type appartenant à un certain domaine de longueurs d'ondes. On voit donc

que dans le vide la lumière, et plus généralement toutes les ondes électromagnétiques, se propagent avec la vitesse c . D'ailleurs dans le vide les équations de Maxwell entraînent comme conséquence les relations $\square \mathbf{h} = 0$ et $\square \mathbf{H} = 0$, avec $\square = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta$, équations qui expriment que dans le vide les champs électromagnétiques se propagent toujours avec la vitesse c . L'expérience a, en effet, montré que la vitesse de propagation des ondes électromagnétiques et lumineuses est numériquement égale à c , ce a qui apporté une belle confirmation des idées de Maxwell.

Dans un milieu matériel isolant ($\mathbf{i} = \mathbf{d} = 0$), où les relations (5) entre champs et inductions sont valables, on trouve des solutions de la forme

$$\mathbf{h} = \mathbf{h}_0 \cos 2\pi\nu \left[t - \frac{\alpha x + \beta y + \gamma z}{V} \right],$$

$$\mathbf{H} = \mathbf{H}_0 \cos 2\pi\nu \left[t - \frac{\alpha x + \beta y + \gamma z}{V} \right],$$

avec

$$V = \frac{c}{\sqrt{k\mu}}; \quad \mu H_0^2 = kh_0^2.$$

Un grand nombre de milieux matériels sont sensiblement non magnétiques, c'est-à-dire qu'on peut y poser sensiblement $\mu = 1$. On a alors simplement $V = \frac{c}{\sqrt{k}}$ et, comme par définition, l'indice de réfraction d'un corps est $n = \frac{c}{V}$, on trouve la célèbre relation de Maxwell

$$k = n^2$$

qui se vérifie bien quand on a soin de mesurer k et n dans des conditions comparables.

Si l'on applique à une onde plane monochromatique dans un milieu non dispersif de constante diélectrique k les expressions (6) et (7) de la densité et du flux de l'énergie, on trouve en moyenne dans le temps

$$\bar{w} = \frac{1}{8\pi} kh_0^2, \quad \bar{\mathbf{S}} = \bar{w} V \mathbf{u},$$

\mathbf{u} étant le vecteur unité porté dans la direction de propagation α, β, γ . Dans les milieux non dispersifs, où k et V ne sont pas fonction de la fréquence, l'énergie électromagnétique se propage donc avec la vitesse V et le vecteur \mathbf{S} donne bien le flux de l'énergie. Dans les

milieux dispersifs, la question est plus compliquée parce qu'il y a localisation partielle de l'énergie dans le milieu même ⁽¹⁾.

On peut encore remarquer que dans le vide la quantité de mouvement transportée par l'onde plane est par unité de volume, $\frac{\bar{S}}{c^2} = \frac{\bar{w}}{c}$ en moyenne.

3. Théorie des électrons (H. A. Lorentz). — Des expériences de plus en plus nombreuses et précises ont suggéré aux physiciens vers 1880-1890 l'idée d'une structure corpusculaire de l'électricité. A la suite de la découverte de l'électron, on a été amené à admettre que la matière contient un nombre énorme de petites particules électrisées. Quand la matière est électriquement neutre, cela signifie donc qu'elle contient autant de particules positives que de particules négatives; quand elle est chargée, cela signifie qu'il y a excès de particules d'un certain signe; quand elle est le siège d'un courant, cela veut dire qu'il y a un déplacement d'ensemble des particules d'un certain signe par rapport à celles de l'autre signe, etc.

La conception d'une structure discontinue de l'électricité a conduit le physicien hollandais H. A. Lorentz vers 1895 à développer la théorie électromagnétique sous une forme nouvelle qu'on a appelée « Théorie des Électrons ». Au lieu de se contenter des champs électriques et magnétiques observables à grande échelle, Lorentz cherche à définir les champs tels qu'ils doivent exister dans la matière, soit entre les particules électrisées, soit même à l'intérieur de ces particules. Il prétend ainsi donner une description du champ électromagnétique plus détaillée, plus fine, que celle qui est fournie par la théorie de Maxwell. Il décrit donc le champ électromagnétique *fin* par deux champs \mathbf{h}' et \mathbf{H}' qui sont les champs microscopiques : il n'y a plus lieu de définir des inductions, car les inductions sont seulement des apparences macroscopiques.

Pour préciser la forme des champs microscopiques, Lorentz a admis que ces champs obéissent à des équations de même forme que celles de Maxwell, et cela, même à l'intérieur des électrons. Il écrit donc

$$(1') \quad \text{rot } \mathbf{h}' = -\frac{1}{c} \frac{\partial \mathbf{H}'}{\partial t},$$

$$(2') \quad \text{rot } \mathbf{H}' = \frac{1}{c} \frac{\partial \mathbf{h}'}{\partial t} + 4\pi\rho \frac{\mathbf{v}}{c},$$

$$(3') \quad \text{div } \mathbf{H}' = 0,$$

$$(4') \quad \text{div } \mathbf{h}' = 4\pi\rho.$$

(1) LÉON BRILLOUIN, *Conférence-rapport sur la Théorie des Quanta*, Presses universitaires, chapitre I.

Ici ρ est la densité microscopique de l'électricité qui est nulle entre les électrons et a des valeurs non nulles à l'intérieur de ceux-ci, par opposition à la densité macroscopique δ qui est définie par l'excès de particules d'un certain signe contenues dans un élément de volume très petit à notre échelle, mais renfermant néanmoins un très grand nombre de particules. \mathbf{v} est la vitesse locale de l'électricité en chaque point à l'intérieur des charges de sorte que $\rho \mathbf{v}$ est la densité de courant microscopique remplaçant ici la densité macroscopique \mathbf{i} .

Quand on a affaire à un nombre immense de charges électriques élémentaires en mouvement, on ne peut observer que les valeurs moyennes des champs : ces valeurs moyennes diffèrent d'ailleurs suivant la méthode de mesure (cavité plate ou cavité allongée), ce qui permet de définir les champs et *les inductions* figurant dans les équations de Maxwell. On peut démontrer ⁽¹⁾ que, si les équations de Lorentz sont exactes microscopiquement, celles de Maxwell s'en déduisent pour les grandeurs macroscopiques par le jeu des moyennes et cela était évidemment nécessaire pour que les équations de la théorie des électrons soient acceptables.

En plus des équations écrites plus haut, la théorie des Électrons admet une équation complémentaire appelée « équation de la force de Lorentz » qui exprime la force « microscopique » agissant sur un élément de charge $\rho d\tau$ animé de la vitesse \mathbf{v} dans un champ électromagnétique défini par les vecteurs \mathbf{h}' et \mathbf{H}' . Cette force est donnée par

$$\mathbf{f} d\tau = \rho d\tau \left[\mathbf{h}' + \frac{1}{c} [\mathbf{v} \times \mathbf{H}'] \right].$$

Cette expression de la force de Lorentz est choisie, elle aussi, par une extrapolation du macroscopique au microscopique. Elle est telle qu'en repassant aux phénomènes à grande échelle, on retrouve la définition usuelle de la force électrique et la loi d'action des champs magnétiques sur l'électricité en mouvement (loi de Laplace).

En combinant les équations de Lorentz, on trouve facilement qu'en dehors des électrons ($\rho = 0$), les champs \mathbf{h}' et \mathbf{H}' se propagent suivant les équations

$$\square \mathbf{h}' = 0, \quad \square \mathbf{H}' = 0$$

qui expriment que ces champs fins se propagent dans le vide et aussi

⁽¹⁾ Voir, par exemple, le livre de R. BECKER (bibliographie [5], p. 121 et suiv.)

entre les électrons de la matière avec la vitesse c . On retrouve aussi la conservation de l'électricité sous la forme

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) = 0$$

et l'on obtient les expressions *microscopiques* de la densité et du flux de l'énergie sous la forme

$$w = \frac{1}{8\pi} (\mathbf{h}'^2 + \mathbf{H}'^2), \quad \mathbf{s} = \frac{c}{4\pi} [\mathbf{h}' \times \mathbf{H}'].$$

Nous devons remarquer ici que la théorie des Électrons n'est jamais parvenue à donner un modèle satisfaisant de l'électron lui-même. Formé d'électricité d'un seul signe, l'électron devrait faire explosion. Henri Poincaré a démontré que pour assurer sa stabilité il fallait introduire une pression, la *pression de Poincaré*, dont la nature est restée inconnue. Il est certain aujourd'hui que le problème de l'existence des particules élémentaires est d'une nature très profonde et doit faire intervenir les quanta, mais sa solution ne paraît pas encore prochaine.

4. Les potentiels électromagnétiques. Formule des potentiels retardés. — Aussi bien dans la théorie de Maxwell que dans celle de Lorentz, il est très commode de faire intervenir des grandeurs nommées « potentiels ». Plaçons-nous dans le cadre de la théorie des Électrons. Le champ électromagnétique γ est défini par les six grandeurs (que nous écrivons en supprimant les accents) $h_x, h_y, h_z, H_x, H_y, H_z$. Au lieu de considérer ces six grandeurs, il est possible de considérer seulement quatre grandeurs, les « potentiels électromagnétiques » dont l'une est une grandeur scalaire, le potentiel scalaire V , et dont les autres sont les composantes d'une grandeur vectorielle, le potentiel vecteur \mathbf{A} . La liaison entre les champs et les potentiels est exprimée par les relations de définition

$$(8) \quad \mathbf{H} = \operatorname{rot} \mathbf{A}, \quad \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \operatorname{grad} V.$$

Il est évident que ces définitions assurent l'exactitude des deux équations de Lorentz « sans second membre »

$$\operatorname{div} \mathbf{H} = 0, \quad \operatorname{rot} \mathbf{h} = -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t}.$$

Les potentiels seront déterminés en fonction des charges et des courants

en introduisant les définitions (8) dans les équations avec second membre

$$\operatorname{rot} \mathbf{h} = 4\pi\rho, \quad \operatorname{rot} \mathbf{H} = \frac{1}{c} \frac{d\mathbf{h}}{dt} + 4\pi\rho \frac{\mathbf{v}}{c}.$$

En se servant des formules suivantes de calcul vectoriel :

$$\operatorname{div} \operatorname{grad} V = \Delta V, \quad \operatorname{rot} \operatorname{rot} \mathbf{A} = \operatorname{grad} \operatorname{div} \mathbf{A} - \Delta \mathbf{A},$$

on obtient ainsi les équations

$$(9) \quad \left\{ \begin{array}{l} -\Delta V - \frac{1}{c} \frac{d}{dt} (\operatorname{div} \mathbf{A}) = 4\pi\rho, \\ \frac{1}{c^2} \frac{d^2 \mathbf{A}}{dt^2} - \Delta \mathbf{A} + \operatorname{grad} \left[\operatorname{div} \mathbf{A} + \frac{1}{c} \frac{dV}{dt} \right] = 4\pi\rho \frac{\mathbf{v}}{c}. \end{array} \right.$$

Lorentz a trouvé une manière ingénieuse de donner à ces équations une forme plus intuitive. Il est parti de la remarque que, si l'on considère les champs comme étant les véritables grandeurs physiques, les potentiels sont de simples artifices mathématiques permettant de calculer les champs. Dès lors, on est libre de choisir pour les potentiels les fonctions que l'on préfère à condition que les champs s'en déduisent par les formules (8). Il en résulte que, si l'on a trouvé des potentiels V et \mathbf{A} répondant à la question, on peut tout aussi bien prendre pour les potentiels les fonctions

$$V' = V + \frac{1}{c} \frac{dF}{dt}, \quad \mathbf{A}' = \mathbf{A} - \operatorname{grad} F,$$

F étant une fonction arbitraire de x, y, z, t , car les valeurs des champs données par (8) restent les mêmes. De là découle que l'on peut imposer aux potentiels de satisfaire à une certaine relation; il suffit pour cela de choisir convenablement la fonction F . Profitant de cette circonstance, Lorentz a imposé aux potentiels la relation suivante :

$$\frac{1}{c} \frac{dV}{dt} + \operatorname{div} \mathbf{A} = 0,$$

dite « relation de Lorentz entre les potentiels ». Il a choisi cette condition parce qu'elle donne aux équations (9) la forme très simple

$$(10) \quad \square V = 4\pi\rho, \quad \square \mathbf{A} = 4\pi\rho \frac{\mathbf{v}}{c}.$$

Dans le vide où $\rho = 0$, on a alors simplement $\square V = 0$ et $\square \mathbf{A} = 0$, équations qui expriment que les potentiels ainsi déterminés se propagent

dans le vide avec la vitesse c tout comme les champs. Ceci donne aux potentiels de Lorentz un sens physique particulier.

Les équations (10) expriment alors la façon dont les potentiels sont créés par la présence et le mouvement des charges et permettent de calculer les potentiels et, par suite, les champs, quand on connaît ρ et \mathbf{v} en fonction de x, y, z, t . La solution du problème est donnée, conformément à un raisonnement classique de Kirchhoff par les formules dites « des potentiels retardés ».

Considérons les potentiels créés en un point P par une distribution de charges en mouvement donnée. Soit $d\tau_M$ l'élément de volume entourant un point M situé à la distance r_{PM} de P.

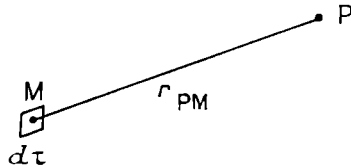


Fig. 1.

Les potentiels retardés sont donnés par les formules

$$(11) \quad V(P, t) = \iiint \frac{[\rho]_{t - \frac{r_{PM}}{c}}}{r} d\tau_M, \quad \mathbf{A}(P, t) = \iiint \frac{1}{c} \frac{[\rho \mathbf{v}]_{t - \frac{r_{PM}}{c}}}{r} d\tau_M.$$

Les quantités $[\rho]$ et $[\rho \mathbf{v}]$ sont les valeurs de la densité de charge et de la densité de courant au point M, non pas à l'instant t , mais à l'instant $t - \frac{r_{PM}}{c}$, c'est-à-dire à l'instant où une perturbation électromagnétique doit partir de M pour parvenir en P à l'instant t : le retardement tient donc compte de la propagation avec une vitesse finie c de toutes les perturbations électromagnétiques.

Notons qu'on pourrait obtenir une autre solution, les « potentiels avancés », en changeant c en $-c$, car l'équation de propagation ne dépend que de c^2 . Cette solution est *en général* considérée comme physiquement inacceptable parce qu'elle ferait dépendre le présent de l'avenir.

§. Calcul du potentiel créé par une distribution statique. — L'emploi des potentiels retardés va nous permettre de calculer les champs qui entourent une distribution d'électricité dont la densité ρ et la vitesse \mathbf{v} sont connues en chaque point.

Nous commencerons par considérer le cas très simple d'une distri-

bution statique ($\nabla = 0$) limitée à un domaine de dimensions finies. Le potentiel vecteur \mathbf{A} étant nul, nous calculerons le potentiel scalaire V en un point P très éloigné de la distribution. Si nous prenons dans la distribution une origine O arbitraire, pour tout point M de la distribution les rapports $\frac{x}{R}$, $\frac{y}{R}$, $\frac{z}{R}$, seront très petits.

La première formule (11) nous donne alors pour V le développement suivant qu'il est très important de connaître :

$$V = \frac{\varepsilon}{R} + \frac{1}{R^2} (\alpha P_x + \beta P_y + \gamma P_z) + \frac{1}{2R^3} \times [(3\alpha^2 - 1)P_{xx} + (3\beta^2 - 1)P_{yy} + (3\gamma^2 - 1)P_{zz} + 6\alpha\beta P_{xy} + 6\beta\gamma P_{yz} + 6\gamma\alpha P_{zx}] + \dots$$

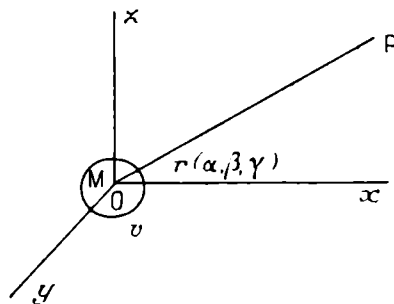


Fig. 2.

avec les définitions

$$\begin{aligned} \varepsilon &= \int_v \rho \, d\tau, & P_x &= \int_v \rho x \, d\tau, & P_y &= \int_v \rho y \, d\tau, & P_z &= \int_v \rho z \, d\tau, \\ P_{xx} &= \int_v \rho x^2 \, d\tau, & P_{yx} &= P_{xy} & & & P_{zx} &= P_{xz} = \int_v \rho xz \, d\tau; \end{aligned}$$

ε est donc la charge totale constante de la distribution. Le vecteur $\vec{\mathcal{P}}$ de composantes $\mathcal{P}_x, \mathcal{P}_y, \mathcal{P}_z$ est son « moment dipolaire »; le tenseur \mathcal{Q} symétrique de composantes $\mathcal{Q}_{xx}, \dots, \mathcal{Q}_{zz}$, est son « moment quadripolaire ». Les termes non écrits du développement feraient intervenir des moments d'ordre supérieur (moments 2^n polaires) : ces termes interviennent assez rarement dans la pratique.

Pour une distribution de charge totale non nulle, le terme $\frac{\varepsilon}{R}$ est le terme prépondérant à grande distance. Pour une distribution dont la charge totale est nulle et le moment dipolaire non nul, les termes en $\vec{\mathcal{P}}$ sont prépondérants : c'est le cas pour un dipôle électrique formé par deux charges ponctuelles très voisines. Si la charge totale et le moment dipolaire sont nuls, ce sont les termes en \mathcal{Q} qui sont prépondérants, etc.

6. Distribution non statique. Rayonnement. — Une distribution statique comme celle que nous venons d'envisager ne rayonne pas d'énergie vers l'infini. Il en est autrement d'une distribution variable qui rayonne de l'énergie vers l'extérieur, tandis que son mouvement s'amortit. Naturellement, comme la charge totale d'une distribution doit rester constante en vertu de la conservation de l'électricité, c'est dans les variations des moments dipolaires, quadrupolaires, etc., qu'il faut chercher l'origine du rayonnement. Dans le cas général où le moment dipolaire n'est pas nul, ce sont les variations de ce moment qui provoquent la partie de beaucoup la plus importante de l'émission, de sorte que pratiquement dans la théorie de l'émission on a guère à se préoccuper des rayonnements quadrupolaires ou d'ordre supérieur. Nous les laisserons ici de côté.

Une distribution quelconque a un moment dipolaire (moment électrique) qui peut toujours se développer en série ou en intégrale de Fourier par rapport au temps, de sorte que l'on a, soit

$$(12) \quad \mathfrak{X}_x = \sum_1^{\infty} \mathfrak{X}_x^{(n)} \cos 2\pi(n\nu t + \varphi_x) = \sum_{-\infty}^{+\infty} p_x^{(n)} e^{2\pi i n \nu t}, \quad \text{avec} \quad p_x^{(n)*} = p_x^{(n)},$$

soit

$$(13) \quad \left\{ \begin{array}{l} \mathfrak{X}_x = \int_0^x \mathfrak{X}_x(\nu) \cos 2\pi[\nu t + \varphi(\nu)] d\nu = \int_{-\infty}^{+\infty} p_x(\nu) e^{2\pi i \nu t} d\nu, \\ \text{avec} \quad p_x^*(\nu) = p_x(\nu), \end{array} \right.$$

Mais il est inutile de traiter le cas général, car on démontre les deux théorèmes suivants :

1° *Le rayonnement émis par la distribution est la somme des rayonnements correspondant aux composantes \mathfrak{X}_x , \mathfrak{X}_y , \mathfrak{X}_z considérées isolément;*

2° *Le rayonnement émis, par exemple, par \mathfrak{X}_x est la somme des rayonnements émis par chaque composante du développement de Fourier de \mathfrak{X} considérée isolément.*

Si donc on sait calculer le rayonnement correspondant à un moment électrique rectiligne et harmonique, on saura résoudre le problème général par simple addition.

Considérons donc le moment électrique

$$\mathfrak{X}_x = \mathfrak{X}_y = 0, \quad \mathfrak{X}_z = A_0 \cos 2\pi(\nu t + \varphi).$$

Le calcul montre qu'en un point P situé à une grande distance R de

la distribution, le champ de rayonnement se compose d'un champ électrique \mathbf{h} tangent au méridien en P et d'un champ magnétique \mathbf{H} tangent au parallèle. La grandeur de ces champs est

$$h = H = \frac{4\pi^2 v^2}{c^2 R} \sin \theta A_0 \cos 2\pi \left[v \left(t - \frac{R}{c} \right) + \varphi \right].$$

L'expression du vecteur de Poynting permet d'évaluer la valeur moyenne par unité de surface du flux de l'énergie électromagnétique en P. On trouve

$$\bar{S}_p = \frac{2\pi^3}{c^3} \frac{v^4}{R^2} \sin^2 \theta A_0^2.$$

En intégrant sur toute la sphère de rayon R, on obtient pour le montant total de l'énergie rayonnée par seconde vers l'infini

$$\Delta W = \frac{16\pi^4 v^4}{3c^3} A_0^2.$$

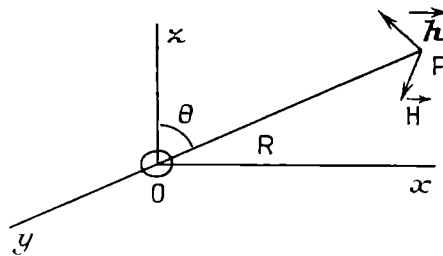


Fig. 3.

Cette expression est indépendante de R comme on devait s'y attendre.

Dans le cas où le moment électrique est représenté par une série de Fourier du type (12), on obtient pour l'énergie rayonnée sous forme de fréquence $n\nu$

$$(14) \quad (\Delta W)_{n\nu} = \frac{64\pi^4 (n\nu)^4}{3c^3} |p_z^{(n)}|^2.$$

Dans le cas où le moment électrique est représenté par une intégrale de Fourier du type (13), le rayonnement dans l'intervalle de fréquence $(\nu, \nu + d\nu)$ est donné par

$$(15) \quad (\Delta W)_{d\nu} = \frac{64\pi^4 \nu^4}{3c^3} |p_z(\nu)|^2 d\nu.$$

Les formules (14) et (15) nous seront utiles dans la suite.

7. Phénomènes de polarisation. Théorie de la dispersion. — La théorie de Lorentz, reprenant d'ailleurs des conceptions qui avaient été intro-

duites dans l'ancienne théorie de l'Électricité, a interprété la différence entre le champ et l'induction macroscopiques comme étant le résultat de l'état de polarisation créé par le champ dans ces milieux. Suivant cette vue, la polarisation d'un milieu matériel consiste en ce que chaque élément de volume $d\tau$ du milieu devient un petit doublet électrique de moment $\mathbf{P} d\tau$, par suite du déplacement sous l'influence du champ des charges électriques qu'il contient. \mathbf{P} est, par définition, le vecteur « polarisation » ou « densité de moment électrique » au point où est situé $d\tau$. On démontre dans tous les traités de Physique qu'un milieu polarisé (supposé électriquement neutre dans sa masse) se comporte comme si chacun des éléments de sa surface extérieure portait une densité de charge superficielle égale à la composante normale P_n du vecteur \mathbf{P} et comme si l'intérieur de son volume était le siège d'une densité de charge électrique δ et d'une densité de courant \mathbf{i} égales à

$$\delta = -\operatorname{div}\mathbf{P}, \quad \mathbf{i} = \frac{d\mathbf{P}}{dt}.$$

Suivant l'ancienne terminologie, ces charges et ces courants représentent de l'électricité libre et non de l'électricité vraie : ceci veut dire qu'à grande échelle ces quantités sont inobservables et ne correspondent pas à une véritable charge globale. Or, l'intervention de ces charges et de ces courants de polarisation conduit à écrire les équations microscopiques de Lorentz dans un milieu diélectrique polarisable sous la forme

$$\begin{aligned} \frac{1}{c} \frac{\partial \mathbf{h}}{\partial t} &= \operatorname{rot} \mathbf{H} - \frac{4\pi}{c} \frac{\partial \mathbf{P}}{\partial t}, & -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} &= \operatorname{rot} \mathbf{h}; \\ \operatorname{div} \mathbf{h} &= -4\pi \operatorname{div} \mathbf{P}, & \operatorname{div} \mathbf{H} &= 0. \end{aligned}$$

Mais, si l'on repasse au point de vue macroscopique de Maxwell, on doit considérer les charges et courants de polarisation comme inobservables directement, ce qui conduit à considérer les termes supplémentaires des équations précédentes comme devant s'incorporer aux premiers membres pour définir une grandeur nouvelle, l'*induction électrique*, et à écrire les équations de Maxwell pour un milieu non magnétisable

$$\begin{aligned} \frac{1}{c} \frac{\partial \mathbf{b}}{\partial t} &= \operatorname{rot} \mathbf{H}, & -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} &= \operatorname{rot} \mathbf{h}; \\ \operatorname{div} \mathbf{b} &= 0, & \operatorname{div} \mathbf{H} &= 0, \end{aligned}$$

avec

$$\mathbf{b} = \mathbf{h} + 4\pi \mathbf{P},$$

relation qui donne l'expression de l'induction électrique en fonction du

champ et de la polarisation. Des considérations analogues pourraient servir à interpréter la différence entre le champ et l'induction magnétiques dans le cas des milieux magnétiquement polarisables.

La polarisation étant produite dans un milieu diélectrique par l'action du champ, on doit avoir une relation vectorielle du type $\mathbf{P} = \mathbf{f}(\mathbf{h})$ et, par suite, $\mathbf{b} = \mathbf{F}(\mathbf{h})$. L'hypothèse la plus simple valable dans beaucoup de milieux isotropes consiste à poser

$$\mathbf{P} = \alpha \mathbf{h} \quad (\alpha \text{ constant}).$$

On a alors

$$\mathbf{b} = (1 + 4\pi\alpha)\mathbf{h} = k\mathbf{h},$$

avec

$$k = 1 + 4\pi\alpha.$$

La constante diélectrique k se trouve ainsi exprimée en fonction de la «susceptibilité» α . Si le corps est non magnétique, la relation de Maxwell donne pour l'indice n ,

$$n^2 = k = 1 + 4\pi\alpha.$$

Considérons le cas simple d'un corps homogène contenant N molécules par unité de volume. On peut, en première approximation, admettre que, sous l'action d'un champ électrique extérieur \mathbf{h} , chaque molécule acquiert un moment électrique $\vec{\mathcal{E}}$ proportionnel au champ

$$\vec{\mathcal{E}} = \alpha' \mathbf{h};$$

α' est la «polarisabilité» du genre de molécule considéré. On a évidemment

$$\mathbf{P} = N \vec{\mathcal{E}} = N \alpha' \mathbf{h},$$

d'où

$$\alpha = N\alpha', \quad n^2 = 1 + 4\pi N\alpha'.$$

La constante α' dépend du corps envisagé, mais est indépendante de sa densité D qui est proportionnelle à N . On a donc

$$n^2 - 1 = CD \quad (\text{loi de Laplace})$$

et pour les corps d'indice voisin de l'unité [$n^2 - 1 \simeq 2(n - 1)$], il vient

$$n - 1 = C'D \quad (\text{loi de Gladstone-Dale}).$$

Cependant, les formules précédentes ne sont pas rigoureuses parce que chaque molécule est soumise, non pas à l'action du champ extérieur \mathbf{h} , mais à celle de la somme du champ \mathbf{h} et des champs créés par

la polarisation des molécules voisines. Un calcul plus exact donne la relation

$$x = \frac{N\alpha'}{1 + \frac{4\pi}{3}N\alpha'}$$

d'où

$$n^2 = 1 + 4\pi x = 1 + \frac{4\pi N\alpha'}{1 + \frac{4\pi}{3}N\alpha'}$$

et l'on en tire une formule obtenue presque simultanément par le hollandais H.-A. Lorentz et par le danois Louis Lorenz

$$\frac{n^2 - 1}{n^2 + 2} = \frac{4\pi N\alpha'}{3} = \text{const. D} \quad (\text{loi de Lorentz-Lorenz}).$$

Cette loi se montre dans beaucoup de cas plus exacte que celles de Laplace et de Gladstone-Dale : elle les admet, d'ailleurs, comme formes approximatives quand n est voisin de 1 ($n^2 + 2 \simeq 3$).

Si la connaissance de la structure de l'atome ou de la molécule permet d'évaluer la polarisabilité α' , les formules précédentes donneront la valeur de l'indice en fonction de la fréquence, c'est-à-dire la formule de dispersion pour la substance envisagée.

La théorie des Électrons était parvenue à trouver (ou plutôt à retrouver, car des théories plus anciennes l'avaient déjà obtenue) une formule qui en gros représente bien les phénomènes de dispersion. Pour cela, il suffit de supposer avec Lorentz que les atomes contiennent des électrons susceptibles d'osciller autour d'une position d'équilibre sous l'action d'une force de rappel proportionnelle à l'élongation. Si l'on suppose que la liaison de chaque électron à sa position d'équilibre est isotrope, chacun d'eux a une fréquence d'oscillation bien définie. Adoptons ce modèle simple et supposons que chaque molécule du milieu réfringent contienne n_1 électrons de fréquence ν_1, \dots, n_i électrons de fréquence ν_i, \dots . On trouve alors, si ν est assez différent de tous les ν_i ,

$$x' = \frac{e^2}{4\pi^2 m} \sum_i \frac{n_i}{\nu_i^2 - \nu^2},$$

e et m étant la charge et la masse de l'électron. Quand on peut se contenter de l'approximation de Laplace, on en déduit

$$(16) \quad n^2 = 1 + \frac{Ne^2}{\pi m} \sum_i \frac{n_i}{\nu_i^2 - \nu^2}.$$

Telle est la formule de dispersion prévue par la théorie des Électrons quand n n'est pas trop différent de 1 et que la fréquence ν n'est pas trop voisine de l'une des valeurs critiques ν_i .

Les fréquences ν_i sont très souvent des fréquences de l'ultraviolet, c'est-à-dire beaucoup plus élevées que la fréquence ν si celle-ci appartient au spectre visible. On a alors approximativement

$$\frac{1}{\nu_i^2 - \nu^2} = \frac{1}{\nu_i^2} \left(1 - \frac{\nu^2}{\nu_i^2} \right)^{-1} \approx \frac{1}{\nu_i^2} + \frac{\nu^2}{\nu_i^4} + \dots$$

et l'on obtient dans ces conditions la formule de Cauchy

$$n^2 = A + \frac{B}{\lambda^2} + \dots$$

généralement valable pour le spectre visible.

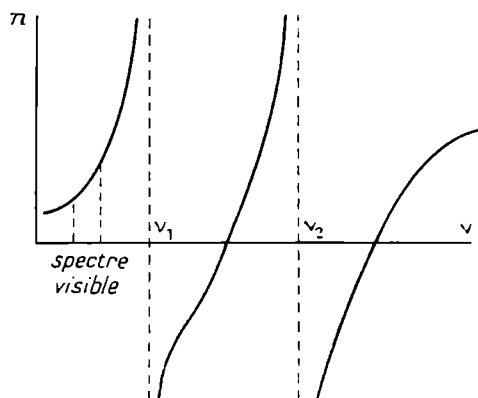


Fig. 4.

L'allure générale de la courbe de dispersion correspondant à la formule (16) est bien connue.

Naturellement, n ne devient pas réellement infini pour les fréquences critiques ν_i . Au voisinage de ces fréquences, la formule ne s'applique plus, car le mouvement des électrons devient trop violent et il faut tenir compte du rayonnement et de l'absorption de la radiation incidente qui en résulte. En tenant compte de cette absorption, on obtient une loi de dispersion valable même dans le voisinage des fréquences critiques : cette loi a une forme acceptable, mais ne rend pas bien compte de la valeur numérique des coefficients d'absorption.

Une autre difficulté de la théorie électronique de la dispersion est relative aux nombres n_i . La comparaison de la formule (16) avec les courbes expérimentales conduit pour les n_i à des nombres généralement

fractionnaires, ce qui est contraire à leur signification physique. Nous verrons que la théorie quantique de la dispersion a pu lever cette difficulté.

8. Succès et échecs de la théorie de Lorentz. — En introduisant systématiquement dans la théorie électromagnétique la notion d'électron, Lorentz et ses continuateurs (Drude, Langevin, etc.) ont pu traiter d'une façon satisfaisante un grand nombre de problèmes d'interaction. Non seulement ils ont obtenu la théorie de la dispersion dont nous venons de parler, mais ils ont pu trouver ou retrouver les lois d'absorption dans les corps éteignants, les lois de la réflexion vitreuse dues à Fresnel, des théories exactes des effets électro- et magnéto-optiques, du diamagnétisme, du paramagnétisme, etc. Le plus grand succès de Lorentz fut la prévision de l'effet Zeeman normal : nous aurons l'occasion d'y revenir. Mais toutes ces fructueuses théories sont venues se heurter à des difficultés fondamentales quand on a voulu analyser de près les processus de l'échelle atomique. La théorie électromagnétique, complétée par la conception classique des électrons, est venue échouer devant les processus atomiques où les quanta interviennent. Nous aurons à revenir longuement sur ce sujet.

BIBLIOGRAPHIE

- [1] H. POINCARÉ, *Électricité et Optique*, Gauthier-Villars, Paris, 1901.
- [2] H.-A. LORENTZ, *Theory of Electrons*, Teubner, Leipzig, 1916.
- [3] P. DRUDE, *Précis d'Optique*, Gauthier-Villars, Paris, 1911.
- [4] M. BORN, *Optik*, Springer, Berlin, 1933.
- [5] R. BECKER, *Théorie des Électrons*, Alcan, Paris, 1938.

CHAPITRE II.

LE PRINCIPE DE RELATIVITÉ.

1. **La transformation de Lorentz.** — Le fait qui a été le point de départ de la théorie de la Relativité, c'est que les phénomènes de propagation de la lumière sont exactement les mêmes pour des observateurs qui sont en mouvement rectiligne et uniforme les uns par rapport aux autres et par rapport à l'ensemble des étoiles fixes (observateurs galiléens). Ce fait est prouvé par des expériences précises dont les plus célèbres sont celles de Michelson. Si l'on admet la réalité de ce fait, on est amené à poser qu'il existe entre les coordonnées d'espace et de temps de deux observateurs galiléens des relations différentes de celles qui étaient admises par la Mécanique classique.

En Mécanique classique, on admet en général que l'équation fondamentale $\mathbf{F} = m\gamma$ et toutes celles qui en dérivent sont valables dans un système de référence attaché à l'ensemble des étoiles fixes, mais on voit aisément qu'elles sont également valables dans tout système de référence en mouvement rectiligne et uniforme par rapport à celui-là. En effet, considérons un observateur lié au système des étoiles fixes : il emploie des coordonnées cartésiennes et le temps t (qui, en Mécanique classique, a un caractère absolu) pour repérer les phénomènes qu'il observe. Un autre observateur en mouvement rectiligne et uniforme par rapport au premier emploiera le temps absolu t , mais ses coordonnées cartésiennes x', y', z' sont reliées à x, y, z par des formules linéaires qui, si les axes cartésiens des deux observateurs sont parallèles et coïncident à l'origine du temps, ont la forme simple

$$x' = x - v_x t, \quad y' = y - v_y t, \quad z' = z - v_z t \quad (t' = t).$$

v_x, v_y, v_z étant les composantes de la vitesse du second observateur dans le système de référence du premier. C'est là « la transformation de

Galilée » qui est admise comme une hypothèse plus ou moins explicite au début de la Mécanique rationnelle. Comme x', y', z' sont fonctions linéaires de x, y, z , les dérivées secondes $\frac{d^2x'}{dt^2}, \dots$ sont respectivement égales à $\frac{d^2x}{dt^2}, \dots$ et comme seules les dérivées secondes interviennent dans les équations fondamentales de la Mécanique classique, il en résulte aisément que celles-ci ont la même forme dans tous les systèmes galiléens.

On peut donc passer des coordonnées employées par un observateur galiléen à celles employées par un autre observateur galiléen à l'aide d'une transformation de Galilée. Les transformations de Galilée forment d'ailleurs un « groupe » en ce sens que deux transformations de Galilée successives sont finalement équivalentes à une seule transformation du même type. La propriété fondamentale des équations de la Dynamique classique de Newton est donc, de notre point de vue actuel, que ces équations sont invariantes par rapport au groupe de Galilée, c'est-à-dire qu'elles restent les mêmes quand on passe d'un observateur galiléen à un autre.

Considérons maintenant la propagation de la lumière dans le vide.

Pour un observateur lié au système des étoiles fixes, nous pouvons admettre que cette propagation se fait d'une façon isotrope suivant l'équation bien connue

$$(1) \quad \left\{ \begin{array}{l} \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} = \Delta \varphi \\ \left(\text{ou } \square \varphi = 0, \text{ avec } \square = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \right) \end{array} \right.$$

avec la vitesse constante c . Le fait fondamental révélé par les expériences de Michelson, c'est que la propagation de la lumière a lieu également d'une façon isotrope et suivant l'équation (1) pour tout observateur en mouvement rectiligne et uniforme par rapport à l'ensemble des étoiles fixes, c'est-à-dire dans tout système de référence galiléen. Et il convient de remarquer dès maintenant qu'étant donnée l'extrême précision des mesures optiques, nous sommes beaucoup plus sûrs de l'exactitude rigoureuse de l'équation de la propagation de la lumière dans le vide que de celle des équations de la Mécanique newtonienne; donc, si nous sommes contraints de modifier soit la théorie de la lumière, soit la Mécanique de Newton, nous devons être prêts à modifier plutôt la seconde que la première à condition, bien entendu, de retomber sur les équations de Newton au degré de précision où elles se vérifient.

Pour que l'équation (1) reste invariante pour tous les observateurs galiléens, il faut que le changement de coordonnées correspondant au passage d'un système galiléen à un autre laisse invariante l'équation des ondes. Or il est facile de vérifier que ce n'est pas le cas pour la transformation de Galilée. Nous devons donc chercher quelle est la transformation de coordonnées qui laisse invariante l'équation des ondes lumineuses.

Pour le voir, supposons que l'observateur galiléen qui emploie les coordonnées x', y', z', t' soit entraîné d'un mouvement rectiligne et uniforme le long de l'axe des z du premier observateur galiléen et supposons de plus que les axes $O'x'y'z'$ soient parallèles aux axes $Oxyz$. On se trouve alors dans le cas particulier de figure ci-contre.

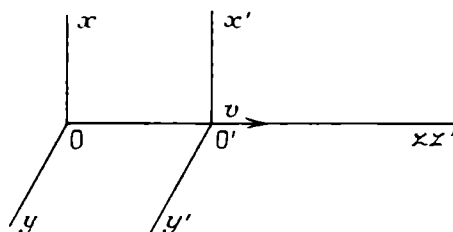


Fig. 5.

Si l'on parvient à trouver les formules qui en ce cas permettent de passer des variables x, y, z aux variables x', y', z' , il sera facile d'en déduire, en faisant subir des rotations aux deux systèmes d'axes autour des origines O et O' , les formules de transformation valables dans le cas général. Nous supposerons d'ailleurs que l'origine du temps est choisie de telle façon que pour $t = 0$, O et O' coïncident. Nous allons donc chercher la transformation des variables x', y', z', t' en les variables x, y, z, t qui laisse invariante l'équation (1).

Des considérations de symétrie simples, mais un peu longues, conduisent alors au résultat suivant : si l'on admet que les formules de transformation cherchées sont linéaires, elles ont nécessairement dans le cas particulier envisagé la forme suivante :

$$(2) \quad x' = x, \quad y' = y, \quad z' = \frac{z - vt}{\sqrt{1 - \beta^2}}, \quad t' = \frac{t - \frac{\beta z}{c}}{\sqrt{1 - \beta^2}}$$

avec par définition $\beta = \frac{v}{c}$. En résolvant (2) par rapport à x, y, z, t , on trouve

$$(2') \quad x = x', \quad y = y', \quad z = \frac{z' + vt'}{\sqrt{1 - \beta^2}}, \quad t = \frac{t' + \frac{\beta z'}{c}}{\sqrt{1 - \beta^2}}.$$

Telle est la fameuse « Transformation de Lorentz » dont le trait essentiel est l'abandon de la notion de temps absolu, car on n'a plus $t' = t$.

La transformation de Lorentz rend non seulement invariante l'équation des ondes lumineuses, mais aussi la quantité $c^2 t^2 - x^2 - y^2 - z^2$, car on vérifie aisément que

$$c^2 t'^2 - x'^2 - y'^2 - z'^2 = c^2 t^2 - x^2 - y^2 - z^2.$$

De plus, si dans l'espace à quatre dimensions formé à l'aide des variables x, y, z, t , on considère un élément de volume $d\tau$, l'expression de $d\tau'$ dans le système des variables primées sera

$$d\tau' = d\tau \frac{D(x, y, z, t)}{D(x', y', z', t')},$$

le symbole $\frac{D(x, y, z, t)}{D(x', y', z', t')}$ désignant le déterminant jacobien des anciennes variables par rapport aux nouvelles. Or, la transformation (2) montre que ce déterminant est égal à 1, d'où $d\tau' = d\tau$. La transformation de Lorentz assure donc aussi l'invariance de l'élément de volume de l'espace-temps.

Dans le cas général où les axes des deux systèmes ne sont pas parallèles, la transformation de Lorentz prend une forme plus générale qu'on écrira plus commodément en posant

$$\begin{aligned} x^1 &= x, & x^2 &= y, & x^3 &= z, & x^4 &= ct; \\ x'^1 &= x', & x'^2 &= y', & x'^3 &= z', & x'^4 &= ct'. \end{aligned}$$

On obtient la relation sous forme condensée

$$x'^i = \sum_k^4 \alpha_k^i x^k,$$

où les seize quantités α_k^i vérifient certaines relations que nous ne pouvons étudier ici. Le déterminant fonctionnel $\frac{D(x', y', z', t')}{D(x, y, z, t)}$ est égal au déterminant des α_k^i et a pour valeur 1, de sorte qu'il y a toujours invariance de l'élément de volume de l'espace-temps. Les transformations de Lorentz forment un groupe comme celles de Galilée.

2. Signification physique de la Transformation de Lorentz. — Il est beaucoup plus difficile qu'on ne le croyait autrefois de définir la simultanéité de deux événements qui n'ont pas lieu au même point de l'espace. C'est en réfléchissant sur ce fait qu'Albert Einstein est parvenu en 1905.

par un effort génial, à dégager la véritable signification de la transformation de Lorentz.

La difficulté signalée tient à ce que, si l'on veut synchroniser les horloges liées à un même système de référence, il faut avoir un moyen de comparer leurs indications et, si les horloges sont distantes l'une de l'autre, il faut pour cela qu'un signal parti de l'emplacement de la première horloge parvienne à l'emplacement de l'autre. Mais nous ne connaissons pas dans la nature de signal allant plus vite qu'un signal électromagnétique ou lumineux, et Einstein a pris comme postulat justifié par ses conséquences qu'aucun transport d'énergie, donc aucun signal, ne peut s'effectuer avec une vitesse supérieure à c : nous verrons que la Dynamique relativiste confirme *a posteriori* cette idée. Il est donc naturel et commode de s'imaginer que la synchronisation des horloges se fait par échange de signaux lumineux. Le résultat de l'opération serait d'ailleurs le même avec un mode de signalisation à propagation moins rapide (signaux sonores, par exemple), mais le raisonnement qui suit serait plus compliqué car il faudrait tenir compte du milieu matériel qui transmet les signaux : l'avantage de raisonner sur les signaux lumineux, c'est que tout se passe, si l'on abandonne avec Einstein l'ancienne conception de l'éther, comme si la lumière se propageait dans l'espace vide.

Pour synchroniser les horloges, dans son système de référence, un observateur galiléen devra donc d'abord mesurer la vitesse de la lumière dans le vide, vitesse par hypothèse supposée isotrope. La connaissance de cette vitesse lui est, en effet, nécessaire pour tenir compte des retards dus à la propagation. Pour mesurer la vitesse de la lumière, il devra, grâce à une réflexion sur un miroir, faire parcourir à un signal lumineux un même trajet aller et retour et noter sur la même horloge placée au point de départ l'heure du départ et celle du retour du signal lumineux. Si la différence de ces heures est égale à τ et si la distance franchie par le signal est $2d$, on aura $c = \frac{2d}{\tau}$. Connaissant ainsi la valeur de la constante c , l'observateur considéré pourra régler toutes les horloges liées à son système de référence de telle façon que, si deux horloges sont distantes de l et si la première émet à l'instant t_1 un signal lumineux, ce signal parvienne à la seconde horloge à l'instant où elle marque le temps $t_2 = t_1 + \frac{l}{c}$. De cette manière sera réalisée une « synchronisation » tout à fait univoque des horloges du système de référence. Ainsi sera bien déterminé le temps t dans ce système et deux événements

qui se produisent en des points P et Q du système de référence au moment où les horloges locales de P et de Q marquent le même temps t peuvent à juste titre être considérés comme « simultanés ». C'est même la seule manière de définir d'une façon précise la simultanéité dans un système de référence galiléen. C'est là un point essentiel mis en lumière par Einstein.

Considérons maintenant deux systèmes de référence galiléens A et B. Un observateur lié à A emploiera des coordonnées cartésiennes x, y, z et un temps t défini comme nous venons de l'expliquer. Un observateur lié à B emploie des coordonnées x', y', z' et un temps t' défini par la même méthode. Nous pouvons alors avec Einstein adopter le Principe général de Relativité dont voici l'énoncé : « L'observateur B peut se considérer comme au repos tout aussi légitimement que l'observateur A et pour lui, quand il emploie les coordonnées x', y', z', t' , qui lui sont propres, toutes les lois de la nature doivent avoir exactement la même forme que pour l'observateur A quand celui-ci utilise ses coordonnées propres x, y, z, t ». Or, nous avons montré que, pour qu'il en soit ainsi en ce qui concerne les phénomènes lumineux, il fallait que x', y', z', t' soient reliées à x, y, z, t par la transformation de Lorentz. Einstein a ainsi dégagé le véritable sens de cette transformation qui est intimement lié à la façon dont on peut définir la simultanéité et la synchronisation des horloges dans chaque système de référence galiléen.

3. La contraction de Lorentz. — Nous allons maintenant pouvoir comprendre ce qu'on appelle « la contraction de Lorentz ». Supposons que l'observateur A ait construit une règle, par exemple en métal, pour s'en servir comme étalon de longueur. Soit l la longueur de cet étalon exprimée à l'aide d'une certaine unité de longueur arbitraire. L'observateur galiléen B qui est en mouvement rectiligne et uniforme par rapport à A peut aussi construire une règle qui sera pour lui identique à ce qu'est la première règle pour l'opérateur A : ceci est une conséquence évidente du principe de Relativité. Supposons (ce qui au fond ne diminue pas la généralité) que nous ayons pris pour axe des z la direction du mouvement de B par rapport à A et que les axes $Oxy z$ et $O'x'y'z'$ soient parallèles. Alors, la transformation de Lorentz sous sa forme (2) sera valable. Supposons maintenant que chacun des deux observateurs pose *sa* règle sur *son* axe des z , l'extrémité de gauche coïncidant avec l'origine des z . Alors, pour A, sa règle couvre l divisions de l'axe des z (supposé gradué) à partir de O et, de même, pour B, sa règle couvre l divisions de l'axe des z' à partir de O' : c'est là

simplement l'expression du fait que, pour chacun des observateurs, sa règle a la longueur l .

Mais demandons-nous quelle longueur la règle de B paraît posséder pour A. Voici le point essentiel de la réponse : la seule manière admissible de définir la longueur que possède pour un observateur une règle mobile ou immobile, c'est de dire que cette longueur est égale à la distance des positions des deux extrémités de la règle *qui sont simultanées pour cet observateur*. Cette définition a bien pour A un sens net, puisque A possède un temps t bien défini et peut, par suite, définir la simultanéité dans son système de référence. Or, pour B, les coordonnées des deux extrémités de sa règle sont

$$x'_1 = y'_1 = z'_1 = 0, \quad x'_2 = y'_2 = 0, \quad z'_2 = l.$$

Pour A, d'après la transformation (2) de Lorentz, les deux extrémités en question auront au temps t les coordonnées

$$\begin{aligned} x_1 = y_1 = 0, \quad z_1 = vt + z'_1 \sqrt{1 - \beta^2} = vt; \\ x_2 = y_2 = 0, \quad z_2 = vt + z'_2 \sqrt{1 - \beta^2} = vt + l \sqrt{1 - \beta^2}. \end{aligned}$$

La longueur de la règle de B évaluée par A par rapport auquel elle est en mouvement sera donc égale à la différence $z_2 - z_1$ pour une même valeur de t , soit

$$z_2 - z_1 = l \sqrt{1 - \beta^2}.$$

Elle est donc inférieure à l , puisque $0 < \beta \leq 1$. La règle est raccourcie dans le rapport $\sqrt{1 - \beta^2}$. Cette règle qui, immobile par rapport à l'observateur A, lui semblerait avoir la longueur l , lui apparaît plus courte parce qu'elle est en mouvement. C'est la contraction de Lorentz.

Si la règle de B n'était pas placée le long de l'axe Oz' , mais avait une direction oblique quelconque, c'est seulement la projection de la règle sur la direction de mouvement qui paraîtrait plus courte à A qu'à B tandis que les projections sur les axes perpendiculaires au mouvement ne subiraient pas de contraction. Il en résulte qu'un corps qui pour B est une sphère apparaît à A comme un ellipsoïde de révolution aplati, l'aplatissement étant mesuré par le facteur $\sqrt{1 - \beta^2}$.

La contraction de Lorentz qui résulte de la troisième équation (2) peut, si l'on admet *a priori* son existence, conduire à l'équation (2). Figurons à nouveau les deux systèmes qui glissent l'un sur l'autre.

Supposons qu'au temps $t = 0$, O' était en O. Alors si l'observateur A fait à l'instant t de son temps propre une observation simultanée de l'abscisse de O' et de celle d'un point P' de $O'z'$, il trouvera ct pour la

première abscisse et une certaine valeur z , pour la seconde, c'est-à-dire qu'à l'instant t un observateur du système A placé devant P' voit ce point passer devant le point d'abscisse z de l'axe Oz . Mais comment l'observateur B pourra-t-il, lui, procéder à la mesure de la longueur $O'P'$? C'est simplement, par exemple, en mettant bout à bout le long de $O'z'$ auquel il est lié des petites réglettes de longueur unité, la première ayant son origine en O' , la dernière son extrémité en P' (en supposant pour simplifier que l'unité de longueur soit telle que $O'P'$ la contienne un nombre entier de fois). Supposons que B ait ainsi disposé ses réglettes : il lui en faut z' pour jalonner $O'P'$ et il écrira $\overline{O'P'} = z'$. Mais que voit l'observateur A quand les z' réglettes sont posées sur $O'z'$? Il voit z' réglettes dont la longueur lui paraît être pour chacune égale à $\sqrt{1 - \beta^2}$.

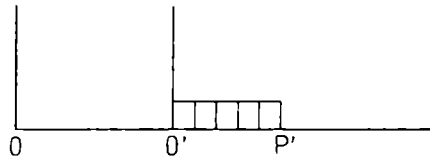


Fig. 6.

puisque nous admettons la contraction de Lorentz. Ces réglettes étant disposées jointivement de O' à P' , A écrit l'équation

$$z - vt = z' \sqrt{1 - \beta^2}$$

et c'est bien la troisième relation (2) de Lorentz.

Pour l'observateur A, B trouve z' au lieu de z parce qu'il emploie une règle unité qui est trop courte. Mais, d'après le principe de Relativité, l'observateur B qui voit A se mouvoir le long de $O'z'$ avec la vitesse $-c$ dira que c'est A qui utilise une règle unité trop courte et cela lui donnera la troisième relation (2') sous la forme

$$z' + vt' = z \sqrt{1 - \beta^2}.$$

Chaque observateur trouve la règle de l'autre plus courte que la sienne. C'est le paradoxe de réciprocité sur lequel nous reviendrons.

4. Le ralentissement des horloges. — La contraction de Lorentz correspond à la troisième relation (2); à la quatrième correspond le « ralentissement des horloges ».

L'observateur lié au système A, qui peut définir une simultanéité dans son système de référence, peut placer en chaque point une horloge cons-

tituée par un système périodique quelconque, par exemple une aiguille tournant d'un mouvement uniforme devant un cadran. On pourra repérer ce mouvement par une variable périodique de la forme $q = q_0 \sin 2\pi\nu t$ qui, dans le cas du cadran à aiguille, sera par exemple l'abscisse de la projection de l'extrémité de l'aiguille sur un certain diamètre du cadran.

Supposons que l'observateur A ait ainsi muni d'une horloge tous les points de son système et qu'il ait réalisé, de la manière que nous avons étudiée, la synchronisation de toutes ces horloges. Supposons aussi que l'observateur B lié à un autre système galiléen ait fait de même, ce qui est possible puisqu'en vertu du principe de Relativité, B peut construire des dispositifs qui sont pour lui identiques à ce que les horloges de A sont pour A. Imaginons maintenant que A observe le mouvement d'une des horloges de B qui pour lui est mobile. Si z_0 est la coordonnée z de cette horloge pour A à l'instant $t = 0$ de son temps, comme elle lui paraît entraînée avec la vitesse v , il lui trouvera comme coordonnée z à l'instant t , $z = z_0 + vt$. Or, cette horloge, ayant été réglée par B, marque le temps t' et son paramètre oscillant est

$$q = q_0 \sin 2\pi\nu' t' = q_0 \sin 2\pi \frac{t'}{T'},$$

où T' est la période de l'horloge pour B. Pour l'observateur A, à l'instant t de son temps propre, l'horloge mobile considérée passe donc au point d'abscisse $z = z_0 + vt$ devant l'une des horloges fixes qui marquent le temps t et, d'après la quatrième relation (2) de Lorentz, le paramètre oscillant de l'horloge s'exprime en fonction de t par

$$\begin{aligned} q &= q_0 \sin 2\pi \frac{t'}{T'} = q_0 \sin 2\pi \frac{t - \frac{\beta}{c}(z_0 + vt)}{\sqrt{1 - \beta^2} T'} \\ &= q_0 \sin 2\pi \left(\frac{t}{T'} \sqrt{1 - \beta^2} - \frac{\beta z_0}{cT' \sqrt{1 - \beta^2}} \right). \end{aligned}$$

La période de l'horloge telle qu'elle est observée par A est donc

$$T = \frac{T'}{\sqrt{1 - \beta^2}} > T'.$$

D'où la conclusion : l'observateur A attribue à l'horloge une période plus longue que T' . Une horloge qui paraîtrait pour l'observateur A battre la seconde (si elle était immobile par rapport à lui) lui paraît aller trop lentement, retarder, si elle est en mouvement. Si les observateurs du système A comparent constamment l'indication de l'horloge mobile avec celles des horloges fixes devant lesquelles elle passe, ils peuvent constater

que l'horloge mobile retarde de plus en plus sur les horloges fixes. Donc le ralentissement relativiste des horloges, tout comme la contraction de Lorentz, est en principe un phénomène réel, observable. Malheureusement, ce phénomène n'est notable que pour des vitesses v voisines de c , ce qui le rend inobservable à grande échelle. On ne peut qu'en constater les conséquences dans les phénomènes de l'échelle atomique.

Pour préciser ce qui précède, nous pouvons nous placer à un autre point de vue en supposant que les observateurs du système A observent simultanément à un même instant t de leur temps propre *toutes* les horloges du système B dont les paramètres oscillants sont $q = q_0 \sin 2\pi\nu't'$, où ν' est la fréquence de l'horloge dans le système B, c'est-à-dire le nombre des oscillations qu'elle effectue par unité du temps t' . Les observateurs A voient donc ces horloges mobiles marquer à l'instant t de leur temps propre l'indication

$$q = q_0 \sin 2\pi \frac{\nu'}{\sqrt{1-\beta^2}} \left(t - \frac{\beta}{c} z \right).$$

Telle est l'indication de l'horloge mobile qui occupe pour A la position z à l'instant t .

Posons

$$\nu = \frac{\nu'}{\sqrt{1-\beta^2}}, \quad V = \frac{c}{\beta} = \frac{c^2}{v}.$$

on peut écrire

$$q = q_0 \sin 2\pi\nu \left(t - \frac{z}{V} \right)$$

et sous cette forme la relation nous indique que pour A tout se passe comme si les valeurs de q se propageaient par ondes planes dans la direction des z avec la fréquence ν et la vitesse de propagation V . C'est là, comme nous le verrons, un résultat qui a joué un grand rôle dans le développement initial de la Mécanique ondulatoire.

Montrons que, si l'on admet le ralentissement des horloges, on peut en tirer la quatrième relation (2) de Lorentz. En effet, pour synchroniser une horloge placée en P' sur son axe $O'z'$, l'observateur B placé en O' enverra, quand son horloge locale marquera un certain temps $t' = 0$ un signal lumineux et un collaborateur placé en P' marquera sur son horloge locale à la réception du signal le temps $\frac{z'}{c}$ si z' est l'abscisse de P' . Supposons que A observe cette expérience de synchronisation et qu'il choisisse comme origine de son temps t l'instant où B envoie son signal. Pour lui, ce signal lumineux chemine le long de $O'z'$ avec la vitesse $c - v$

et, comme $O'P'$ lui paraît avoir d'après la contraction de Lorentz la longueur $z'\sqrt{1-\beta^2}$, il mettra pour franchir la distance $O'P'$ le temps

$$t = \frac{z'\sqrt{1-\beta^2}}{c-v}.$$

Donc, quand le collaborateur de B marquera l'heure $t' = \frac{z'}{c}$ sur son horloge locale en P' , l'horloge du système A en face de laquelle il passe marque le temps $t = \frac{z'\sqrt{1-\beta^2}}{c-v}$ et l'on peut écrire

$$t = \frac{z'}{c} \frac{\sqrt{1-\beta^2}}{1-\beta} = \frac{z'}{c} \frac{1+\beta}{\sqrt{1-\beta^2}}$$

et puisque $t' = \frac{z'}{c}$, on voit qu'au temps t' marqué par l'horloge au point z' correspond dans le temps t l'instant

$$t = \frac{t' + \frac{v}{c} z'}{\sqrt{1-\beta^2}},$$

ce qui est la quatrième relation (2') de Lorentz.

5. Mesure de la vitesse de la lumière par deux observateurs galiléens. — Il est aisé de voir comment le ralentissement des horloges permet à A de s'expliquer que B puisse trouver pour la vitesse de la

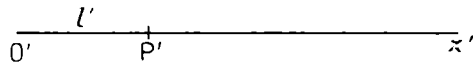


Fig. 7.

lumière dans le vide la même valeur c que lui-même. En effet, pour mesurer la vitesse de la lumière, B doit, nous l'avons dit, faire accomplir par un signal lumineux un même trajet aller et retour. Supposons que l'observateur placé en O' fasse voyager la lumière le long de son axe $O'z'$.

La lumière part de O' , se réfléchit en P' à la distance $\overline{O'P'} = l'$ et revient en O' . Admettant l'isotropie de son espace et constatant que la lumière a mis une durée τ de son temps propre pour aller et revenir, B pose

$$(3) \quad c = \frac{|\overline{O'P'}| + |\overline{P'O'}|}{\tau} = \frac{2l'}{\tau}.$$

Si A regarde B effectuer cette mesure, que dit-il? Il dit : la

distance $|\overline{O'P'}|$ a pour longueur, non pas l' , mais $l'\sqrt{1-\beta^2}$ et la lumière ne chemine pas avec même vitesse par rapport à $O'z'$ à l'aller et au retour, car à l'aller elle se déplace par rapport à $O'z'$ avec la vitesse $c-v$ et au retour avec la vitesse $c+v$. De plus, pour A, l'horloge de B marche trop lentement et la durée de l'expérience est, non pas τ , mais $\frac{\tau}{\sqrt{1-\beta^2}}$. A écrit donc l'équation

$$\frac{\tau}{\sqrt{1-\beta^2}} = \frac{l'\sqrt{1-\beta^2}}{c-v} + \frac{l'\sqrt{1-\beta^2}}{c+v} = \frac{2l'}{c\sqrt{1-\beta^2}}.$$

Il retrouve donc pour c la valeur (3).

Mais, pour A, l'exactitude de la mesure faite par B résulte d'une compensation entre plusieurs erreurs sur les longueurs, les temps et les vitesses. A se dit : « B a bien de la chance. Il raisonne comme s'il était au repos, alors qu'il est en mouvement; il emploie des mètres trop courts et des horloges qui retardent et finalement toutes ces erreurs se compensent et il trouve le résultat exact ».

Voilà comment raisonne A s'il n'est pas relativiste, mais naturellement B raisonnerait de même si c'était A qui faisait une mesure de c . Notons que si B faisait sa mesure en prenant le point de réflexion P' en dehors de l'axe $O'z'$, un calcul un peu plus long conduirait au même résultat.

6. Paradoxe de la réciprocité. — Le paradoxe de la réciprocité consiste en ce que les observateurs A et B s'accusent mutuellement d'erreur et qu'en raison du principe de Relativité ils ont exactement le même droit de le faire. Nous venons d'en voir un exemple, en voici un autre.

A et B ont chacun une règle qui leur paraît avoir la longueur l . Si ces règles sont placées dans le sens du mouvement relatif, chacun des observateurs attribue à la règle de l'autre la longueur réduite $l\sqrt{1-\beta^2}$. Comment est-ce possible? Cela vient de la manière dont chaque observateur évalue la longueur de la règle de l'autre : nous avons vu qu'il ne peut définir cette longueur que comme égale à la distance des deux extrémités de la règle mobile à un même instant de son temps propre. Mais si chaque observateur peut définir la simultanéité dans son système propre, la simultanéité dans les deux systèmes n'est pas la même.

Regardons A mesurer la longueur de la règle de B.

Pour plus de clarté, nous séparons les axes Oz et $O'z'$. Supposons

qu'à l'instant où l'extrémité arrière O' de la règle de B passe devant l'extrémité arrière O de la règle de A, les horloges placées en O et O' marquent la même heure que nous nommerons l'heure o . Le collaborateur de A placé en P voit à cette même heure o passer devant lui l'extrémité avant de la règle de B, et A dira que la règle B a pour longueur $\overline{OP} = l\sqrt{1-\beta^2} < l$ inférieure à la longueur \overline{OL} de sa propre règle. Mais, les horloges en O et en O' marquant la même heure, celles en L'_1 et en P marquent des heures différentes quand elles passent l'une devant l'autre puisque d'après le ralentissement des horloges, l'horloge L'_1 paraît dans le système A prendre du retard. D'après la quatrième relation (2) de Lorentz, on a pour $t = o$

$$t'_z - t'_0 = - \frac{\frac{\beta}{c} z}{\sqrt{1-\beta^2}} < 0.$$

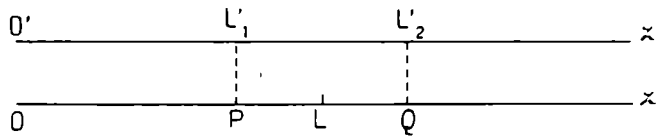


Fig. 8.

Le collaborateur de B placé en L' doit donc, après son passage devant P , attendre encore un temps

$$\Delta t' = \frac{\frac{\beta}{c} z}{\sqrt{1-\beta^2}} = \frac{\beta}{c} l$$

avant d'observer sur son horloge locale le temps $t' = 0$.

Puisque $t = \frac{t' + \frac{\beta}{c} z'}{\sqrt{1-\beta^2}}$, le temps Δt marqué sur l'horloge locale de A située en face de L' quand $t' = 0$ est (puisque $z' = l$)

$$\Delta t = \frac{\frac{\beta}{c} z'}{\sqrt{1-\beta^2}} = \frac{\frac{\beta}{c} l}{\sqrt{1-\beta^2}}$$

Comme, pour A, le point L' a la vitesse βc , ce point sera au temps $o + \Delta t$ en face d'un point Q de Oz tel que, avec les unités de A, on ait

$$\overline{OQ} = z_Q = \overline{OP} + \overline{PQ} = l\sqrt{1-\beta^2} + \beta c \frac{\frac{\beta l}{c}}{\sqrt{1-\beta^2}} = \frac{l}{\sqrt{1-\beta^2}} = \frac{\overline{OL}}{\sqrt{1-\beta^2}}.$$

Pour B, la longueur \overline{OQ} vaut l , puisque les positions simultanées des deux extrémités de sa règle de longueur propre l au temps $t' = 0$ sont en face des points O et Q. Comme on a dans le système A

$$\overline{OL} = \overline{OQ} \sqrt{1 - \beta^2},$$

on voit que pour B, la règle \overline{OL} de A paraît avoir pour longueur $l \sqrt{1 - \beta^2}$.

Ce que nous venons de dire dans le paragraphe précédent et dans celui-ci montre à quels raisonnements profonds et subtils conduit l'analyse de la mesure des longueurs et des durées quand on adopte le point de vue d'Einstein. Malgré les nombreuses critiques que l'on a en vain cherché à leur adresser, ces raisonnements refaits de bien des manières par des théoriciens éminents (Einstein, Lorentz, Langevin, etc.) sont inattaquables si l'on admet l'hypothèse que le mouvement rectiligne et uniforme est sans influence sur les phénomènes physiques et en particulier sur la propagation de la lumière. On ne pourrait ébranler la théorie de la Relativité restreinte dont nous venons de donner une esquisse qu'en mettant en évidence par l'expérience une influence du mouvement rectiligne et uniforme sur les phénomènes optiques ou électromagnétiques. Aucune preuve sérieuse d'une telle influence n'a jusqu'ici été apportée.

7. Formule de composition des vitesses. — Soit un système d'axes galiléens $O_1 x_1 y_1 z_1$ et un autre système d'axes parallèles $O_2 x_2 y_2 z_2$ dont l'origine glisse d'un mouvement rectiligne et uniforme sur l'axe $O_1 z_1$. Nous pouvons toujours nous ramener à ce cas quand nous comparons deux systèmes galiléens. La transformation a alors la forme simple (2).

Considérons un point matériel mobile animé dans le système (2) d'une vitesse \mathbf{v}_2 .

La vitesse \mathbf{v}_1 du point M dans le système (1) a pour composantes

$$\begin{aligned} (v_1)_x &= \frac{dx_1}{dt_1} = \frac{dx_2 \sqrt{1 - \beta^2}}{dt_2 + \frac{\beta}{c} dz_2} = \frac{(v_2)_x \sqrt{1 - \beta^2}}{1 + \frac{\beta}{c} (v_2)_z} = \frac{(v_2)_x \sqrt{1 - \beta^2}}{1 + \frac{1}{c^2} v (v_2)_z}, \\ (v_1)_y &= \frac{dy_1}{dt_1} = \frac{dy_2 \sqrt{1 - \beta^2}}{dt_2 + \frac{\beta}{c} dz_2} = \frac{(v_2)_y \sqrt{1 - \beta^2}}{1 + \frac{\beta}{c} (v_2)_z} = \frac{(v_2)_y \sqrt{1 - \beta^2}}{1 + \frac{1}{c^2} v (v_2)_z}, \\ (v_1)_z &= \frac{dz_1}{dt_1} = \frac{dz_2 + \beta c dt_2}{dt_2 + \frac{\beta}{c} dz_2} = \frac{(v_2)_z + v}{1 + \frac{\beta}{c} (v_2)_z} = \frac{(v_2)_z + v}{1 + \frac{1}{c^2} v (v_2)_z}. \end{aligned}$$

Telles sont les formules de composition des vitesses de la cinématique relativiste qui donnent les composantes de la vitesse dans le système (1) à partir des composantes dans le système (2). Elles remplacent les formules de la cinématique classique

$$(v_1)_x = (v_2)_x, \quad (v_1)_y = (v_2)_y, \quad (v_1)_z = (v_2)_z + v$$

auxquelles elles se réduisent quand v est beaucoup plus petit que c .

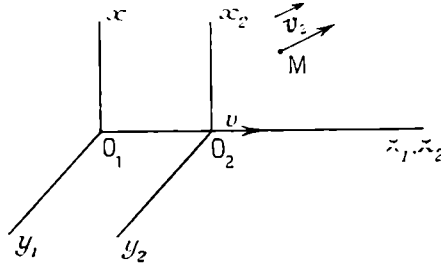


Fig. 9.

En changeant β en $-\beta$, on obtient les formules inverses

$$(v_2)_x = \frac{(v_1)_x \sqrt{1 - \beta^2}}{1 - \frac{\beta}{c} (v_1)_z}, \quad (v_2)_y = \frac{(v_1)_y \sqrt{1 - \beta^2}}{1 - \frac{\beta}{c} (v_1)_z}, \quad (v_2)_z = \frac{(v_1)_z - v}{1 - \frac{\beta}{c} (v_1)_z}.$$

8. Interprétation de l'expérience de Fizeau par la cinématique relativiste. — Dans l'expérience classique de Fizeau sur « l'entraînement partiel de l'éther par les corps réfringents en mouvement », une substance réfringente d'indice n est animée d'un mouvement de

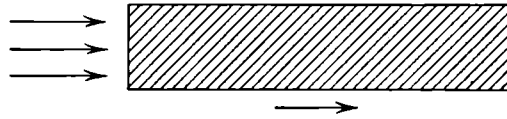


Fig. 10.

translation rectiligne et uniforme et un faisceau de lumière parallèle se propage dans la substance dans la direction du mouvement.

L'expérience prouve que, du moins en première approximation quand on néglige la dispersion, la vitesse V de la lumière dans la substance réfringente en mouvement est donnée par la formule

$$(4) \quad V = \frac{c}{n} + v \left(1 - \frac{1}{n^2} \right).$$

Dans l'ancienne théorie ondulatoire de la lumière, on admettait qu'elle se propageait dans un « éther », où baignerait la substance mobile, avec la vitesse $\frac{c}{n}$ d'après la définition même de l'indice de réfraction. Si cet éther est « entraîné » par le mouvement de la substance réfringente de façon à prendre lui-même une vitesse qui soit une certaine fraction z de la vitesse v , la lumière possédera par rapport à l'expérimentateur, par suite de cet entraînement partiel de l'éther, une vitesse de propagation égale à

$$V = \frac{c}{n} + z v.$$

La formule (4) peut alors s'interpréter en disant que le coefficient d'entraînement partiel z est égal à $1 - \frac{1}{n^2}$.

Les théories anciennes, en particulier la théorie des Électrons, permettaient de retrouver ainsi la formule (4), mais la théorie de la Relativité en fournit une démonstration purement cinématique d'une remarquable élégance.

En effet, la vitesse de la lumière dans la substance réfringente étant $\frac{c}{n}$, la vitesse par rapport à l'observateur fixe est donnée d'après le cinématique relativiste par la formule

$$v_1 = \frac{v_2 + v}{1 + \frac{1}{c^2} v v_2}, \quad \text{avec} \quad v_1 = V \quad \text{et} \quad v_2 = \frac{c}{n}.$$

La vitesse d'entraînement que l'on peut pratiquement imposer à la substance réfringente étant très petite devant c , β est petit et β^2 négligeable devant l'unité. On trouve donc

$$V = \frac{\frac{c}{n} + v}{1 + \frac{1}{c^2} v \frac{c}{n}} \simeq \left(\frac{c}{n} + v \right) \left(1 - \frac{\beta}{n} \right) = \frac{c}{n} (1 + n\beta) \left(1 - \frac{\beta}{n} \right) \simeq \frac{c}{n} + v \left(1 - \frac{1}{n^2} \right),$$

ce qui est bien la formule de Fizeau.

Nous remarquerons que Lorentz en tenant compte de la dispersion et de l'effet Doppler, a trouvé la formule plus exacte

$$V = \frac{c}{n(\lambda)} + v \left[1 - \frac{1}{n^2(\lambda)} - \frac{\lambda}{n(\lambda)} \left(\frac{dn}{d\lambda} \right)_\lambda \right],$$

où λ est la longueur d'onde de l'onde incidente pour l'expérimentateur.

Cette formule a été vérifiée par Zeeman qui a repris les expériences de Fizeau sur des liquides en mouvement et les a étendues à des corps solides en mouvement. Mais, comme la correction de Lorentz n'est pas spécifiquement relativiste, nous n'y insisterons pas.

BIBLIOGRAPHIE.

- [1] MAX VON LAUE, *Théorie de la Relativité* (2 vol.), Gauthier Villars, 1922.
- [2] JEAN BECQUEREL, *Le principe de Relativité*, Gauthier-Villars, Paris, 1922.
- [3] H. A. LORENTZ, *Relativitätstheorie*, Akademische Verlagsgesellschaft, Leipzig, 1929.
- [4] R. BECKER, *Théorie des Électrons*, Alcan, Paris, 1938.
- [5] O. COSTA DE BEAUREGARD, *La théorie de la Relativité restreinte*, Masson, Paris, 1949.

CHAPITRE III.

COMPLÉMENTS SUR LA THÉORIE DE RELATIVITÉ RESTREINTE.

1. **L'espace-temps.** — Nous supposons connus du lecteur les principes du calcul tensoriel.

La relation entre les coordonnées x, y, z, t et x', y', z', t' de deux observateurs galiléens qui est exprimée analytiquement par la transformation de Lorentz peut être représentée géométriquement.

Pour y parvenir, nous introduisons d'abord un espace à quatre dimensions défini par les quatre coordonnées

$$x^1 = x, \quad x^2 = y, \quad x^3 = z, \quad x^4 = ct.$$

La métrique de cet espace sera déterminée par le fait que la distance des deux points infiniment voisins de coordonnées x^1, x^2, x^3, x^4 et $x^1 + dx^1, x^2 + dx^2, x^3 + dx^3, x^4 + dx^4$ est donnée par la formule

$$(1) \quad ds^2 = (dx^4)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 = c^2 dt^2 - dx^2.$$

Chaque point de cet espace à quatre dimensions représente un « événement », c'est-à-dire un point de l'espace physique considéré à un instant donné. L'espace à quatre dimensions ainsi introduit est « l'espace-temps » ou « Univers » et le ds^2 est le carré de l'élément de longueur dans l'espace-temps. L'espace-temps est une multiplicité pseudo-euclidienne parce que dans l'expression (1) du ds^2 ne figurent que des carrés comme dans un ds^2 euclidien, mais que ces carrés n'ont pas tous le même signe. On appelle transformations orthogonales des coordonnées dans l'espace-temps celles qui conservent la forme du ds^2 : elles s'expriment précisément par les formules de Lorentz.

Dans l'espace-temps, un vecteur est un être mathématique à quatre composantes rectangulaires A^1, A^2, A^3, A^4 qui se transforment comme x^1, x^2, x^3, x^4 lors d'un changement de système de référence galiléen.

Le carré de sa longueur est donné par une formule qui a même forme que celle du ds^2 , savoir

$$|\mathbf{A}|^2 = (\mathbf{A}^4)^2 - (\mathbf{A}^1)^2 - (\mathbf{A}^2)^2 - (\mathbf{A}^3)^2 = (\mathbf{A}^4)^2 - (\mathbf{A}_t)^2.$$

Cette quantité, comme le ds^2 , est invariante pour tous les changements de système de référence galiléen.

Le produit scalaire de deux vecteurs d'espace-temps \mathbf{A} et \mathbf{B} est, par définition,

$$(\mathbf{A} \cdot \mathbf{B}) = \mathbf{A}^4 \mathbf{B}^4 - \mathbf{A}^1 \mathbf{B}^1 - \mathbf{A}^2 \mathbf{B}^2 - \mathbf{A}^3 \mathbf{B}^3 = \mathbf{A}^4 \mathbf{B}^4 - (\mathbf{A}_t \cdot \mathbf{B}_t),$$

\mathbf{A}_t et \mathbf{B}_t étant les projections sur l'espace physique de \mathbf{A} et de \mathbf{B} . Le produit scalaire $(\mathbf{A} \cdot \mathbf{B})$ est invariant pour les transformations de Lorentz. Le produit scalaire d'un vecteur par lui-même est évidemment égal au carré de sa longueur.

On introduit souvent à côté des composantes « contrevariantes » $\mathbf{A}^1, \mathbf{A}^2, \mathbf{A}^3$ et \mathbf{A}^4 les composantes « covariantes » $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ et \mathbf{A}_4 reliées aux premières par les relations

$$\mathbf{A}_1 = -\mathbf{A}^1, \quad \mathbf{A}_2 = -\mathbf{A}^2, \quad \mathbf{A}_3 = -\mathbf{A}^3, \quad \mathbf{A}_4 = \mathbf{A}^4$$

et l'on peut alors écrire

$$|\mathbf{A}|^2 = \sum_1^4 \mathbf{A}_i \mathbf{A}^i, \quad (\mathbf{A} \cdot \mathbf{B}) = \sum_1^4 \mathbf{A}_i \mathbf{B}^i = \sum_1^4 \mathbf{A}^i \mathbf{B}_i.$$

Par définition, deux vecteurs d'espace-temps sont orthogonaux si leur produit scalaire d'espace-temps est nul.

Si nous voulons nous représenter graphiquement l'espace-temps, nous nous heurtons à deux difficultés : il a quatre dimensions et sa métrique est pseudo-euclidienne. Pour tourner ces difficultés, nous nous placerons d'abord dans le système de référence d'un certain observateur galiléen et nous tracerons sur notre papier l'axe $x_4 = ct$ correspondant au temps propre de cet observateur.

L'ensemble des points de l'espace-temps qui sont simultanés pour cet observateur et qui forment à un instant t de son temps propre son espace physique est représenté par un hyperplan à trois dimensions perpendiculaire à l'axe x_4 au point de cote ct . Nous devons imaginer trois axes de coordonnées normaux à l'axe x_4 et normaux entre eux qui servent à repérer les points de l'espace physique pour l'observateur considéré.

Si nous considérons le cône C formé par l'ensemble des droites qui

passent par O et sont inclinées de 45° sur Ox_4 , tout vecteur ayant son origine en O et situé sur ce cône aura une longueur nulle dans l'espace-temps, car sa projection A_l sur l'espace est égale à sa projection A_4 sur le temps. Tout vecteur contenu dans le cône double aura son $|A|^2$ positif : on dit que c'est un « vecteur du genre temps ». Tout vecteur extérieur au cône double aura son $|A|^2$ négatif, on dira que c'est un « vecteur du genre espace ».

Deux vecteurs symétriques par rapport au cône sont orthogonaux, car la projection de l'un sur l'espace est égale à la projection de l'autre sur le temps [$A_l = B_4$ et $B_l = A_4$, avec A_l parallèle à B_l , d'où $(\mathbf{A}, \mathbf{B}) = A_4 B_4 - A_l B_l = 0$].

La propagation d'un signal lumineux émis au point O à l'instant

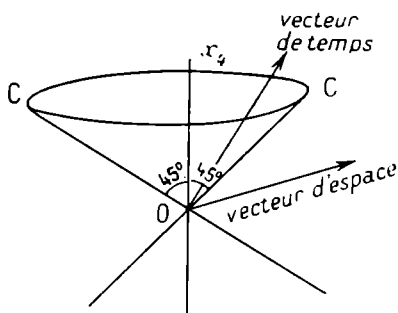


Fig. 11.

$t = 0$ est représenté par les génératrices du cône C , puisque pour tout point P de ce cône, on a

$$\overline{OP}^2 = c^2 t^2 - x^2 - y^2 - z^2 = 0.$$

C'est ce qui a fait donner au cône C le nom de « cône de lumière » relatif au point O .

Les signaux autres que les signaux lumineux ou électromagnétiques ont des vitesses inférieures à c et leur propagation est représentée par des courbes contenues dans C . La partie supérieure du cône C contient donc tous les événements qui peuvent être influencés par l'événement O . Pour cette raison, on appelle « cône de l'avenir » relatif au point O , la nappe supérieure du double cône C ; on appelle de même « cône du passé » relatif à O , la nappe inférieure du double cône, car elle contient tous les événements qui sont susceptibles d'influencer l'événement O . La région extérieure au double cône contient les événements qui ne peuvent influencer O , ni être influencés par lui.

Il est essentiel de remarquer que, l'écoulement du temps présentant un sens privilégié, l'espace-temps possède une sorte de polarité dans le sens des x_4 croissants.

Considérons maintenant un objet ponctuel, un point matériel. A chaque instant, il occupe une certaine position dans l'espace. Son mouvement définit donc pour ses coordonnées x^1, x^2, x^3, x^4 une suite de valeurs représentée par une courbe continue de l'espace-temps. Cette courbe, c'est la « ligne d'Univers » du point matériel qui symbolise l'ensemble de son évolution. Comme, d'après Einstein, aucun corps matériel ne peut aller plus vite que la lumière et que le temps a un sens d'écoulement, la ligne d'Univers d'un corps quelconque doit toujours faire un angle inférieur à 45° avec la direction positive de l'axe des x_4 . En effet, si ds est un élément de la ligne d'Univers en question dont la projection sur x^4 est $dx^4 = c dt$ et dont la projection sur l'espace est dl ,

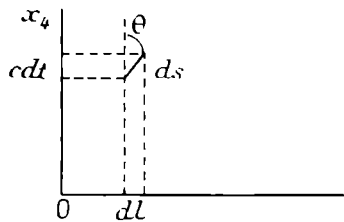


Fig. 12.

la vitesse du mobile dans la phase de son mouvement que représente ds est égale au produit par c de l'angle θ que fait ds avec Ox^4

$$\operatorname{tg} \theta = \frac{dl}{dx^4} = \frac{1}{c} \frac{dl}{dt}.$$

D'où la proposition énoncée, car v doit être inférieure à c et, par suite, tous les éléments de la ligne d'Univers doivent être des vecteurs de temps.

Seule la lumière a dans le vide une ligne d'Univers qui fait 45° avec Ox^4 . La ligne d'Univers de la lumière dans le vide est une courbe de longueur nulle, car $ds^2 = c^2 dt^2 - dl^2 = 0$.

2. Démonstration géométrique dans l'espace-temps des relations de Lorentz. — Nous allons maintenant chercher à utiliser l'espace-temps pour retrouver par un raisonnement géométrique la transformation de Lorentz. Nous prendrons le cas simple où il y a coïncidence des axes Oz et $O'z'$ et mouvement relatif le long de ces axes. Nous parlerons

toujours des observateurs A et B et, nous servant de la représentation graphique indiquée ci-dessus, nous représenterons l'axe des t et celui des z de l'observateur A.

Dans cette représentation de l'espace-temps, l'origine O' du système de référence de B a la droite Ot' comme ligne d'Univers, si nous supposons que l'on a choisi pour instant initial l'instant où O et O' coïncidaient. Ot' fait avec Ot l'angle θ et l'on a $\frac{\overline{mM}}{\overline{Om}} = \operatorname{tg}\theta = \beta$. A l'instant $t = 0$, l'axe des z' de B doit se représenter par Oz' ; cela résulte du fait que la ligne d'Univers OD de la lumière, bissectrice de zOt , doit aussi être bissectrice de $z'Ot'$, puisque pour B comme pour A, la lumière a la vitesse c . Les événements du plan zOt qui sont simultanés pour B

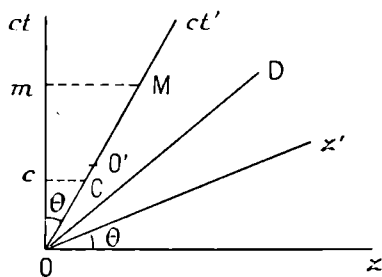


Fig. 13.

sont donc situés sur des parallèles à Oz' ; ils ne sont pas simultanés pour A. Les événements du plan zOt qui ont même abscisse z' pour B sont situés sur des parallèles à Ot' ; ils n'ont pas même abscisse z pour A.

Pour retrouver sur la figure les formules de Lorentz, il faut tenir compte du caractère pseudo-euclidien du plan zOt , qui nous donne dans le triangle $O m M$ la relation

$$\overline{OM}^2 = \overline{Om}^2 - \overline{mM}^2.$$

Marquons sur l'axe Ot' un point C tel que $\overline{OC} = 1$ et soit c la projection de C sur Ot . Nous aurons

$$\overline{Oc}^2 - \overline{cC}^2 = 1,$$

d'où, puisque $\frac{\overline{cC}}{\overline{Oc}} = \beta$,

$$\overline{Oc} = \frac{1}{\sqrt{1 - \beta^2}}, \quad \overline{cC} = \frac{\beta}{\sqrt{1 - \beta^2}}.$$

Les coordonnées de M étant $\overline{Om} = ct$ et $\overline{mM} = z$ et \overline{OM} étant égal à ct' , nous obtenons pour le produit scalaire $(\overline{OC}, \overline{OM})$ l'expression

$$(\overline{OC}, \overline{OM}) = {}_1ct' = \overline{Oc} \cdot \overline{Om} - \overline{mM} \cdot \overline{cC} = \frac{ct}{\sqrt{1-\beta^2}} - \frac{\beta z}{\sqrt{1-\beta^2}},$$

d'où

$$t' = \frac{t - \frac{\beta}{c} z}{\sqrt{1-\beta^2}}.$$

C'est la formule de Lorentz pour le passage de t' à t . On trouverait de même la formule donnant le passage de z' à z en prenant un point M sur l'axe Oz' .

La figure que nous avons donnée plus haut semble faire jouer un rôle dissymétrique aux deux observateurs A et B , puisqu'on croit lire sur la figure

$$\widehat{zOt} = 90^\circ \quad \text{et} \quad \widehat{z'Ot'} < 90^\circ$$

et ce fait paraît contraire au principe de relativité. Mais ce n'est là qu'une apparence due à la difficulté de représenter un plan pseudo-euclidien sur une feuille de papier qui est un plan euclidien. En réalité, avec nos définitions, Oz' symétrique de Ot' par rapport à D lui est orthogonal (se reporter à une remarque précédente selon laquelle deux vecteurs symétriques par rapport au cône C sont orthogonaux). On pourrait refaire la figure en partant des axes de l'observateur B figurés perpendiculaires sur le papier et tous les résultats seraient les mêmes.

3. Vecteurs et tenseurs d'espace-temps. — Rappelons quelques définitions.

Un invariant ou tenseur de rang 0 dans l'espace-temps est une expression qui a la même valeur dans tous les systèmes galiléens. Le ds^2 ou carré de l'élément de longueur en est un exemple; la valeur d'un élément de volume de l'espace-temps en est un autre.

Un vecteur d'espace-temps \mathbf{A} ou tenseur de rang 1 est un être mathématique à quatre composantes (contrevariantes) A^1, A^2, A^3, A^4 telles que dans une transformation de Lorentz

$$x'^i = \sum_k \alpha_k^i x^k,$$

elles se transforment comme les coordonnées, c'est-à-dire suivant le schéma

$$A'^i = \sum_k \alpha_k^i A^k$$

avec les mêmes α_k^i . De cette définition, résulte que le carré de la longueur d'un vecteur \mathbf{A} , soit

$$|\mathbf{A}|^2 = (A^4)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2,$$

et le produit scalaire de deux vecteurs \mathbf{A} et \mathbf{B} , soit

$$(\mathbf{A}, \mathbf{B}) = A^4 B^4 - A^1 B^1 - A^2 B^2 - A^3 B^3$$

sont des invariants.

Un exemple typique de vecteur d'espace-temps est le vecteur densité-courant d'un fluide en mouvement. Considérons un fluide formé de molécules. Pour un certain observateur galiléen, ce fluide a en chaque point une certaine densité ρ (nombre de molécules par unité de volume) et une certaine vitesse de composantes u_x, u_y, u_z . On peut démontrer en s'appuyant sur la formule de composition des vitesses en cinématique relativiste que les quatre grandeurs

$$A^1 = \rho u_x, \quad A^2 = \rho u_y, \quad A^3 = \rho u_z, \quad A^4 = \rho c$$

se transforment comme x^1, x^2, x^3, x^4 dans un changement de système galiléen. On obtient ainsi un vecteur \mathbf{A} dont la composante de temps est, au facteur c près, égale à la densité ρ , tandis que les composantes d'espace représentent le flux du fluide par unité de surface. La quantité $|\mathbf{A}|^2 = \rho^2 c^2 \left(1 - \frac{u^2}{c^2}\right)$ doit donc être un invariant. Il en est bien ainsi car, au facteur c^2 près, cette quantité est le carré de $\rho \sqrt{1 - \frac{u^2}{c^2}}$ et ceci est la densité ρ_0 du fluide pour un observateur entraîné par le mouvement local de ce fluide. En effet, pour cet observateur, il y a, par suite de la contraction de Lorentz, ρ molécules du fluide dans le volume $\frac{1}{\sqrt{1 - \beta^2}}$, d'où pour lui une densité $\rho_0 = \rho \sqrt{1 - \beta^2}$. Nous trouverons plus loin d'autres exemples de vecteurs d'espace-temps.

Un tenseur de rang 2 est un être mathématique à seize composantes A^{ij} qui forment un tableau à quatre lignes et quatre colonnes et qui sont telles que, pour un changement de système galiléen, les A^{ij} se transforment comme les produits $x^i x^j$. Sans étudier ici en détail les propriétés

de ces tenseurs, disons qu'il résulte de leur définition que la quantité

$$|\mathbf{A}|^2 = \sum_{ik}^4 \varepsilon_{ik} (\mathbf{A}^{ik})^2,$$

avec $\varepsilon_{ik} = -1$ si l'un des deux indices est égal à 4 et $\varepsilon_{ik} = 1$ dans le cas contraire, est un invariant. En introduisant des composantes covariantes $\mathbf{A}_{ik} = \varepsilon_{ik} \mathbf{A}^{ik}$, on peut écrire plus simplement

$$|\mathbf{A}|^2 = \sum_{ik}^4 \mathbf{A}_{ik} \mathbf{A}^{ik}.$$

À côté des tenseurs symétriques ($\mathbf{A}^{ik} = \mathbf{A}^{ki}$), une catégorie très importante de tenseurs du second rang sont les tenseurs antisymétriques tels que $\mathbf{A}^{ik} = -\mathbf{A}^{ki}$. Les termes symétriques par rapport à la diagonale dans le tableau des \mathbf{A}^{ik} sont alors égaux et de signe contraire, tandis que les termes diagonaux \mathbf{A}^{kk} sont nuls : il n'y a donc plus que six composantes distinctes \mathbf{A}^{14} , \mathbf{A}^{24} , \mathbf{A}^{34} , \mathbf{A}^{23} , \mathbf{A}^{31} et \mathbf{A}^{12} . On démontre que pour une rotation des axes xyz dans l'espace, les grandeurs \mathbf{A}^{14} , \mathbf{A}^{24} , \mathbf{A}^{34} d'une part, les grandeurs \mathbf{A}^{23} , \mathbf{A}^{31} , \mathbf{A}^{12} d'autre part, se transforment comme les composantes d'un vecteur à trois dimensions, de sorte que, pour les changements d'axes galiléens sans mouvement relatif, tout se passe comme si le tenseur d'espace-temps se décomposait en deux vecteurs d'espace. Mais pour une transformation de Lorentz qui fait varier la variable temps, il n'en est plus ainsi et dans le cas d'une transformation simple de Lorentz, on trouve

$$\begin{aligned} \mathbf{A}'^{14} &= \frac{\mathbf{A}^{14} - \beta \mathbf{A}^{31}}{\sqrt{1 - \beta^2}}, & \mathbf{A}'^{24} &= \frac{\mathbf{A}^{24} + \beta \mathbf{A}^{23}}{\sqrt{1 - \beta^2}}, & \mathbf{A}'^{34} &= \mathbf{A}^{34}; \\ \mathbf{A}'^{23} &= \frac{\mathbf{A}^{23} + \beta \mathbf{A}^{24}}{\sqrt{1 - \beta^2}}, & \mathbf{A}'^{31} &= \frac{\mathbf{A}^{31} - \beta \mathbf{A}^{14}}{\sqrt{1 - \beta^2}}, & \mathbf{A}'^{12} &= \mathbf{A}^{12}, \end{aligned}$$

formules importantes, nous le verrons, pour l'électromagnétisme relativiste.

Pour un tenseur antisymétrique de rang 2, l'invariant « carré de la longueur » prend la forme

$$|\mathbf{A}|^2 = 2[(\mathbf{A}^{23})^2 + (\mathbf{A}^{31})^2 + (\mathbf{A}^{12})^2 - (\mathbf{A}^{14})^2 - (\mathbf{A}^{24})^2 - (\mathbf{A}^{34})^2].$$

4. La Dynamique relativiste du point matériel. — La Mécanique classique de Newton admet le groupe de transformation de Galilée, mais non le groupe de transformation de Lorentz. Mais, comme les formules de transformation de ces deux groupes ne diffèrent que par des termes

de l'ordre de β^2 , la divergence est très faible pour tous les mouvements considérés par la Mécanique rationnelle, y compris le mouvement des corps célestes. La transformation de Lorentz se trouvant confirmée par des expériences sur la propagation de la lumière et sur les interférences qui sont infiniment plus précises que les expériences de Mécanique, il est naturel de penser que les équations de la Mécanique doivent être réformées de manière à devenir invariantes pour la transformation de Lorentz et l'on pressent que ceci n'entraînera dans les équations de la Mécanique que des modifications de l'ordre de β^2 qui seront tout à fait négligeables quand on aura à traiter des problèmes de Mécanique à grande échelle. C'est en suivant cette ligne d'idée que M. Einstein a pu constituer une dynamique nouvelle du point matériel qui a été en son temps qualifiée de « Mécanique nouvelle » et qui est à son tour devenue une Mécanique ancienne depuis le développement des Mécaniques ondulatoires et quantiques.

La condition essentielle que doit *a priori* remplir la Dynamique relativiste du point matériel, c'est de se confondre avec la Dynamique classique chaque fois que β^2 est négligeable devant l'unité, car il faut retrouver l'ancienne Dynamique comme première approximation dans le cas des vitesses faibles devant celle de la lumière dans le vide (approximation newtonienne).

On doit donc chercher à mettre, à la base de la Dynamique relativiste du point matériel, un principe de moindre action qui se réduise pour $\beta \ll 1$ au principe de Hamilton. Celui-ci nous apprend que, si un point matériel se déplace dans un champ de force défini par une fonction potentielle $U(x, y, z, t)$ telle que la force subie par le point matériel soit égale au gradient de U changé de signe, depuis une position x_0, y_0, z_0 occupée à l'instant initial t_0 jusqu'à une position x_1, y_1, z_1 occupée à l'instant final t_1 , le mouvement du point matériel est tel que l'intégrale

$$\int_{t_0}^{t_1} (T - U) dt$$

(où T est l'énergie cinétique $\frac{1}{2}mv^2$ du mobile) est minimum, ou plus généralement stationnaire, par rapport aux mouvements virtuellement voisins qui amèneraient le point matériel de la position x_0, y_0, z_0 à l'instant t_0 à la position x_1, y_1, z_1 à l'instant t_1 . On dit que

$$\mathcal{L} \left(x, y, z, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}, t \right) = T - U$$

est la « fonction de Lagrange » du point matériel dans la Dynamique classique. On a donc à exprimer que

$$\delta \int_{t_1}^{t_2} \mathcal{L} dt = 0.$$

Posons

$$\frac{dx}{dt} = \dot{x}, \quad \frac{dy}{dt} = \dot{y}, \quad \frac{dz}{dt} = \dot{z}.$$

Un calcul classique, s'appuyant sur les méthodes du calcul des variations, montre que cette condition d'extremum est équivalente aux trois équations

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = \frac{\partial \mathcal{L}}{\partial x}, \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{y}} \right) = \frac{\partial \mathcal{L}}{\partial y}, \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{z}} \right) = \frac{\partial \mathcal{L}}{\partial z}.$$

Ce sont les « équations de Lagrange » qu'on peut aussi écrire si U ne dépend pas des vitesses

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{x}} \right) - \frac{\partial T}{\partial x} &= - \frac{\partial U}{\partial x}, & \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{y}} \right) - \frac{\partial T}{\partial y} &= - \frac{\partial U}{\partial y}, \\ \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{z}} \right) - \frac{\partial T}{\partial z} &= - \frac{\partial U}{\partial z}. \end{aligned}$$

Définissons les composantes de la quantité de mouvement (ou moments de Lagrange) par les formules

$$p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}}, \quad p_y = \frac{\partial \mathcal{L}}{\partial \dot{y}}, \quad p_z = \frac{\partial \mathcal{L}}{\partial \dot{z}}.$$

On peut écrire simplement

$$\frac{dp_x}{dt} = \frac{\partial \mathcal{L}}{\partial x}, \quad \frac{dp_y}{dt} = \frac{\partial \mathcal{L}}{\partial y}, \quad \frac{dp_z}{dt} = \frac{\partial \mathcal{L}}{\partial z}.$$

En coordonnées rectangulaires,

$$T = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2),$$

d'où, par exemple,

$$p_x = m\dot{x}, \quad \frac{d}{dt} (m\dot{x}) = - \frac{\partial U}{\partial x}$$

et l'on retrouve les équations classiques de Newton

$$m\ddot{x} = f_x, \quad m\ddot{y} = f_y, \quad m\ddot{z} = f_z.$$

Après ce bref rappel de notions classiques en Mécanique analytique, revenons au point de vue relativiste. Considérons dans l'espace-temps

la ligne d'Univers d'un point matériel et étudions-la en nous servant des coordonnées x, y, z, t d'un observateur galiléen. L'élément de longueur de la ligne d'Univers est donné par

$$ds^2 = \sqrt{c^2 dt^2 - dx^2 - dy^2 - dz^2} = \sqrt{c^2 dt^2 - dt^2},$$

dt étant l'élément de la trajectoire parcouru par le point matériel pendant le temps dt , de sorte que la vitesse du mobile pour l'observateur galiléen est $v = \frac{dl}{dt}$ et l'on a

$$ds = c dt \sqrt{1 - \beta^2} \quad \left(\beta = \frac{v}{c} \right).$$

Posons encore $d\tau = dt \sqrt{1 - \beta^2}$; $d\tau$ est l'élément de « temps propre » du point matériel, c'est-à-dire du temps marqué par une horloge entraînée par le point matériel : en effet, la relation qui définit $d\tau$ exprime le ralentissement de l'horloge vue du système de l'observateur. L'élément de ligne d'Univers $ds = c d\tau$ est donc proportionnel à la variation du temps propre.

En Dynamique relativiste, on considérera les positions du point matériel aux temps t_0 et t_1 comme définissant deux points de l'espace-temps, $P_0(x_0, y_0, z_0, t_0)$ et $P_1(x_1, y_1, z_1, t_1)$. Il est alors naturel de chercher un principe d'action stationnaire de la forme

$$\delta \int_P^Q \mathcal{L} dt = 0,$$

l'intégrale étant prise le long de la ligne d'Univers de P en Q . La fonction \mathcal{L} sera différente par des termes de l'ordre de β^2 de la fonction de Lagrange classique; elle devra être telle que l'intégrale puisse s'écrire

sous la forme invariante $\int_P^Q \mathcal{L}_0 d\tau$, l'invariant \mathcal{L}_0 étant défini par

$$\mathcal{L}_0 = \frac{\mathcal{L}}{\sqrt{1 - \beta^2}}.$$

Considérons d'abord la partie de \mathcal{L} qui dépend du mouvement et non du champ et qui correspond au terme T de l'expression classique. Si nous caractérisons le point matériel envisagé par une constante invariante, sa masse propre m_0 , nous obtiendrons une forme satisfaisante pour cette partie cinétique de la fonction de Lagrange en posant

$$\mathcal{L} = - m_0 c^2 \sqrt{1 - \beta^2}.$$

Il est d'abord évident que $\mathcal{L}_0 = \frac{\mathcal{L}}{\sqrt{1 - \beta^2}}$ est un invariant.

De plus, pour β petit nous aurons approximativement

$$\mathcal{L} = -m_0 c^2 + \frac{1}{2} m_0 v^2$$

et, comme un terme constant n'a aucune importance dans l'expression de \mathcal{L} , nous retombons bien sur la fonction de Lagrange classique. En l'absence de champ, l'intégrale d'action sera donc

$$\int_p^Q (-m_0 c^2) dt = \int_p^Q (-m_0 c^2 \sqrt{1-\beta^2}) dt.$$

En présence d'un champ dérivant d'une fonction potentielle, nous nous contenterons pour l'instant d'ajouter à l'expression du terme cinétique de \mathcal{L} le terme $-U$. Nous examinerons plus loin la validité de cette hypothèse. Finalement, nous écrivons le principe de Hamilton en Dynamique relativiste du point matériel sous la forme

$$\delta \int_p^Q (-m_0 c^2 \sqrt{1-\beta^2} - U) dt = 0$$

et nous obtenons encore par le calcul des variations des équations formellement analogues à celles obtenues en Mécanique analytique classique; par exemple

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = \frac{\partial \mathcal{L}}{\partial x} = - \frac{\partial U}{\partial x}.$$

Définissons encore les « moments de Lagrange » par

$$p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}}, \quad p_y = \frac{\partial \mathcal{L}}{\partial \dot{y}}, \quad p_z = \frac{\partial \mathcal{L}}{\partial \dot{z}};$$

on trouve

$$p_x = \frac{m_0 \dot{x}}{\sqrt{1-\beta^2}}, \quad p_y = \frac{m_0 \dot{y}}{\sqrt{1-\beta^2}}, \quad p_z = \frac{m_0 \dot{z}}{\sqrt{1-\beta^2}}.$$

La quantité de mouvement sera donc ici

$$\mathbf{p} = \frac{m_0 \mathbf{v}}{\sqrt{1-\beta^2}}.$$

On peut encore l'écrire $\mathbf{p} = m \mathbf{v}$, mais la « masse en mouvement » m du point matériel est égale à $m = \frac{m_0}{\sqrt{1-\beta^2}}$ et varie avec la vitesse. La masse m d'un point matériel pour un observateur augmente avec sa vitesse. Pour un observateur lié au point en mouvement, $\beta = 0$ et $m = m_0$; d'où l'origine des noms de « masse propre » et « masse au

repos » donnés à m_0 . Pour v tendant vers c , m tend vers l'infini. Aucun mobile de masse finie ne peut donc atteindre la vitesse c et ainsi se trouve justifiée *a posteriori* l'hypothèse de départ d'Einstein.

Les équations de la Dynamique relativiste peuvent s'écrire

$$\frac{dp_x}{dt} = \frac{d}{dt} \left(\frac{m_0 v_x}{\sqrt{1-\beta^2}} \right) = - \frac{\partial U}{\partial x}, \quad \frac{dp_y}{dt} = \frac{d}{dt} \left(\frac{m_0 v_y}{\sqrt{1-\beta^2}} \right) = - \frac{\partial U}{\partial y},$$

$$\frac{dp_z}{dt} = \frac{d}{dt} \left(\frac{m_0 v_z}{\sqrt{1-\beta^2}} \right) = - \frac{\partial U}{\partial z}.$$

Reprenons ces équations du mouvement sous la forme générale

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = \frac{\partial \mathcal{L}}{\partial x}, \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{y}} \right) = \frac{\partial \mathcal{L}}{\partial y}, \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{z}} \right) = \frac{\partial \mathcal{L}}{\partial z}$$

qui est valable à la fois en Mécanique classique et en Mécanique relativiste, mais avec des formes différentes de la fonction \mathcal{L} . Posons par définition

$$W = \dot{x} \frac{\partial \mathcal{L}}{\partial \dot{x}} + \dot{y} \frac{\partial \mathcal{L}}{\partial \dot{y}} + \dot{z} \frac{\partial \mathcal{L}}{\partial \dot{z}} - \mathcal{L} = \dot{x} p_x + \dot{y} p_y + \dot{z} p_z - \mathcal{L}.$$

On démontre aisément à partir des équations précédentes que

$$\frac{dW}{dt} = - \frac{\partial \mathcal{L}}{\partial t} = \frac{\partial U}{\partial t}.$$

Donc, si U ne dépend pas du temps (champ permanent), W reste constant. W est appelée « l'énergie » du point matériel. Dans un champ permanent, l'énergie d'un point matériel reste constante : ce théorème est exact en Mécanique classique et en Mécanique relativiste, mais avec des expressions différentes de W que nous allons donner.

a. En Mécanique classique, on a

$$\mathcal{L} = T - U, \quad p_x = m\dot{x} \quad \dots, \quad \text{avec } m \text{ constant.}$$

On trouve alors

$$W = m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - (T - U) = T + U = \frac{1}{2} m v^2 + U,$$

expression bien connue :

b. En Mécanique relativiste,

$$\mathcal{L} = - m_0 c^2 \sqrt{1-\beta^2} - U, \quad p_x = \frac{m_0 \dot{x}}{\sqrt{1-\beta^2}},$$

On trouve

$$\begin{aligned} W &= \frac{m_0(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)}{\sqrt{1-\beta^2}} + m_0 c^2 \sqrt{1-\beta^2} + U, \\ &= \frac{m_0 c^2}{\sqrt{1-\beta^2}} + U. \end{aligned}$$

Le terme $\frac{m_0 c^2}{\sqrt{1-\beta^2}}$ est nouveau et intéressant. Dans le système lié au mobile (système propre), $\beta = 0$, le terme se réduit à $m_0 c^2$. C'est « l'énergie propre » du point matériel. Pour un observateur qui voit le point matériel se mouvoir, l'énergie croît comme $\frac{1}{\sqrt{1-\beta^2}}$. En généralisant ces faits, Einstein a été conduit au « principe de l'inertie de l'énergie » dont voici l'énoncé : « A toute masse est toujours associée une énergie égale au produit de cette masse par c^2 ».

Ainsi pour l'observateur lié au mobile, la masse est m_0 et l'énergie $m_0 c^2$; pour l'observateur qui voit passer le mobile avec la vitesse βc , la masse est $\frac{m_0}{\sqrt{1-\beta^2}}$ et l'énergie $\frac{m_0 c^2}{\sqrt{1-\beta^2}}$. Le principe de l'inertie de l'énergie peut être vérifié sur un grand nombre d'exemples plus compliqués : il joue aujourd'hui un rôle capital dans la Physique du noyau de l'atome.

En Dynamique relativiste, on appelle souvent « énergie cinétique » l'expression

$$T = m_0 c^2 \left(\frac{1}{\sqrt{1-\beta^2}} - 1 \right).$$

C'est l'augmentation de l'énergie due au mouvement. Pour $\beta \ll 1$, on trouve

$$T = \frac{1}{2} m_0 v^2,$$

c'est-à-dire qu'on retombe sur l'expression classique.

On a $W = m_0 c^2 + E$ en posant $E = T + U$. La Mécanique classique qui ignore le terme constant $m_0 c^2$ introduit par la relativité, désigne toujours par énergie la quantité $E = T + U$ en laissant indéterminée une constante additive. La constante de l'énergie est donc fixée par la théorie de la Relativité; nous verrons que la constante de l'entropie est fixée par la théorie des Quanta. Remarquons que W est toujours positif, tandis que E peut être négatif.

Nous allons définir maintenant dans l'espace-temps le vecteur

« impulsion d'Univers » \mathbf{I} du point matériel. Il a comme composantes

$$\begin{aligned} I^1 &= p_x = \frac{m_0 v_x}{\sqrt{1-\beta^2}}, & I^2 &= p_y = \frac{m_0 v_y}{\sqrt{1-\beta^2}}, \\ I^3 &= p_z = \frac{m_0 v_z}{\sqrt{1-\beta^2}}, & I^4 &= \frac{W}{c} = \frac{m_0 c}{\sqrt{1-\beta^2}}. \end{aligned}$$

Sa longueur invariante est donnée par

$$|\mathbf{I}|^2 = \left(\frac{W}{c}\right)^2 - p_x^2 - p_y^2 - p_z^2 = m_0^2 c^2.$$

Il a pour composante de temps (au facteur $\frac{1}{c}$ près) l'énergie du mobile et pour composantes d'espace les composantes de sa quantité de mouvement : il réunit donc en un seul être mathématique les notions d'énergie et de quantité de mouvement qui étaient séparées en Mécanique classique.

Pour représenter géométriquement l'impulsion d'Univers, nous remarquerons que la tangente en un point de la ligne d'Univers du point matériel a pour cosinus directeurs les quantités $\frac{dx^i}{ds}$ dont voici l'expression

$$\begin{aligned} u^1 &= \frac{dx^1}{ds} = \frac{dx^1}{dt} \frac{dt}{ds} = \frac{v_x}{c \sqrt{1-\beta^2}}, & u^2 &= \frac{v_y}{c \sqrt{1-\beta^2}}, \\ u^3 &= \frac{v_z}{c \sqrt{1-\beta^2}}, & u^4 &= c \frac{dt}{ds} = \frac{1}{\sqrt{1-\beta^2}}. \end{aligned}$$

Les u^i sont les composantes d'un vecteur d'espace-temps, la vitesse d'Univers \mathbf{u} , et l'on vérifie aisément que

$$\mathbf{I} = m_0 c \mathbf{u}.$$

L'impulsion d'Univers s'obtient donc en portant en chaque point de la ligne d'Univers sur la tangente en ce point une longueur égale à $m_0 c$.

La partie de l'intégrale d'action indépendante du champ est $-c \int m_0 c d\tau$: elle est donc, au signe près, la circulation de l'impulsion d'Univers le long de la ligne d'Univers que l'on peut écrire

$$\begin{aligned} \int (\mathbf{I} \cdot d\mathbf{s}) &= \int [I^1 dx^1 - I^2 dx^2 - I^3 dx^3] \\ &= \int (W dt - p_x dx - p_y dy - p_z dz). \end{aligned}$$

Sur cette expression, on lit que l'énergie W joue le rôle de moment conjugué de la variable t .

5. **L'Électromagnétisme relativiste.** — Nous avons vu dans le Chapitre I que les équations de Maxwell sont

$$\begin{aligned} \operatorname{rot} \mathbf{H} &= \frac{1}{c} \frac{\partial \mathbf{b}}{\partial t} + \frac{4\pi}{c} \mathbf{i}, & \operatorname{div} \mathbf{B} &= 0; \\ \operatorname{rot} \mathbf{h} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, & \operatorname{div} \mathbf{b} &= 4\pi \delta. \end{aligned}$$

On peut démontrer qu'il est possible de grouper les champs et les inductions en deux tenseurs antisymétriques du second ordre dont le premier \mathbf{F} réunit le champ électrique et l'induction magnétique et le deuxième \mathbf{G} , le champ magnétique et l'induction électrique. Voici leurs composantes :

$$\begin{aligned} F_{14} &= h_x, & F_{24} &= h_y, & F_{34} &= h_z; \\ F_{23} &= B_x, & F_{31} &= B_y, & F_{12} &= B_z; \\ G_{14} &= b_x, & G_{24} &= b_y, & G_{34} &= b_z; \\ G_{23} &= H_x, & G_{31} &= H_y, & G_{12} &= H_z. \end{aligned}$$

Naturellement, dans le vide, les deux tenseurs \mathbf{F} et \mathbf{G} se confondent.

Les grandeurs

$$s^1 = i_x, \quad s^2 = i_y, \quad s^3 = i_z, \quad s^4 = \delta c$$

sont les composantes d'un vecteur d'espace-temps, le vecteur densité-courant d'électricité. Ceci posé, les équations de Maxwell peuvent s'écrire sous forme tensorielle

$$\begin{aligned} \frac{\partial F^{ij}}{\partial x^k} + \frac{\partial F^{jk}}{\partial x^i} + \frac{\partial F^{ki}}{\partial x^j} &= 0, \\ \sum_j \frac{\partial F^{lj}}{\partial x^j} &= \frac{4\pi}{c} s^l. \end{aligned}$$

Dans les milieux polarisables, la différence entre le champ et l'induction s'explique par l'existence d'un état de polarisation et l'on a

$$\mathbf{b} = \mathbf{h} + 4\pi \mathbf{P}, \quad \mathbf{B} = \mathbf{H} + 4\pi \mathbf{M},$$

où \mathbf{P} est la polarisation électrique (ou densité de moment électrique) et où \mathbf{M} est l'intensité d'aimantation (ou densité de moment magnétique). En faisant la différence des tenseurs \mathbf{F} et \mathbf{G} , on voit que les quantités M_x, M_y, M_z et $-P_x, -P_y, -P_z$ forment les composantes contre-variantes d'un tenseur antisymétrique du second rang, le tenseur Polarisation.

Passons maintenant aux équations de Lorentz qui, avec les champs fins \mathbf{h} et \mathbf{H} , s'écrivent

$$\begin{aligned} \operatorname{rot} \mathbf{h} &= -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t}, & \operatorname{div} \mathbf{H} &= 0; \\ \operatorname{rot} \mathbf{H} &= \frac{1}{c} \frac{\partial \mathbf{h}}{\partial t} + 4\pi \rho \frac{\mathbf{v}}{c}, & \operatorname{div} \mathbf{h} &= 4\pi \rho. \end{aligned}$$

Il suffit ici d'introduire un seul tenseur F antisymétrique de rang 2 tel que

$$\begin{aligned} F^{14} &= h_x, & F^{24} &= h_y, & F^{34} &= h_z; \\ F^{23} &= H_x, & F^{31} &= H_y, & F^{12} &= H_z \end{aligned}$$

et de poser

$$s^1 = \rho v_x, \quad s^2 = \rho v_y, \quad s^3 = \rho v_z, \quad s^4 = \rho c.$$

Les équations de Lorentz s'écrivent alors

$$\frac{\partial F^{ij}}{\partial x^k} + \frac{\partial F^{jk}}{\partial x^i} + \frac{\partial F^{ki}}{\partial x^j} = 0, \quad \sum_j \frac{\partial F^{ij}}{\partial x^j} = \frac{4\pi}{c} s^i.$$

Montrons qu'on peut obtenir les champs \mathbf{h} et \mathbf{H} à partir des potentiels V et \mathbf{A} par

$$\mathbf{H} = \operatorname{rot} \mathbf{A}, \quad \mathbf{h} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \operatorname{grad} V$$

en définissant un vecteur $\vec{\mathcal{X}}$ d'espace-temps, le potentiel d'Univers, dont les composantes sont

$$\mathcal{X}^1 = -\mathcal{X}_1 = A_x, \quad \mathcal{X}^2 = -\mathcal{X}_2 = A_y, \quad \mathcal{X}^3 = -\mathcal{X}_3 = A_z, \quad \mathcal{X}^4 = \mathcal{X}_4 = V.$$

Il suffit en effet de poser

$$F_{ik} = \frac{\partial \mathcal{X}_i}{\partial x^k} - \frac{\partial \mathcal{X}_k}{\partial x^i} \quad \text{ou} \quad \vec{\mathbf{F}} = \operatorname{Rot} \vec{\mathcal{X}}$$

comme on le vérifie aisément.

Le vecteur $\vec{\mathcal{X}}$ n'est d'ailleurs défini qu'à un gradient d'espace-temps près, car si l'on a $\vec{\mathcal{X}}' = \vec{\mathcal{X}} + \operatorname{Grad} \Phi$, on a

$$\operatorname{Rot} \mathbf{P}' = \operatorname{Rot} \mathbf{P} \quad \text{car} \quad \operatorname{Rot} \operatorname{Grad} = 0.$$

Nous avons vu que Lorentz avait restreint l'indétermination des potentiels en leur imposant la condition

$$\frac{1}{c} \frac{\partial V}{\partial t} + \operatorname{div} \mathbf{A} = 0$$

qui s'écrit, en termes tensoriels,

$$\text{Div } \vec{\mathcal{R}} = \sum_k \frac{\partial \mathcal{R}^i}{\partial x^k} = 0.$$

6. Dynamique relativiste du corpuscule électrisé dans un champ électromagnétique. — Nous avons pris (au § 4) comme fonction de Lagrange

$$\mathcal{L} = -m_0 c^2 \sqrt{1 - \beta^2} - U,$$

mais le terme U n'est pas satisfaisant parce que sa variance relativiste n'est pas précise. Nous connaissons actuellement trois sortes de champs : le champ de gravitation, le champ électromagnétique qui agit sur les particules quand elles sont chargées, et enfin le champ nucléaire dont le domaine d'action est extrêmement réduit et qui n'intervient que dans les phénomènes à l'échelle du noyau de l'atome. Nous laisserons ici de côté le champ de gravitation dont l'interprétation relève de la théorie de la Relativité généralisée, ainsi que le champ nucléaire dont la découverte est récente et la nature encore mal connue. Nous nous bornerons à considérer le mouvement d'une charge ponctuelle dans un champ électromagnétique (dynamique relativiste de l'électron).

Dans le cas où dans le système de référence employé le champ est purement électrostatique ($\mathbf{A} = 0$), un corpuscule électrisé de charge ε a une énergie potentielle égale à $U = \varepsilon V$. Mais, du point de vue relativiste, ceci ne peut avoir lieu que dans certains systèmes galiléens au repos les uns par rapport aux autres, puisque V n'est que la composante de temps d'un vecteur d'Univers. En passant à un système galiléen en mouvement par rapport à ceux-ci, il apparaît un champ magnétique et il faut tenir compte de son action sur la charge en mouvement.

Au terme cinétique de l'intégrale d'action $\int -m_0 c ds$, il faut donc ajouter un terme invariant qui se réduise à $-\int \varepsilon V dt$ dans les systèmes galiléens où le champ est purement électrostatique. On voit immédiatement qu'il s'impose de choisir le terme $-\int \frac{\varepsilon}{c} (\vec{\mathcal{R}} \cdot \vec{ds})$, $\vec{\mathcal{R}}$ étant le potentiel d'Univers. Nous aurons donc pour fonction de Lagrange relativiste

$$\mathcal{L} = -m_0 c^2 \sqrt{1 - \beta^2} - \varepsilon V + \frac{\varepsilon}{c} (\mathbf{A} \cdot \mathbf{v}),$$

\mathbf{v} étant la vitesse du corpuscule pour l'observateur galiléen qui emploie le temps t et $(\mathbf{A} \cdot \mathbf{v})$ étant le produit scalaire à trois dimensions

$$A_x v_x + A_y v_y + A_z v_z.$$

Il nous faut vérifier qu'avec ce choix de \mathcal{L} , les équations de Lagrange

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = \frac{\partial \mathcal{L}}{\partial x}, \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{y}} \right) = \frac{\partial \mathcal{L}}{\partial y}, \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{z}} \right) = \frac{\partial \mathcal{L}}{\partial z}$$

nous donnent correctement la Dynamique de l'électron avec la variation relativiste de la masse avec la vitesse.

Nous poserons d'abord comme d'habitude

$$p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}}, \quad p_y = \frac{\partial \mathcal{L}}{\partial \dot{y}}, \quad p_z = \frac{\partial \mathcal{L}}{\partial \dot{z}},$$

ce qui nous donne

$$p_x = \frac{m_0 v_x}{\sqrt{1-\beta^2}} + \frac{\varepsilon}{c} A_x, \quad p_y = \frac{m_0 v_y}{\sqrt{1-\beta^2}} + \frac{\varepsilon}{c} A_y, \quad p_z = \frac{m_0 v_z}{\sqrt{1-\beta^2}} + \frac{\varepsilon}{c} A_z.$$

On voit alors qu'ici le vecteur \mathbf{p} dont les composantes sont les moments de Lagrange est la somme géométrique de la quantité de mouvement $\frac{m_0 \mathbf{v}}{\sqrt{1-\beta^2}}$ d'origine cinétique et d'une sorte de quantité de mouvement « potentielle » égale à $\frac{\varepsilon}{c} \mathbf{A}$.

En portant cette expression des p_x , p_y , p_z dans les équations de Lagrange, nous obtenons, par exemple pour la première de ces équations :

$$\frac{d}{dt} \left[\frac{m_0 v_x}{\sqrt{1-\beta^2}} + \frac{\varepsilon}{c} A_x \right] = -\varepsilon \frac{\partial V}{\partial x} + \frac{\varepsilon}{c} \left(\frac{\partial A_x}{\partial x} v_x + \frac{\partial A_y}{\partial x} v_y + \frac{\partial A_z}{\partial x} v_z \right),$$

ce qui peut aussi s'écrire

$$\begin{aligned} \frac{d}{dt} \left[\frac{m_0 v_x}{\sqrt{1-\beta^2}} \right] &= \varepsilon \left[-\frac{\partial V}{\partial x} - \frac{1}{c} \frac{\partial A_x}{\partial t} \right] + \frac{\varepsilon}{c} \left[v_y \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) - v_z \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \right] \\ &= \varepsilon \left(h_x + \frac{1}{c} [\mathbf{v} \times \mathbf{H}]_x \right) = f_x, \end{aligned}$$

où f_x est la composante x de la force de Lorentz. En tenant compte des deux autres relations analogues, on voit que ceci s'écrit sous forme vectorielle :

$$\frac{d}{dt} \left[\frac{m_0 \mathbf{v}}{\sqrt{1-\beta^2}} \right] = \mathbf{f}.$$

C'est bien l'équation attendue, car elle exprime que la dérivée

par rapport au temps de la quantité de mouvement (définie en tenant compte de la variation de la masse avec la vitesse) est égale à la force.

Comme expression de l'énergie, on trouve aisément, d'après la définition générale de W ,

$$W = (\mathbf{p} \cdot \mathbf{v}) - \mathcal{L} = \frac{m_0 c^2}{\sqrt{1 - \beta^2}} + \varepsilon V.$$

On constate alors aisément que les quatre quantités p_x , p_y , p_z et $\frac{W}{c}$ forment les composantes d'un vecteur d'espace-temps qui est la somme géométrique du vecteur impulsion d'Univers \mathbf{I} défini antérieurement et du vecteur $\frac{\varepsilon}{c} \vec{\mathcal{L}}$.

Rappelons maintenant qu'on nomme « expression hamiltonienne de l'énergie » l'expression de l'énergie en fonction des variables x , y , z , p_x , p_y , p_z , t et qu'on la dénomme $\mathbf{H}(x, y, z, p_x, p_y, p_z, t)$.

En Mécanique classique, en éliminant v_x , v_y , v_z entre les expressions

$$E = \frac{1}{2} m v^2, \quad p_x = m v_x, \quad p_y = m v_y, \quad p_z = m v_z,$$

on trouve

$$E = \mathbf{H}(x, y, z, p_x, p_y, p_z, t) = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + \mathbf{U}(x, y, z, t).$$

En Mécanique relativiste de l'électron, en éliminant v_x , v_y , v_z entre les expressions

$$W = \frac{m_0 c^2}{\sqrt{1 - \beta^2}} + \varepsilon V, \quad \mathbf{p} = \frac{m_0 \mathbf{v}}{\sqrt{1 - \beta^2}} + \frac{\varepsilon}{c} \mathbf{A},$$

on trouve

$$W = \mathbf{H}(x, y, z, p_x, p_y, p_z, t) = c \sqrt{m_0^2 c^2 + \sum_{xyz} \left(p_x - \frac{\varepsilon}{c} \mathbf{A}_x \right)^2} + \varepsilon V.$$

7. Masse transversale et masse longitudinale. — Comme en Mécanique classique on est habitué à employer la formule de Newton $\mathbf{f} = m \boldsymbol{\gamma}$, on a cherché au début du développement de la Mécanique relativiste de l'électron à obtenir des formes analogues, et ceci a amené à distinguer une masse longitudinale et une masse transversale.

Considérons un corpuscule occupant une position M sur sa trajectoire et prenons trois axes rectangulaires tels que l'axe des x soit la tangente en M .

Alors, à l'instant considéré,

$$v_y = v_z = 0.$$

L'équation du mouvement pour l'axe des x nous donne

$$\begin{aligned} f_x &= \frac{d}{dt} \left(\frac{m_0 v_x}{\sqrt{1-\beta^2}} \right) = \frac{m_0 \dot{\gamma}_x}{\sqrt{1-\beta^2}} + \frac{m_0 v_x}{(1-\beta^2)^{\frac{3}{2}}} \dot{\beta} \frac{d\beta}{dt} \\ &= \frac{m_0 \dot{\gamma}_x}{\sqrt{1-\beta^2}} + \frac{m_0 \beta^2}{(1-\beta^2)^{\frac{3}{2}}} \dot{\gamma}_x = \frac{m_0}{(1-\beta^2)^{\frac{3}{2}}} \dot{\gamma}_x. \end{aligned}$$

Par contre, pour les axes transversaux y et z , on obtient plus simplement

$$f_y = \frac{d}{dt} \left(\frac{m_0 v_y}{\sqrt{1-\beta^2}} \right) = \frac{m_0}{\sqrt{1-\beta^2}} \dot{\gamma}_y, \quad f_z = \frac{d}{dt} \left(\frac{m_0 v_z}{\sqrt{1-\beta^2}} \right) = \frac{m_0}{\sqrt{1-\beta^2}} \dot{\gamma}_z,$$

puisque $v_y = v_z = 0$.

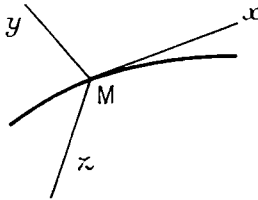


Fig. 14.

Si donc, au lieu d'écrire les équations

$$\frac{dp_x}{dt} = f_x, \quad \frac{dp_y}{dt} = f_y, \quad \frac{dp_z}{dt} = f_z,$$

on cherche à écrire

$$m \dot{\gamma}_x = f_x, \quad m \dot{\gamma}_y = f_y, \quad m \dot{\gamma}_z = f_z,$$

on est conduit à définir deux masses différentes suivant qu'il s'agit du mouvement longitudinal le long de la tangente à la trajectoire ou du mouvement transversal normal à cette tangente. La masse longitudinale m_l et la masse transversale m_t ont pour expressions

$$m_l = \frac{m_0}{(1-\beta^2)^{\frac{3}{2}}}, \quad m_t = \frac{m_0}{\sqrt{1-\beta^2}}.$$

Cette dernière coïncide avec la masse en mouvement précédemment définie. On évite la complication de définir deux masses et l'on peut se servir uniquement de la masse m_l si l'on abandonne l'expression $\mathbf{f} = m \dot{\boldsymbol{\gamma}}$ pour ne conserver que la forme $\mathbf{f} = \frac{d\mathbf{p}}{dt}$ qui est, elle, aussi bien valable dans la Mécanique relativiste que dans la Mécanique classique.

8. Vérifications expérimentales de la théorie de la Relativité restreinte. — Les vérifications expérimentales de la théorie de la Relativité restreinte sont extrêmement nombreuses. Il y a d'abord la célèbre expérience de Michelson et Morley plusieurs fois refaite dans des conditions différentes, les expériences de Trouton et Noble et celles de Röntgen et Eichenwald montrant l'insensibilité au mouvement de translation de divers phénomènes électromagnétiques. Ces expériences sont exposées dans les traités de Relativité, notamment dans le livre de Laue indiqué dans la bibliographie du précédent chapitre.

Les expériences de Fizeau [1], reprises par Zeeman [2], sur ce qu'on nommait naguère l'entraînement partiel de l'éther par le mouvement des corps réfringents, peuvent, nous l'avons vu, être considérées comme une confirmation de la cinématique relativiste. Le phénomène du ralentissement des horloges a été mis très directement en évidence par Ives et Stilwell [3] au moyen de l'effet Doppler et par les calculs de Rossi et Hall et de Rossi et Nereson [4] relatifs à la vie moyenne des mésons. La variation de la masse avec la vitesse prévue par la Dynamique relativiste a été démontrée expérimentalement par les expériences de Guye et Lavanchy [5] et par celles plus récentes de Nacken [6]. Nous verrons qu'elle a aussi été confirmée indirectement par le succès de la théorie de Sommerfeld et par la vérification des lois de l'effet photoélectrique et de la diffraction des électrons dans le cas des grandes énergies. Enfin, les vérifications du principe de l'inertie de l'énergie dans le domaine des phénomènes nucléaires sont aujourd'hui innombrables.

On trouvera une vue d'ensemble sur toutes ces vérifications de la théorie de la Relativité restreinte et de quelques autres dans un très bel article de M. Robert Lennuyer [7].

BIBLIOGRAPHIE.

- [1] FIZEAU, *C. R. Acad. Sc.*, t. 33, 1851, p. 329.
- [2] P. ZEEMAN, *Archives néerlandaises des Sciences exactes et naturelles*, t. X, 1927, p. 132.
- [3] H. E. IVES et G. R. STILWELL, *J. Opt. Soc. Am.*, t. 28, 1938, p. 369.
- [4] ROSSI et HALL, *Phys. Rev.*, t. 54, 1941, p. 223; NERESON et ROSSI, *Phys. Rev.*, t. 64, 1943, p. 199.
- [5] GUYE et LAVANCHY, *Annales des Sciences physiques et naturelles de Genève*, t. 42, 1916, p. 286.
- [6] M. NACKEN, *Ann. Physik*, t. 25, 1935, p. 313.
- [7] R. LENNUYER, *Revue scientifique*, 85^e année, 1947, p. 740.

CHAPITRE IV.

LA MÉCANIQUE STATISTIQUE CLASSIQUE.

1. But de la Mécanique statistique classique. — Le but de la Mécanique statistique classique développée notamment par Clausius, par Maxwell, par Boltzmann et par Gibbs est d'étudier les propriétés statistiques moyennes des systèmes très complexes définis par un nombre extrêmement grand de paramètres. Son principal succès est d'être parvenue à interpréter les lois de la Thermodynamique de telle façon qu'elles apparaissent comme une conséquence du fait que la Thermodynamique envisage toujours les propriétés moyennes globales des corps macroscopiques dont la description détaillée exacte exigerait l'emploi d'un nombre énorme de paramètres. Par exemple, la Thermodynamique traite des propriétés globales des gaz et, aux yeux de la Mécanique statistique, un gaz est formé par un nombre immense d'atomes ou de molécules, chacun de ces éléments étant lui-même en général décrit à l'aide de plusieurs paramètres. Les lois des gaz apparaissent alors comme des lois statistiques résultant globalement des mouvements incoordonnés d'innombrables molécules.

En Mécanique statistique classique, on admet que les éléments en nombre immense (atomes ou molécules) dont sont formés les corps matériels obéissent aux lois de la Mécanique classique, de sorte que, s'il était possible de connaître exactement à un instant donné les positions et les vitesses de tous ces éléments, on pourrait en principe calculer rigoureusement toute leur histoire ultérieure. Mais, en fait, on ne peut observer les évolutions des molécules et on ne constate que des effets statistiques moyens dus à l'ensemble de ces évolutions et obéissant à des lois de probabilité dont la Mécanique statistique se propose de déterminer la forme. Ainsi, dans cette théorie classique, on admet l'existence d'un déterminisme mécanique sous-jacent, le caractère « probabiliste » des lois obtenues résultant uniquement

de l'impossibilité d'observer autre chose que des effets globaux. Il en est tout autrement dans les théories quantiques actuelles où la probabilité s'introduit jusque dans la Mécanique des particules *individuelles* sans qu'aucun déterminisme du type classique puisse subsister, même à cette échelle.

2. Extension-en-phase. Théorème de Liouville. — Pour commencer, nous considérons un système matériel dont la configuration est définie par un certain nombre de paramètres $q_1, \dots, q_i, \dots, q_N$. En Mécanique statistique, nous serons généralement amené à admettre que le nombre N est énorme; par exemple, pour un gaz monoatomique formé par un nombre immense n d'atomes assimilés à des points matériels, on aura $N = 3n$, ce qui sera un nombre énorme. Néanmoins pour l'instant, nous ne ferons aucune hypothèse sur l'ordre de grandeur de N .

D'une façon générale, nous admettons que notre système obéit aux lois de la Mécanique classique exprimées par les équations canoniques de Hamilton. Pour écrire ces équations, nous devons introduire les moments de Lagrange p_1, \dots, p_N respectivement conjugués de q_1, \dots, q_N et définis par $p_i = \frac{\partial T}{\partial \dot{q}_i}$, où T est l'énergie cinétique totale du système et \dot{q}_i la dérivée de q_i par rapport au temps. L'énergie du système s'exprime en fonction des q_i et des p_i par la fonction hamiltonienne du système $H(q_1, \dots, q_N, p_1, \dots, p_N, t)$ qui peut naturellement dépendre explicitement du temps si les actions extérieures sont variables. Les équations de Hamilton sont alors

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \quad (i = 1, 2, \dots, N).$$

Si le système n'est pas soumis à des actions extérieures variables, et *a fortiori* s'il est isolé, l'énergie H est une constante; en effet, on a

$$\frac{dH}{dt} = \sum_i^N \left(\frac{\partial H}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial H}{\partial p_i} \frac{dp_i}{dt} \right) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}$$

et $\frac{\partial H}{\partial t} = 0$ si H ne dépend pas explicitement du temps.

L'état dynamique du système à un certain instant est entièrement défini à chaque instant par la donnée des q_i et des p_i . On peut donc représenter cet état par un point dans un espace à $2N$ dimensions formé des q_i et des p_i : c'est l'*extension-en-phase* de Gibbs. Au cours du temps, le point figuratif du système décrit une certaine courbe,

une trajectoire, dans l'extension-en-phase : cette trajectoire représente l'évolution du système au cours du temps.

Considérons maintenant un petit élément de volume $d\tau$ dans l'extension-en-phase : il correspond à de petites variations dq_1, \dots, dp_N des variables q_1, \dots, p_N . Un tel élément de volume possède deux propriétés qui lui donnent une grande importance en Mécanique statistique.

La première de ces propriétés est la suivante : si l'on opère un changement de variables faisant passer des variables canoniques q_1, \dots, p_N à de nouvelles variables canoniques Q_1, \dots, P_N (pour lesquelles les équations de Hamilton sont également satisfaites), les éléments de volume qui se correspondent dans les deux extensions-en-phase sont égaux. Cette propriété, que je me contenterai d'énoncer, résulte de la façon dont sont définies les variables conjuguées q_i et p_i et montre que l'élément de volume d'extension-en-phase a une signification intrinsèque indépendante du choix particulier des variables canoniques servant à définir le système.

Mais c'est surtout la deuxième propriété de l'élément $d\tau$ qui est essentielle pour la Mécanique statistique : elle s'exprime par le « *théorème de Liouville* ». Pour énoncer ce théorème, nous allons considérer, non plus un seul exemplaire de notre système à N degrés de liberté, mais un grand nombre d'exemplaires de ce système. A un instant t_0 , chacun de ces exemplaires est représenté par un point de l'extension-en-phase et à l'intérieur de l'élément $d\tau$ qui nous intéresse, il y a un certain nombre de ces points représentatifs. Fixons notre attention sur le nuage des points représentatifs qui à l'instant t occupent l'élément $d\tau$. Au cours du temps, ces points vont se déplacer et, à un instant postérieur t' , il y aura encore un élément de volume $d\tau'$ de l'extension-en-phase qui contiendra ces points et ne contiendra que ceux-là. Le théorème de Liouville s'exprime par l'égalité $d\tau' = d\tau$. L'élément $d\tau'$ pourra avoir une forme tout à fait différente de l'élément $d\tau$, être par exemple un long ruban replié sur lui-même alors que $d\tau$ était un petit cube : les volumes des deux éléments seront néanmoins égaux. On peut démontrer le théorème de Liouville en remarquant que le mouvement du nuage de points représentatifs dans l'extension-en-phase est comparable au mouvement des molécules d'un fluide dans un espace à $2N$ dimensions : le théorème signifie donc qu'un même nombre de molécules de ce fluide occupe toujours le même volume au cours du mouvement, c'est-à-dire que le fluide se comporte comme un fluide incompressible. Or, la condition d'incompressibilité d'un fluide est que

la divergence de sa vitesse soit nulle, et cette condition reste valable pour un fluide à $2N$ dimensions où la vitesse a $2N$ composantes. Pour notre fluide fictif dans l'extension-en-phase, les $2N$ composantes de vitesses d'une molécule sont évidemment

$$\frac{dq_1}{dt}, \dots, \frac{dq_N}{dt}, \frac{dp_1}{dt}, \dots, \frac{dp_N}{dt}$$

et le théorème sera démontré si l'on peut établir que

$$\sum_i^N \left[\frac{\partial}{\partial q_i} \left(\frac{dq_i}{dt} \right) + \frac{\partial}{\partial p_i} \left(\frac{dp_i}{dt} \right) \right] = 0.$$

Or ceci résulte immédiatement des équations de Hamilton.

Le théorème de Liouville montre qu'une répartition uniforme de points représentatifs dans l'extension-en-phase reste indéfiniment uniforme. Il est évident que ce résultat nous incite à prendre l'élément d'extension-en-phase comme mesure de la probabilité *a priori* pour que notre système se trouve avoir son point représentatif dans l'élément $d\tau$.

Il y a cependant quelques remarques à faire à ce sujet. Tout d'abord, il arrive très souvent qu'un système mécanique admette des intégrales premières, c'est-à-dire évolue de façon qu'une certaine fonction des q_i et des p_i reste constante : un exemple très usuel est celui d'un système isolé dont l'énergie $H(q_i, p_i)$ reste constante. Le point représentatif du système est alors assujéti à se déplacer sur une multiplicité à moins de $2N$ dimensions de l'extension-en-phase (par exemple sur une multiplicité à $2N-1$ dimensions dans le cas d'une seule intégrale première). Pour un système isolé n'ayant que l'énergie comme intégrale première uniforme, le point représentatif se déplacera sur l'hyper-surface $H = E$ ou, plus exactement, comme la valeur de l'énergie n'est jamais connue qu'avec une certaine incertitude dE , il est contenu dans une couche de l'extension-en-phase limitée par les hypersurfaces $H = E$ et $H = E + dE$: ce sont les éléments de volume $d\tau$ de cette couche qui doivent servir à l'évaluation des probabilités.

Pour achever de justifier le rôle attribué à l'élément d'extension-en-phase, les créateurs de la Mécanique statistique avaient admis le postulat suivant :

Quand un système admet comme seule intégrale première uniforme l'énergie et qu'il part d'un état d'énergie comprise entre E et $E + dE$, on peut admettre qu'au bout d'un temps suffisamment

long son point représentatif a balayé d'une façon uniforme toute la couche d'extension-en-phase comprise entre les multiplicités $H = E$ et $H = E + dE$.

C'est là « l'hypothèse ergodique » qu'on peut énoncer sous une forme un peu moins stricte constituant « l'hypothèse quasi ergodique ». Il est impossible de justifier rigoureusement ces hypothèses : il y a d'ailleurs des cas simples où elles sont manifestement en défaut comme, par exemple, le cas d'un mouvement périodique où le point représentatif reste cantonné sur une multiplicité à une dimension. On peut cependant admettre que ces cas exceptionnels ont une probabilité évanouissante, mais néanmoins les hypothèses ergodiques ou quasi ergodiques sont très difficiles à justifier dans le cadre des théories classiques : cette difficulté est liée au caractère trop rigoureux du déterminisme mécanique dans ces théories et s'atténue dans les théories quantiques. Sans nous arrêter ici sur ces difficultés, nous admettrons que l'hypothèse ergodique est pratiquement vérifiée ; par suite, un système admettant uniquement l'énergie comme intégrale première, étant parti d'un état d'énergie comprise entre E et $E + dE$, la probabilité de trouver son point représentatif au bout d'un temps très long dans un élément de volume $\Delta\tau$ de la couche $(E, E + dE)$ de l'extension-en-phase sera prise proportionnelle à $\Delta\tau$.

3. Entropie et probabilité. Relation de Boltzmann. — La grandeur la plus caractéristique et la plus mystérieuse introduite par la Thermodynamique était l'entropie qui a une tendance à toujours augmenter dans toutes les transformations physiques spontanées. Le grand succès de la Mécanique statistique a été de parvenir à interpréter l'entropie de l'état d'un corps comme mesurant le degré de probabilité de cet état. Il est facile de pressentir quelle doit être la relation fonctionnelle entre l'entropie d'un état et sa probabilité. En effet, si l'on considère deux systèmes sans interaction dont les entropies sont S_1 et S_2 , l'entropie totale du système formé par l'ensemble des deux systèmes est $S_1 + S_2$. Or, si P_1 est la probabilité de l'état du premier système et P_2 celle de l'état du second système, l'hypothèse de l'indépendance des deux systèmes et le théorème des probabilités composées montrent que la probabilité de l'état global du système $(1 + 2)$ est $P_1 P_2$. La liaison entre l'entropie et la probabilité doit donc être telle qu'elle fasse correspondre le produit des probabilités à la somme des entropies et l'on est ainsi conduit à poser

$$S = k \log P.$$

C'est la fameuse relation de Boltzmann où k est la « constante de Boltzmann » dont nous pourrions plus loin préciser la valeur. Pour P , on prendra une quantité proportionnelle au volume d'extension-en-phase correspondant à l'état du système.

En Thermodynamique, on définit la variation de l'entropie d'un corps dont l'énergie interne varie de dE et qui fournit le travail extérieur $d\mathfrak{E}$ par la formule

$$dS_{\text{th}} = \frac{dQ}{T} = \frac{dE + d\mathfrak{E}}{T}.$$

Nous écrivons ici S_{th} , car l'entropie thermodynamique définie seulement pour des états d'équilibre est moins générale que l'entropie S de Boltzmann définie pour un état quelconque et n'en est qu'un cas particulier.

Si le corps ne fournit aucun travail $dS_{\text{th}} = \frac{dE}{T}$ et comme l'entropie S_{th} est fonction de E et d'autres paramètres (par exemple le volume V occupé par le corps), on conclut de la relation précédente que l'on peut définir la température absolue d'un corps par la relation

$$\frac{1}{T} = \frac{\partial S_{\text{th}}}{\partial E}.$$

Quand deux corps 1 et 2 sont en contact et peuvent échanger de la chaleur, mais pas de travail mécanique macroscopique, l'expérience montre que les températures des deux corps tendent à s'égaliser et, quand l'équilibre thermique est atteint, on a $\left(\frac{\partial S_{\text{th}}}{\partial E}\right)_1 = \left(\frac{\partial S_{\text{th}}}{\partial E}\right)_2$.

Examinons maintenant cette question de l'équilibre thermique du point de vue de la Mécanique statistique. Le système 1 + 2 est supposé isolé : il a donc une énergie constante E . Cette énergie est répartie entre les corps 1 et 2 et l'on a $E_1 + E_2 = E$. Les deux corps étant supposés sans interaction, la probabilité de l'état de (1 + 2), où le premier corps a l'énergie E_1 et le second l'énergie E_2 est

$$P = P_1(E_1) \cdot P_2(E - E_1).$$

L'état *le plus probable* du système global est donné par le maximum de P défini par la condition

$$\frac{\partial \log P}{\partial E_1} = 0 \quad \text{ou} \quad \frac{\partial \log P_1}{\partial E_1} + \frac{\partial \log P_2}{\partial E_1} = 0,$$

ou encore, comme $dE_2 = -dE_1$,

$$\frac{\partial \log P_1}{\partial E_1} = \frac{\partial \log P_2}{\partial E_2}.$$

En multipliant par k et en introduisant la relation de Boltzmann, il vient

$$\frac{\partial S_1}{\partial E_1} = \frac{\partial S_2}{\partial E_2},$$

et cette équation exprime l'égalité des températures au moment de l'équilibre.

Ces considérations nous permettent de préciser nos idées en disant que l'entropie S_{th} définie par la Thermodynamique classique correspond à la probabilité P_m de l'état le plus probable. On a donc

$$S_{th} = k \log P_m, \quad \frac{\partial \log P_m}{\partial E} = \frac{1}{kT},$$

P_m étant l'état de probabilité maximum du corps considéré pour l'énergie E . La Thermodynamique statistique va donc plus loin que la Thermodynamique ancienne, car elle permet de définir la probabilité d'un état quelconque d'un système et, par suite, son entropie grâce à la relation générale de Boltzmann alors que la Thermodynamique classique ne définissait que les entropies des états de probabilité maximum qui d'ailleurs, en raison du grand nombre de variables intervenant dans la spécification des états macroscopiques, sont en général énormément plus probables que tous les autres.

4. La loi de répartition canonique de Gibbs. — Considérons toujours le système formé par la réunion de deux corps 1 et 2. Nous supposons maintenant que le corps 2 est un très grand réservoir de chaleur, c'est-à-dire d'énergie moléculaire, un « thermostat », de telle façon que le corps 1 n'ait qu'une chance tout à fait négligeable de lui prendre une fraction appréciable de son énergie. En d'autres termes, si E_2^m est l'énergie du corps 2 quand il est dans l'état d'équilibre le plus probable avec le corps 1, on peut admettre que pour tous les états possibles du corps 2, les énergies correspondantes sont telles que $E_2 - E_2^m$ est énormément plus petit que E_2^m . Nous supposons toujours que 1 et 2 peuvent échanger de la chaleur, mais que leurs états sont indépendants. La probabilité de l'état du corps 2 étant P_2 , on peut écrire

$$\log P_2 = \log P_2^m + \frac{\partial \log P_2^m}{\partial E_2} (E_2 - E_2^m) + \dots$$

Les termes non écrits sont d'ordre supérieur au premier en $E_2 - E_2^m$

et l'hypothèse faite permet de les négliger. Si T est la température absolue du thermostat, on aura

$$\log P_2 = \log P_2''' + \frac{1}{kT} (E_2 - E_2'''),$$

d'où

$$P_2 = P_2''' e^{\frac{E_2 - E_2'''}{kT}}.$$

Mais si E_1''' désigne l'énergie du corps 1 quand il est dans son état le plus probable d'équilibre avec le thermostat, on a

$$E_2 - E_2''' = E_1''' - E_1$$

d'après la conservation de l'énergie. La probabilité de l'état global où le thermostat a l'énergie E_2 et le corps 1 l'énergie E_1 est donc

$$P = P_1(E_1) \cdot P_2(E_2) = P_1(E_1) P_2''' e^{\frac{E_2''' - E_1}{kT}}.$$

La probabilité pour que le corps 1 *en contact avec le thermostat* ait l'énergie E_i est donc de la forme

$$P = P_1(E_i) e^{\frac{\Psi - E_i}{kT}},$$

où $P_1(E_i)$ représente la probabilité de l'énergie E_i envisagée *a priori*, c'est-à-dire sans imposer au corps 1 aucune liaison avec un thermostat. La constante Ψ peut se déterminer en écrivant que la somme des probabilités de toutes les valeurs possibles de E_i est l'unité

$$\sum_i P_1(E_i) e^{\frac{\Psi - E_i}{kT}} = 1$$

ou

$$e^{-\frac{\Psi}{kT}} = \sum_i P_1(E_i) e^{-\frac{E_i}{kT}} = Z.$$

La somme Z s'appelle la « *somme d'états* » (Planck).

Les énergies E_i du système 1 peuvent d'ailleurs former une suite continue ou une suite discontinue (états quantifiés). Les formules précédentes sont valables dans les deux cas à condition de remplacer le signe Σ par une intégration dans le cas continu.

Nous allons démontrer que, *si le corps 1 est un corps très complexe défini par un nombre énorme de paramètres*, la quantité Ψ est son potentiel thermodynamique $E - TS$ au sens classique. En effet, dans ce cas, le corps 1 dont l'état macroscopique est défini par la tempé-

rature T du thermostat et par d'autres variables (par exemple son volume V) possède une entropie thermodynamique $S_{th}(T, \dots)$ et une énergie interne $E(T, \dots)$ qui, nous l'avons vu, correspond à l'état le plus probable du système r . De plus, le système r étant supposé défini par un très grand nombre de paramètres, on peut montrer, comme nous l'avons déjà signalé, que son état le plus probable est infiniment plus probable que tous les autres. Dès lors, dans la somme d'état Z il suffit de conserver le terme se rapportant à l'état le plus probable et d'écrire

$$P_1(E_1^m) e^{-\frac{E_1^m}{kT}} = e^{-\frac{\Psi}{kT}}$$

et comme d'après la relation de Boltzmann,

$$S_{th}(E_1^m) = k \log P_1(E_1^m),$$

il vient

$$\frac{S_{th}}{e^k} e^{-\frac{E_1^m}{kT}} = e^{-\frac{\Psi}{kT}},$$

d'où

$$\Psi = E_1^m - TS_{th}(E_1^m) = E - TS;$$

Ψ est donc bien en ce cas le potentiel thermodynamique.

Pour le cas d'un corps r défini par un petit nombre de paramètres, la loi de distribution canonique de Gibbs est toujours valable, mais Ψ n'a plus le sens d'un potentiel thermodynamique.

Considérons par exemple une molécule dans un gaz; elle peut être considérée comme en contact avec un thermostat formé par le reste du gaz et l'on pourra lui appliquer la loi de distribution canonique. Or la probabilité *a priori* pour que les coordonnées et les moments de Lagrange de cette molécule aient des valeurs comprises dans les intervalles $(x, x + dx), \dots, (p_z, p_z + dp_z)$ est, d'après le théorème de Liouville, proportionnelle à l'élément d'extension-en-phase $dx dy dz dp_x dp_y dp_z$. La probabilité de l'état d'énergie E de la molécule dans le gaz s'obtient donc en multipliant cette expression par $e^{-\frac{E}{kT}}$ et le nombre des molécules ayant une énergie correspondant à l'intervalle $dp_x dp_y dp_z$ et contenues dans le volume $dx dy dz$ est

$$dn = C e^{-\frac{E}{kT}} dx dy dz dp_x dp_y dp_z = C m^3 e^{-\frac{E}{kT}} dx dy dz dv_x dv_y dv_z.$$

La valeur de la constante s'obtient en écrivant que $\int dn$ est égal au nombre total N des molécules du gaz.

Nous avons ainsi obtenu la célèbre formule de distribution des énergies moléculaires due à Maxwell. Supposons que les molécules ne soient soumises à aucun champ; on a

$$E = \frac{1}{2} m (v_x^2 + v_y^2 + v_z^2) = \frac{p^2}{2m}.$$

Comme

$$dp_x dp_y dp_z = 4\pi p^2 dp = 2\pi (2m)^{\frac{3}{2}} \sqrt{E} dE,$$

on obtient pour le nombre de molécules par unité de volume ayant une énergie comprise entre E et $E + dE$

$$dn = C e^{-\frac{E}{kT}} \sqrt{E} dE.$$

Si le gaz est placé dans un champ de force, il faut tenir compte de l'énergie potentielle dans l'expression de E . Par exemple, le gaz étant placé dans le champ de la pesanteur

$$E = \frac{p^2}{2m} + mgz$$

et en intégrant sur $dp_x dp_y dp_z$ l'expression de dn , on voit que la densité du gaz diminue avec l'altitude z comme $e^{-\frac{mgz}{kT}}$, ce qui conduit à la loi barométrique de Laplace.

La formule de Boltzmann permet aussi d'étudier les fluctuations statistiques de l'état d'un corps autour de son état le plus probable. Nous ne pouvons développer ici cette étude malgré son très grand intérêt.

§. Application au cas d'un gaz parfait. — Pour donner un exemple concret des considérations générales précédentes, nous envisagerons le cas simple et classique du « gaz parfait » formé de molécules qu'on assimile à des points matériels et qui ne réagissent entre elles qu'au moment de leur choc. En cherchant la répartition la plus probable des énergies entre les molécules, nous allons retrouver par une méthode nouvelle la loi de répartition de Maxwell.

Pour développer ce calcul, considérons un élément $d\tau$ d'extension-en-phase $dx dy dz dp_x dp_y dp_z$ de la molécule. A cet élément correspond un certain nombre d'états possibles pour la molécule qui, d'après le théorème de Liouville et nos hypothèses générales, est de la forme $\frac{dx dy dz dp_x dp_y dp_z}{a}$, où a est une constante. Autrement dit,

nous admettons que tout se passe comme si l'extension-en-phase devait être divisée en *cellules* de grandeur a , chacune de ces cellules contenant un point représentatif et un seul correspondant à un état possible de la molécule : c'est là un point de vue dont nous verrons plus tard l'importance dans la Mécanique statistique quantique. En Mécanique statistique classique, ce point de vue paraît un peu artificiel car tout point de l'extension-en-phase représente un état possible *a priori* et alors a doit être pris infiniment petit; mais nous pouvons toujours raisonner comme si a était fini et le faire tendre vers zéro dans le résultat final.

Considérons alors un certain état du gaz. Soit n_i le nombre des molécules qui, dans cet état du gaz, ont des énergies comprises entre E_i et $E_i + dE_i$. L'énergie totale du gaz E est $\sum_i n_i E_i$, la somme représentant d'ailleurs ici une intégrale; E est une constante si nous supposons le gaz isolé, c'est-à-dire contenu dans un récipient à parois fixes et adiathermanes. Soit g_i le nombre des états possibles *a priori* de la molécule considérée correspondant à l'intervalle $(E_i, E_i + \Delta E_i)$; il est égal au quotient par a de la couche d'extension-en-phase correspondant à cet intervalle d'énergie, soit à

$$g_i = \frac{V \cdot 4\pi p_i^2 dp_i}{a}.$$

La probabilité de l'état considéré du gaz est proportionnelle au nombre de manières de répartir les N molécules du gaz entre les divers intervalles d'énergie, soit à $\frac{N!}{n_1! \dots n_i! \dots}$; elle est aussi proportionnelle au nombre de manières égal à $(g_i)^{n_i}$, de répartir les n_i molécules de chaque intervalle entre les g_i états correspondants. On a donc

$$P = \text{const.} \cdot N! \prod_i \frac{g_i^{n_i}}{n_i!} = \frac{N!}{a^N} \prod_i \frac{(V \cdot 4\pi p_i^2 \Delta p_i)^{n_i}}{n_i!}.$$

Les ΔE_i sont des infiniment petits physiques suffisamment petits pour qu'on puisse remplacer les sommes par des intégrales, mais suffisamment grands pour qu'on puisse supposer que les n_i sont grands. Dès lors, une formule classique, la formule de Stirling, nous apprend que nous pouvons avec une bonne approximation remplacer $\log n_i!$ par $n_i(\log n_i - 1)$. Il vient alors

$$\log P = \sum_i n_i [\log g_i - \log n_i + 1] + \text{const.}$$

Cherchons à déterminer l'état le plus probable du gaz dans les conditions qui lui sont imposées, état qu'envisage toujours la Thermodynamique classique. Pour le déterminer nous devons écrire que P ou $\log P$ est maximum, l'énergie totale du gaz et le nombre total de ses molécules ayant des valeurs données E et N . Nous écrirons donc

$$0 = \delta \log P = \sum_i \left[\delta n_i (\log g_i - \log n_i + 1) - n_i \frac{1}{n_i} \delta n_i \right] = \sum \delta n_i (\log g_i - \log n_i)$$

avec les conditions

$$\delta N = \sum_i \delta n_i = 0 \quad \text{et} \quad \delta E = \sum_i E_i \delta n_i = 0$$

Pour tenir compte de ces conditions, nous emploierons la méthode des multiplicateurs de Lagrange et, en introduisant deux multiplicateurs α et β , nous écrirons que l'on doit avoir pour une variation quelconque des δn_i

$$\delta \log P - \alpha \delta N - \beta \delta E = \sum_i \delta n_i [\log g_i - \log n_i - \alpha - \beta E_i] = 0,$$

ce qui conduit à annuler les crochets et donne

$$n_i = g_i e^{-\alpha - \beta E_i}.$$

On retrouve donc la loi de Maxwell si $\beta = \frac{1}{kT}$; nous verrons ci-dessous qu'il en est bien ainsi.

Portant l'expression trouvée pour les n_i dans celle de $\log P$, on trouve à une constante près

$$\log P_m = \sum_i n_i \left[\log \frac{g_i}{n_i} + 1 \right] = \sum_i n_i [\alpha + \beta E_i + 1] = (\alpha + 1) N + \beta E,$$

d'où, d'après la relation de Boltzmann,

$$S_m = k \log P_m = kN(\alpha + 1) + k\beta E,$$

où α est par sa définition même indépendant de E . Pour calculer β , supposons pour un instant que notre gaz soit contenu dans un récipient à parois fixes, mais diathermanes c'est-à-dire qu'il puisse échanger de la chaleur (mais non du travail) avec le milieu extérieur : un tel échange de chaleur fera varier E de δE , les autres paramètres fixant l'état du gaz (paramètres dont α dépend aussi) restant fixes et l'on aura

$$\delta S = k\beta \delta E + kE \delta\beta + kN \delta\alpha,$$

mais la condition $\delta N = 0$ nous donne

$$\sum_i g_i e^{-\alpha - \beta E_i} (-\delta\alpha - E_i \delta\beta) \quad \text{ou} \quad N\delta\alpha + E\delta\beta = 0,$$

d'où

$$\delta S = k\beta \delta E$$

et puisque $\frac{1}{T} = \frac{\delta S}{\delta E}$, on trouve bien, comme nous l'avions annoncé,

$$\beta = \frac{1}{kT}.$$

La constante α se calcule en écrivant $\sum_i n_i = N$, ce qui donne aisément

$$\begin{aligned} \alpha &= \log \sum_i g_i e^{-\frac{E_i}{kT}} - \log N = \log \frac{V}{N} \int_0^\infty e^{-\frac{p^2}{2mkT}} \frac{4\pi p^2 dp}{a} \\ &= \log \left[\frac{V}{aN} (2\pi mkT)^{\frac{3}{2}} \right]. \end{aligned}$$

En portant cette valeur de α ainsi que $\beta = \frac{1}{kT}$ dans l'expression de S_{th} , on trouve

$$S_{\text{th}} = \frac{E}{T} + kN \log \left[\frac{Ve}{aN} (2\pi mkT)^{\frac{3}{2}} \right],$$

d'où pour le potentiel thermodynamique

$$\Psi = E - TS_{\text{th}} = -kNT \log \left[\frac{Ve}{aN} (2\pi mkT)^{\frac{3}{2}} \right].$$

Une formule classique en thermodynamique donne alors pour la pression du gaz

$$p = - \left(\frac{\partial \Psi}{\partial V} \right)_T = \frac{kNT}{V} \quad \text{ou} \quad pV = kNT.$$

C'est la loi de Mariotte-Gay-Lussac. Appliquons cette formule à une molécule-gramme du gaz pour laquelle N est égal au nombre d'Avogadro \mathcal{N} , nous obtenons

$$k = \frac{R}{\mathcal{N}}.$$

La constante k de Boltzmann est donc égale à la constante des gaz rapportée à la molécule. Numériquement, dans le système C. G. S. avec l'échelle des températures absolues

$$k = \frac{8,31 \cdot 10^7}{6,06 \cdot 10^{23}} = 1,37 \cdot 10^{-16}.$$

Calculons maintenant l'énergie totale E du gaz en fonction de T . Nous trouvons

$$\begin{aligned} E &= \sum_i E_i n_i = e^{-\alpha} \frac{V}{a} \int_0^\infty \frac{p^2}{2m} e^{-\frac{p^2}{2mkT}} 4\pi p^2 dp \\ &= e^{-\alpha} \frac{V}{a} (2\pi mkT)^{\frac{3}{2}} \frac{3}{2} kT = N \frac{3}{2} kT, \end{aligned}$$

résultat important qui nous apprend que l'énergie moyenne par molécule $\frac{E}{N}$ est égale à $\frac{3}{2} kT$. Il vient alors

$$S_{th} = kN(1 + \alpha) + k\beta E = kN \left(\alpha + \frac{5}{2} \right) = kN \log V + \frac{3}{2} kN \log T + \text{const.},$$

d'où pour une molécule-gramme

$$S_{th} = R \log V + \frac{3}{2} R \log T + \text{const.}$$

Or, pour un gaz parfait, on démontre dans tous les Traités de Thermodynamique que

$$S_{th} = R \log V + c \log T + \text{const.},$$

où c est la chaleur spécifique à volume constant telle que $E = cT$. Pour notre gaz parfait à molécules ponctuelles, nous avons $c = \frac{3}{2}R$, ce qui montre l'identité des deux dernières expressions de S_{th} .

Notons que nos formules nous donneraient une valeur précise de la constante de l'entropie si nous connaissions la valeur de a . Les théories quantiques qui posent $a = h^3$ permettront donc de calculer cette constante comme la théorie de la Relativité permet de calculer la constante de l'énergie. Mais en théorie classique, on doit faire tendre a vers zéro dans les résultats, et ceci donne une constante infinie négative dans l'expression de l'entropie. Au point de vue classique, on peut essayer de se débarrasser de cette difficulté en disant qu'après tout, l'entropie thermodynamique n'étant définie qu'à une constante près, il importe peu que cette constante soit infinie, mais c'est là évidemment une bien mauvaise échappatoire.

6. Théorème de l'équipartition de l'énergie. — Dans l'expression de l'énergie d'une molécule ponctuelle, les moments de Lagrange figurent par leurs carrés, car $E = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2)$. Dans l'expression de l'énergie potentielle d'un point matériel rappelé vers une position d'équilibre (prise comme origine des coordonnées) par une force

proportionnelle à l'élongation, les coordonnées figurent par leurs carrés, car

$$E = \frac{1}{2m} p^2 + \frac{K}{2} (x^2 + y^2 + z^2).$$

D'une façon générale, nous dirons qu'une variable canonique q_i ou p_i est un « momentoïde » si elle figure par son carré dans l'expression hamiltonienne de l'énergie du système. L'exemple le plus simple d'un momentoïde est un moment de Lagrange en coordonnées cartésiennes dans le cas d'un corpuscule libre : de là le nom même de momentoïde.

Le théorème de l'équipartition s'énonce alors ainsi :

« Si l'une des variables canoniques d'un système est un momentoïde, quand le système se trouve en équilibre thermique à la température absolue T, le terme correspondant à ce momentoïde dans l'expression de l'énergie a pour valeur moyenne $\frac{1}{2} kT$. »

Ainsi l'énergie se répartit uniformément en moyenne entre tous les momentoïdes, d'où le nom du théorème.

Démontrons le théorème en supposant que la variable q_k soit un momentoïde. Alors l'énergie du système est de la forme

$$E = \alpha q_k^2 + f(q_1, \dots, q_{k-1}, q_{k+1}, \dots, q_N, p_1, \dots, p_N),$$

où α est une constante. La valeur moyenne du terme αq_k^2 dans l'équilibre thermique est

$$\overline{\alpha q_k^2} = \frac{\int \dots \int \alpha q_k^2 e^{-\frac{\alpha q_k^2}{kT}} e^{-\frac{f}{kT}} dq_1 \dots dp_N}{\int \dots \int e^{-\frac{\alpha q_k^2}{kT}} e^{-\frac{f}{kT}} dq_1 \dots dp_N},$$

d'où, en simplifiant et en posant $x = \sqrt{\frac{\alpha}{kT}} q_k$,

$$\begin{aligned} \overline{\alpha q_k^2} &= \frac{\int_{-\infty}^{+\infty} \alpha q_k^2 e^{-\frac{\alpha q_k^2}{kT}} dq_k}{\int_{-\infty}^{+\infty} e^{-\frac{\alpha q_k^2}{kT}} dq_k} = kT \frac{\int_{-\infty}^{+\infty} x^2 e^{-x^2} dx}{\int_{-\infty}^{+\infty} e^{-x^2} dx} \\ &= \frac{kT}{2}. \end{aligned}$$

C. Q. F. D.

Naturellement, on démontrerait de même que la valeur moyenne de l'énergie pour un momentoïde du type p_k est également $\frac{1}{2} kT$.

7. Applications du théorème de l'équipartition de l'énergie. — Considérons d'abord une molécule ponctuelle dans un gaz. Son énergie est $\frac{1}{2m} (p_x^2 + p_y^2 + p_z^2)$ et, comme les $3 p_i$ sont des momentaïdes, l'énergie moyenne de la molécule en équilibre thermique avec le gaz à la température T est $\frac{3}{2} kT$, résultat déjà trouvé. Il en résulte pour l'énergie d'une molécule-gramme la valeur

$$\frac{3}{2} \mathcal{N} kT = \frac{3}{2} RT.$$

La chaleur spécifique moléculaire à volume constant d'un gaz monoatomique parfait est donc $C_M = \frac{3}{2} R$ et comme d'après un résultat bien connu de la Thermodynamique classique $C_M - c_M = R$, il vient

$$c_M = \frac{5}{2} R, \quad \gamma = \frac{C_M}{c_M} = \frac{5}{3} = 1,66,$$

résultat bien vérifié en général pour un gaz monoatomique.

Soit maintenant une molécule biatomique ayant la forme d'une haltère. La position de la molécule est déterminée par les coordonnées x, y, z de son centre de gravité et par les deux angles θ et φ qui fixent l'orientation de l'axe de l'haltère par rapport à un trièdre fixe. Si l'on suppose fixe le moment d'inertie de l'haltère par rapport à son axe, l'énergie de l'haltère est de la forme

$$\frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + ap_0^2 + bp_\varphi^2,$$

p_0 et p_φ étant les moments de Lagrange par rapport aux axes θ et φ . Nous avons donc ici cinq momentaïdes et l'énergie moyenne de la molécule à la température T est $\frac{5}{2} kT$. On a donc ici

$$c_M = \frac{5}{2} R, \quad C_M = \frac{7}{2} R, \quad \gamma = \frac{C_M}{c_M} = \frac{7}{5} = 1,4,$$

valeurs usuellement vérifiées pour les gaz biatomiques.

Pour des molécules plus compliquées, on aura en général n momentaïdes avec $n > 5$ et l'on prévoit les valeurs suivantes :

$$c_M = \frac{n}{2} R, \quad C_M = \frac{n+2}{2} R, \quad \gamma = \frac{n+2}{n} = 1 + \frac{2}{n}.$$

Pour des molécules de plus en plus compliquées, γ doit tendre vers un résultat qui se vérifie en gros.

Pour prendre un autre exemple, considérons un corps solide et admettons que dans ce corps solide les atomes puissent osciller autour de leur position d'équilibre et soient rappelés vers elle par une force proportionnelle à l'élongation. Comme nous l'avons déjà vu, il y a alors six momentaïdes par atome et l'énergie moyenne de l'atome à la température T doit être

$$6 \frac{1}{2} kT = 3 kT.$$

La chaleur spécifique atomique à volume constant est alors

$$C_M = 3R \text{ en unités mécaniques} = \frac{3R}{J} \text{ en unités calorifiques} \approx 6 \text{ cal.},$$

car

$$3 \frac{8,31 \cdot 10^7}{4,18 \cdot 10^7} = 5,97.$$

C'est la loi de Dulong et Petit d'après laquelle la chaleur spécifique atomique à volume constant d'un corps solide est voisine de 6 calories, loi très utile aux physiciens et aux chimistes et, en général, bien vérifiée dans la pratique.

Le théorème de l'équipartition de l'énergie nous conduit ainsi à beaucoup de prévisions exactes, mais il y a des ombres au tableau. D'abord les valeurs $\frac{3}{2}R$ et $\frac{5}{2}R$ pour les chaleurs spécifiques des gaz monoatomiques et biatomiques sont pour ainsi dire, trop bien vérifiées. En effet nous les avons prévues en assimilant la molécule monoatomique à un point matériel qui ne peut prendre aucune rotation et la molécule biatomique à une haltère qui ne peut tourner autour de son axe de figure; ces hypothèses sont physiquement peu vraisemblables et l'on a l'impression que certains degrés de liberté restent « ankylosés » et ne prennent pas la part d'énergie qui leur reviendrait d'après le théorème de l'équipartition. De plus, aux très basses températures, la chaleur spécifique des gaz biatomiques, approximativement égale à $\frac{5}{2}R$ aux températures usuelles, décroît et tend vers la valeur $\frac{3}{2}R$ des gaz monoatomiques. Ici encore, tout se passe comme si aux très basses températures les degrés de liberté de rotation des molécules biatomiques s'ankylosaient progressivement.

Des faits analogues se manifestent pour les corps solides et la loi de Dulong et Petit. Cette loi, bien vérifiée pour un grand nombre de corps solides aux températures usuelles, est complètement en défaut pour des

corps très durs comme le diamant dont la chaleur spécifique est très inférieure à la valeur prévue par la loi de Dulong et Petit. Il en est d'ailleurs de même pour tous les corps solides à température suffisamment basse.

Il y a là toute une série de faits qui semblent montrer l'insuffisance de la Mécanique statistique classique. Elle a été entièrement confirmée par l'échec total des méthodes de la Mécanique statistique dans la théorie du rayonnement noir. C'est pour interpréter cet échec que Max Planck a été amené à introduire en Physique une idée tout à fait inconnue dans les anciennes théories : celle de Quantum d'action. C'est ce que nous allons voir au prochain Chapitre.

BIBLIOGRAPHIE.

- [1] L. BOLTZMANN, *Leçons sur la théorie des gaz*, Gauthier-Villars, Paris, 1902.
- [2] W. GIBBS, *Principes élémentaires de Mécanique statistique*, Hermann, Paris, 1926.
- [3] H. A. LORENTZ, *Les théories statistiques en Thermodynamique*, Teubner, Leipzig, 1916.
- [4] J. H. JEANS, *The dynamical theory of gases*, Cambridge University press, 1916.
- [5] LÉON BRILLOUIN, *Les statistiques quantiques*, t. I, Presses universitaires Paris, 1930.

CHAPITRE V.

LA THÉORIE DU RAYONNEMENT NOIR.

1. **Généralités sur le problème du Rayonnement noir.** — La théorie du rayonnement d'équilibre thermique dit « Rayonnement du corps noir » ou, par abréviation, « Rayonnement noir », a joué un rôle considérable dans l'évolution de la Physique théorique, il y a une quarantaine d'années. C'est des difficultés que cette théorie a soulevées qu'est sortie la Théorie des Quanta dont le rôle dans le développement de la Physique atomique a été décisif. Bien que la répartition spectrale de l'énergie dans le Rayonnement noir ne paraisse plus aujourd'hui que comme un cas particulier de l'intervention des quanta dans les phénomènes physiques, l'importance que cette question a présentée dans l'histoire de la Science justifie une étude particulière approfondie.

Considérons une enceinte dont les parois sont maintenues à une température absolue constante T . Si les parois étaient parfaitement réfléchissantes et l'enceinte vide de matière, le rayonnement qui pourrait se trouver contenu dans l'enceinte garderait une composition spectrale constante, car le caractère linéaire des équations de l'électromagnétisme fait qu'il ne peut pas y avoir d'interactions entre les composantes spectrales; le rayonnement garderait donc sa distribution initiale quelconque et ne tendrait vers aucun état d'équilibre. Il y a là une différence importante entre le rayonnement et un gaz matériel, car un gaz enfermé dans une enceinte à température T tend de lui-même, par suite des chocs entre molécules, à prendre une répartition d'équilibre. Mais supposons, au contraire, qu'il y ait dans l'enceinte ou sur ses parois une certaine quantité de matière susceptible d'émettre et d'absorber des rayonnements de toutes les fréquences, alors la composition du rayonnement se modifiera et évoluera vers un état d'équilibre moyen : cet état d'équilibre sera atteint lorsque les processus d'émission et d'absorption pour chaque fréquence se compenseront en moyenne.

Nous admettrons donc que dans ces conditions le rayonnement atteint un état d'équilibre qui correspond à une parfaite isotropie.

Le rayonnement d'équilibre dans une enceinte à la température T est alors caractérisé par son isotropie et par la présence d'une certaine densité d'énergie par unité de volume, qui ne peut dépendre que de la température, soit $\rho(T)$. Si l'on analyse la répartition de l'énergie contenue dans l'unité de volume entre les fréquences, on est amené à écrire la décomposition spectrale

$$\rho(T) = \int_0^\infty \rho(\nu, T) d\nu.$$

Le problème essentiel de la théorie du Rayonnement noir est en somme la détermination de la fonction $\rho(\nu, T)$.

On peut introduire une autre grandeur directement reliée à $\rho(\nu, T)$.

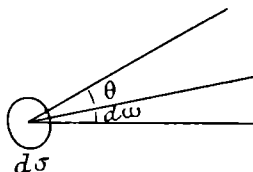


Fig. 15.

Soit une petite surface plane $d\sigma$ plongée dans le rayonnement noir et et considérons les rayonnements de fréquences comprises entre ν et $\nu + d\nu$ qui tombent en une unité de temps sur l'une des faces de $d\sigma$ suivant des directions contenues dans un cône infiniment délié d'ouverture $d\omega$ dont l'axe fait l'angle θ avec la normale à $d\sigma$.

Comme le rayonnement est isotrope et que, dans les parties vides de l'enceinte, il se propage avec la vitesse c , le flux des radiations considérées à travers $d\sigma$ pendant l'unité de temps sera

$$\rho(\nu, T) d\nu \frac{d\omega}{4\pi} c d\sigma \cos \theta.$$

Définissons alors une grandeur $I(\nu, T)$ telle que ce flux puisse aussi s'exprimer sous la forme

$$I(\nu, T) d\omega d\nu d\sigma \cos \theta.$$

En comparant cette expression à la précédente, on obtient

$$I(\nu, T) = \frac{c}{4\pi} \rho(\nu, T);$$

$I(\nu, T)$ est nommée *l'intensité spécifique de rayonnement d'équilibre*

pour la fréquence ν et cette grandeur peut servir aussi bien que $\rho(\nu, T)$ à caractériser le rayonnement.

Supposons maintenant que la petite surface $d\sigma$ soit formée par une lame d'une substance matérielle émettante et absorbante. Par seconde dans un angle $d\omega$, cette lame émet une quantité d'énergie que l'on peut écrire sous la forme

$$\varepsilon(\nu, T, \theta) d\nu d\omega d\sigma \cos\theta.$$

Elle reçoit une énergie radiante égale à $I(\nu, T) d\omega d\nu d\sigma \cos\theta$ et en absorbe une fraction $\alpha(\nu, T, \theta)$ qui, en général, dépend de l'orientation θ . L'équilibre thermique exige alors que l'on ait

$$\varepsilon(\nu, T, \theta) d\nu d\omega d\sigma \cos\theta = \alpha(\nu, T, \theta) I(\nu, T) d\nu d\omega d\sigma \cos\theta$$

ou, plus simplement,

$$I(\nu, T) = \frac{\varepsilon(\nu, T, \theta)}{\alpha(\nu, T, \theta)} = \frac{c}{4\pi} \rho(\nu, T);$$

ε est appelé « le pouvoir émissif » de la substance considérée et α son « pouvoir absorbant » pour les valeurs ν, T, θ . Leur quotient doit donc être indépendant de θ et entièrement déterminé par la répartition spectrale du rayonnement d'équilibre à la température T .

En particulier, appliquons ces résultats à un corps « noir », c'est-à-dire à un corps qui absorbe intégralement toutes les radiations qui tombent sur lui quelle que soit la fréquence. Certains corps réels comme le noir de fumée remplissent cette condition dans un domaine spectral étendu. On a alors $\alpha = 1$ et

$$\varepsilon(\nu, T) = I(\nu, T) = \frac{c}{4\pi} \rho(\nu, T);$$

ε est ici indépendant de θ et égal à l'intensité spécifique $I(\nu, T)$ pour chaque fréquence. Déterminer le rayonnement émis par un corps noir à la température T fait donc connaître la composition spectrale du rayonnement d'équilibre à cette température. D'où le nom, peu correct, de « Rayonnement noir à la température T » donné au rayonnement d'équilibre à cette température.

On peut étudier la répartition spectrale du rayonnement d'équilibre en mesurant l'émission d'un corps noir. Mais on peut l'étudier aussi en analysant la radiation qui s'échappe par une petite ouverture percée dans la paroi d'un vaste four maintenu à la température T , car il est évident que le flux d'énergie vers l'extérieur à travers cette ouverture ne trouble pas sensiblement l'équilibre du rayonnement à l'intérieur du four et

que, par suite, ce flux renseigne directement sur la densité spectrale du rayonnement d'équilibre thermique à la température du four.

La tâche essentielle de la théorie du Rayonnement noir étant d'arriver à préciser la forme de la fonction $\rho(\nu, T)$, on a commencé par appliquer à cette étude les méthodes générales et rigoureuses de la Thermodynamique. Comme nous allons le voir, on est arrivé ainsi à obtenir des propriétés très importantes de la fonction ρ qui sont bien vérifiées par l'expérience. Mais, par cette voie, on n'a pas pu déterminer entièrement la fonction $\rho(\nu, T)$ et, pour y parvenir, d'autres modes de raisonnement que nous étudierons ensuite ont été nécessaires.

2. Loi de Stefan-Boltzmann. — Un raisonnement purement thermodynamique permet d'abord d'obtenir un résultat important en ce qui concerne la quantité

$$\varphi(T) = \int_0^{\infty} \rho(\nu, T) d\nu.$$

Ce résultat est le suivant :

$$\varphi(T) = aT^4 \quad (a = \text{const.}).$$

C'est loi de Stefan-Boltzmann dont nous allons donner la démonstration.

Considérons une enceinte à parois parfaitement réfléchissantes maintenue à une température uniforme T ; dans cette enceinte se trouve un peu de matière dont la présence assure constamment la réalisation de l'équilibre thermodynamique. Le rayonnement est comparable à un gaz : il est homogène, il a une température T , il occupe un volume V et il exerce une pression uniforme p sur les parois de l'enceinte qu'il remplit. Cette pression, c'est la « pression de radiation ».

Calculons la pression exercée par le Rayonnement noir sur les parois réfléchissantes de l'enceinte. Nous savons ⁽¹⁾ qu'une onde électromagnétique plane de densité d'énergie w transporte dans le sens de sa propagation une quantité de mouvement égale à $\frac{w}{c}$ par unité de volume. Soit alors une onde plane tombant sous l'angle d'incidence θ sur un élément $d\sigma$ de surface parfaitement réfléchissante. La quantité de mouvement apportée par seconde par le rayonnement sur $d\sigma$ est $\frac{w}{c} c d\sigma \cos\theta$ et sa composante normale à $d\sigma$ est $w \cos^2\theta d\sigma$. D'après le théorème de l'impulsion, la force normale subie par $d\sigma$ doit être égale à la variation

(1) Voir chapitre I.

de cette composante normale de l'impulsion au moment de la réflexion, soit $2\omega \cos^2\theta d\sigma$: la pression qui en résulte sur $d\sigma$ est donc égale à $2\omega \cos^2\theta$. Maintenant, si $d\sigma$ est un élément de la paroi de l'enceinte contenant le Rayonnement noir, la densité d'énergie correspondant à des angles d'incidence compris entre θ et $\theta + d\theta$ est

$$\rho(\mathbf{T}) \frac{d\omega}{4\pi} = \rho(\mathbf{T}) \frac{1}{2} \sin\theta d\theta,$$

Pour avoir la pression totale, il faut intégrer sur θ de 0 à $\frac{\pi}{2}$ l'expression

$$2\rho(\mathbf{T}) \frac{1}{2} \sin\theta d\theta \cos^2\theta,$$

ce qui donne

$$p = \rho(\mathbf{T}) \left| -\frac{1}{3} \cos^3\theta \right|_0^{\frac{\pi}{2}} = \frac{1}{3} \rho(\mathbf{T}).$$

La pression de radiation est donc égale au tiers de la densité du rayonnement noir. L'énergie totale du rayonnement est évidemment $\rho(\mathbf{T})\mathbf{V}$ et l'on a pour la différentielle de son entropie

$$\begin{aligned} dS &= \frac{dE + p dV}{\mathbf{T}} = \frac{\mathbf{V}}{\mathbf{T}} \frac{d\rho}{d\mathbf{T}} d\mathbf{T} + \left(\rho + \frac{1}{3} \rho \right) dV \\ &= \frac{\mathbf{V}}{\mathbf{T}} \frac{d\rho}{d\mathbf{T}} d\mathbf{T} + \frac{4}{3} \rho dV. \end{aligned}$$

Pour que ceci soit une différentielle exacte, il faut que l'on ait

$$\frac{\partial}{\partial V} \left(\frac{\mathbf{V}}{\mathbf{T}} \frac{d\rho}{d\mathbf{T}} \right) = \frac{\partial}{\partial \mathbf{T}} \left(\frac{4}{3} \frac{\rho}{\mathbf{T}} \right) \quad \text{ou} \quad \frac{d\rho}{d\mathbf{T}} = 4 \frac{\rho}{\mathbf{T}}.$$

On trouve, par intégration, en supposant $\rho(0) = 0$

$$\rho(\mathbf{T}) = a\mathbf{T}^4,$$

ce qui est la loi de Stefan.

L'étude du Rayonnement du corps noir confirme l'exactitude de la loi et montre que l'on a numériquement en unités C. G. S. avec l'échelle absolue des températures

$$a = 7,64 \cdot 10^{-15}.$$

3. Loi du déplacement de Wien. — La loi de Stefan ne nous donne qu'un renseignement global sur l'intégrale de $\rho(\nu, \mathbf{T})$. La Thermodynamique permet cependant, par un raisonnement plus compliqué, d'aller

plus loin et d'obtenir un résultat précis sur la forme de $\rho(\nu, T)$. Il est exprimé par la loi du déplacement de Wien

$$\varepsilon(\nu, T) = \nu^3 F\left(\frac{\nu}{T}\right),$$

où F est une fonction que le raisonnement thermodynamique ne permet pas de déterminer.

Le raisonnement thermodynamique qui conduit à la loi de Wien est assez long et nous ne le reproduirons pas en détail. On le trouvera très minutieusement exposé dans le livre *Wärmestrahlung*, de Planck [1]. Donnons-en seulement le principe. Considérons une enceinte à parois parfaitement réfléchissantes dans laquelle se trouve le rayonnement noir de la température T , mais aucune trace de matière pondérable. Supposons qu'on fasse mouvoir très lentement une partie de la paroi de l'enceinte formant piston. Par suite de la pression de radiation, le rayonnement accomplit un travail (positif ou négatif) sur cette paroi mobile, mais il n'y a aucun échange de chaleur avec l'extérieur : nous avons donc affaire à une transformation adiabatique réversible à entropie constante. La température du rayonnement varie pendant ce processus, mais on peut démontrer qu'à chaque instant le rayonnement a la composition qui correspond à l'équilibre pour la température instantanée. Ce fait n'est pas évident *a priori*, car en l'absence de toute matière émettante et absorbante nous ne pouvons pas affirmer que l'équilibre se rétablisse de lui-même, mais c'est là une propriété de la transformation réversible isentropique considérée car, s'il n'en était pas ainsi, on pourrait constituer un cycle d'états successifs ramenant exactement le système à son état primitif au cours duquel l'entropie aurait augmenté, résultat inadmissible. Mais alors, puisqu'il n'y a pas dans l'enceinte de matière susceptible de provoquer par absorption et émission une redistribution de l'énergie entre les diverses fréquences, comment peut se réaliser cette redistribution qui est évidemment nécessaire au maintien de la répartition d'équilibre au cours du processus adiabatique ? Elle s'opère grâce à l'effet Doppler qui se produit lors de la réflexion de la radiation sur la partie mobile de la paroi. Ce processus d'effet Doppler sur un miroir mobile est facile à analyser et l'on déduit de cette étude que la seule loi de répartition spectrale $\rho(\nu, T)$ qui ne soit pas modifiée par le processus adiabatique est précisément la loi de Wien. Mais le résultat obtenu est incomplet puisque la fonction F reste inconnue.

On peut d'ailleurs retrouver la loi de Wien par un raisonnement

très élégant développé par Léon Brillouin en partant de la théorie générale de l'invariance adiabatique [4].

Les expérimentateurs préfèrent énoncer le résultat de leurs recherches en fonction de la longueur d'onde λ plutôt que de la fréquence ν , les deux grandeurs étant reliées dans le cas du rayonnement par la relation $\lambda\nu = c$. Comme on a évidemment

$$\varphi(\nu, T) d\nu = \varphi(\lambda, T) d\lambda,$$

avec $|d\lambda| = \frac{c}{\nu^2} |d\nu|$, on a aussi

$$\varphi(\lambda, T) = \varphi(\nu, T) \frac{c}{\lambda^2}$$

et la loi de Wien exprimée dans l'échelle des longueurs d'onde prendra la forme

$$\varphi(\lambda, T) = \frac{1}{\lambda^5} \mathcal{F}(\lambda T),$$

\mathcal{F} étant une fonction inconnue du produit λT . L'étude expérimentale du Rayonnement noir permet de déterminer $\varphi(\lambda, T)$ pour une température T donnée et montre que cette fonction de λ présente un maximum pour une valeur $\lambda = \lambda_m(T)$ variable avec T . En écrivant que la dérivée de $\varphi(\lambda, T)$ par rapport à λ est nulle pour $\lambda = \lambda_m$, on obtient

$$-\frac{5}{\lambda_m^6} \mathcal{F}(\lambda_m T) + \frac{T}{\lambda_m^5} \mathcal{F}'(\lambda_m T) = 0,$$

où \mathcal{F}' est la dérivée de \mathcal{F} par rapport à la variable λT . Le produit $\lambda_m T$ est donc donné par la relation

$$\lambda_m T \mathcal{F}'(\lambda_m T) = 5 \mathcal{F}(\lambda_m T).$$

C'est donc un nombre indépendant de T et l'on obtient la loi du déplacement du maximum

$$\lambda_m T = b \quad (b = \text{const.}).$$

L'expérience vérifie bien cette loi et donne en unités C. G. S., avec l'échelle des températures absolues,

$$b = 0,294.$$

Avec les lois de Stefan et de Wien, on a épuisé tout ce que la Thermodynamique classique peut nous apprendre sur la fonction $\varphi(\nu, T)$. Pour achever de préciser la forme de cette fonction, on a tenté de faire appel à la Mécanique statistique jointe à la théorie électromagnétique

de l'émission et de l'absorption du rayonnement. Ces tentatives ont abouti à un échec complet et c'est cet échec qui a été le point de départ des théories quantiques.

4. **La loi de distribution spectrale de Rayleigh-Jeans.** — En admettant les conceptions de la Mécanique statistique et de l'électromagnétisme classique, on a pu aisément trouver la forme que devrait avoir la fonction $\rho(\nu, T)$. Cette forme est la suivante :

$$\rho(\nu, T) = \frac{8\pi\nu^2}{c^3} kT.$$

C'est la loi de Rayleigh-Jeans que l'on peut obtenir de plusieurs façons.

Un premier mode de raisonnement, qui fut notamment employé par Max Planck au début de ses recherches sur ce sujet, part de la remarque suivante : l'équilibre thermodynamique doit être indépendant des mécanismes qui, dans la matière, assurent l'émission et l'absorption du rayonnement à la seule condition que ces mécanismes soient conformes aux théories admises pour ces phénomènes, et il est dès lors suffisant d'imaginer les mécanismes les plus simples. Nous pouvons donc nous contenter, dans le cadre des idées classiques de Lorentz, d'imaginer que l'émission et l'absorption des rayonnements par la matière ait lieu par l'intermédiaire d'oscillateurs électroniques linéaires, c'est-à-dire d'électrons présents dans la matière et susceptibles d'osciller le long d'une droite autour d'une position d'équilibre sous l'action d'une force proportionnelle à l'élongation. Pour un tel oscillateur, la coordonnée x le long de la droite et le moment p_x conjugué sont des momentaïdes : d'après le théorème de l'équipartition de l'énergie, dans l'état d'équilibre thermique à la température T , l'énergie moyenne de ces oscillateurs doit être égale à kT . Or, la théorie électromagnétique montre qu'un oscillateur électronique linéaire dont la fréquence est ν , émet par seconde une quantité d'énergie $\frac{8\pi^2\nu^2}{3c^3} \frac{e^2}{m} \overline{W}$ ergs si \overline{W} est son énergie d'oscillation. Un oscillateur linéaire en équilibre à la température T émet donc en moyenne par seconde $\frac{8\pi^2\nu^2}{3c^3} \frac{e^2}{m} kT$ ergs, puisque $\overline{W} = kT$. Or, l'oscillateur étant plongé dans le rayonnement de la température T , on peut calculer l'énergie qu'il absorbe en moyenne par seconde à ce rayonnement et l'on trouve $\frac{\pi e^2}{3m} \rho(\nu, T)$. Pour exprimer l'équilibre thermodynamique, on doit égaler les énergies

émises et absorbées en moyenne par seconde et l'on obtient ainsi aussitôt la loi de Rayleigh.

Un autre raisonnement dû à Rayleigh et à Jeans conduit au même résultat. Considérons une enceinte vide à parois parfaitement réfléchissantes ayant la forme d'un parallélépipède rectangle. Dans cette enceinte peuvent s'établir un certain nombre d'ondes électromagnétiques stationnaires, déterminées par la condition aux limites que sur les parois réfléchissantes le champ électrique doit être normal à la paroi. Si les dimensions sont grandes par rapport aux longueurs d'onde qui interviennent, on démontre (voir [5], I, p. 32) que le nombre des ondes électromagnétiques stationnaires caractérisées par une fréquence, une direction de propagation et un état de polarisation rectiligne (on sait qu'il y a deux tels états pour chaque fréquence et chaque direction de propagation) est pour l'intervalle de fréquence $(\nu, \nu + d\nu)$ égal à

$$dn_\nu = \frac{8\pi\nu^2}{c^3} V d\nu \quad (\text{formule de Jeans}).$$

V étant le volume de l'enceinte. Ce résultat démontré pour un parallélépipède rectangle est valable pour des formes beaucoup plus générales de l'enceinte (Weyl). Maintenant, Lorentz a démontré que chacune de ces ondes stationnaires se comporte au point de vue énergétique comme un oscillateur linéaire et, en particulier, doit posséder dans l'équilibre thermique à température T une énergie moyenne égale à kT . Cela donne encore

$$\rho(\nu, T) d\nu = \frac{dn_\nu kT}{V} \quad \text{ou} \quad \rho(\nu, T) = \frac{8\pi\nu^2}{c^3} kT.$$

Cette nouvelle démonstration ne fait pas intervenir explicitement le mécanisme de l'émission du rayonnement par la matière.

5. Échec de la loi de Rayleigh. — On peut d'abord chercher si la loi de Rayleigh est en accord avec les deux lois thermodynamiques de Stefan et de Wien. Pour la loi de Wien, la vérification est immédiate, car il suffit de poser

$$F\left(\frac{\nu}{T}\right) = \frac{8\pi^2}{c^3} \frac{kT}{\nu}.$$

Mais pour la loi de Stefan, on trouve une grave difficulté provenant du fait inacceptable qu'avec la formule de Rayleigh, l'intégrale $\int_0^\infty \rho(\nu, T) d\nu$ est infinie.

D'autres difficultés viennent de la comparaison avec les résultats expérimentaux.

Dans la période 1895-1900 où l'étude du rayonnement noir était à l'ordre du jour, un grand nombre de recherches expérimentales furent effectuées sur la forme de la fonction $\rho(\nu, T)$. On a tracé les courbes « isothermes » donnant $\rho(\nu, T)$ en fonction de ν pour T constant, et aussi les courbes « isochromatiques » donnant $\rho(\nu, T)$ en fonction de T pour ν constant, ce qui permet de construire la surface représentant $\rho(\nu, T)$. En comparant les résultats avec la loi de Rayleigh, on a reconnu que cette loi se vérifie bien pour les petites valeurs du quotient $\frac{\nu}{T}$. Ainsi, pour une valeur donnée de T , la valeur exacte de $\rho(\nu, T)$ est bien donnée par la loi de Rayleigh pour les petites valeurs de ν ; elle

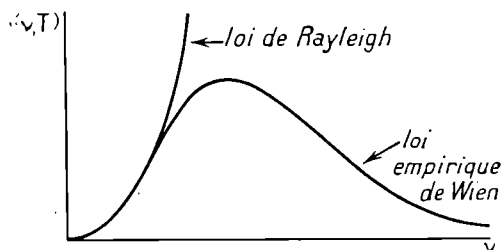


Fig. 16.

croît alors comme ν^2 , mais pour ν croissant suffisamment elle commence par croître moins vite que ν^2 , puis passe par un maximum dont nous avons déjà parlé et enfin décroît. La courbe isotherme a donc une forme « en cloche », ce qui permet à l'intégrale $\int_0^\infty \rho(\nu, T) d\nu$ d'être finie contrairement à ce qui se passe si l'on admet la loi de Rayleigh.

Pour les grandes valeurs du quotient $\frac{\nu}{T}$, la loi de Rayleigh devient donc tout à fait inexacte et Wien a trouvé, pour représenter cette partie de la courbe, la loi empirique

$$\rho(\nu, T) = A \nu^3 e^{-B \frac{\nu}{T}},$$

analogue par sa forme à la loi de Maxwell. Il ne faut pas confondre cette loi empirique de répartition de Wien avec sa loi du déplacement [ici $F\left(\frac{\nu}{T}\right) = A e^{-B \frac{\nu}{T}}$].

Ainsi, les lois classiques de la Mécanique statistique et de l'Électromagnétisme, si souvent vérifiées dans d'autres domaines, se sont trouvées

tout à fait en échec dans la théorie du rayonnement noir. Henri Poincaré a montré que ces lois conduisaient nécessairement à la formule inexacte de Rayleigh. Pour trouver la forme exacte de $\rho(\nu, T)$, il fallait introduire un élément tout à fait étranger aux théories classiques. C'est ce qu'a fait Planck en 1900.

6. La loi de répartition spectrale de Planck. — Max Planck a aperçu la raison profonde des difficultés rencontrées dans la théorie du rayonnement noir : elles proviennent de ce que le théorème classique de l'équipartition de l'énergie avantage d'une façon excessive les grandes fréquences. En effet, le nombre des ondes stationnaires dans une enceinte croissant comme ν^2 , si toutes ces ondes stationnaires prennent en moyenne la même énergie kT , les ondes de haute fréquence prendront une énergie infinie et la matière devra céder toute son énergie au rayonnement. C'est cette conséquence absurde qui vicie la théorie classique du Rayonnement noir.

Pour éviter cette conséquence, il faut imaginer une nouvelle loi de la nature qui défavorise les hautes fréquences. Planck l'a fait très heureusement en introduisant l'hypothèse des Quanta. Selon cette hypothèse, un oscillateur de fréquence ν ne peut émettre ou absorber l'énergie radiante que par quantité finie égale à $h\nu$, h étant une nouvelle constante universelle introduite par l'hypothèse elle-même. Alors, plus un oscillateur a une fréquence élevée, plus il exige qu'on lui offre une grosse quantité d'énergie pour pouvoir l'accepter, ce qui, dans la répartition d'équilibre de l'énergie, défavorise évidemment les très hautes fréquences. Si l'énergie moyenne kT à la température considérée est très grande par rapport au quantum $h\nu$ d'une certaine catégorie d'oscillateurs, tout se passera comme si le quantum était infiniment petit et l'on doit retrouver les lois classiques : c'est pourquoi la loi de Rayleigh se vérifie pour les petites valeurs du quotient $\frac{\nu}{T}$. Mais, quand le quantum $h\nu$ est très grand par rapport à l'énergie moyenne kT , les oscillateurs ne trouvent que rarement l'occasion d'absorber un quantum, leur énergie moyenne reste inférieure à celle que prévoit le théorème classique de l'équipartition et l'on doit s'écarter de la théorie classique : c'est pourquoi la loi de Rayleigh ne se vérifie plus pour les grandes valeurs du quotient $\frac{\nu}{T}$.

L'hypothèse des Quanta explique également les échecs de la Mécanique statistique classique dans le domaine des chaleurs spécifiques : si à une

température T certains degrés de liberté correspondent à un quantum très grand par rapport à l'énergie moyenne kT , ces degrés de liberté n'entreront pratiquement pas en jeu dans l'équipartition et ils sembleront être « ankylosés » : les chaleurs spécifiques prévues par le théorème classique de l'équipartition seront alors plus grandes que les chaleurs spécifiques réelles. Nous avons vu qu'il en est bien ainsi pour les chaleurs spécifiques des gaz à basse température et pour celles des corps solides très durs (diamant), même à la température ordinaire.

L'hypothèse des Quanta présente donc *a priori* l'avantage d'expliquer qualitativement plusieurs faits impossibles à interpréter avec les idées classiques. Nous allons montrer avec Planck qu'elle fournit aussi une forme satisfaisante de $\rho(\nu, T)$: la formule de Planck. C'est de plusieurs façons différentes que Planck est parvenu à démontrer sa formule (voir [1] et [2]). Nous ne donnerons que l'une de ces démonstrations, celle qui correspond à la seconde démonstration de la loi de Rayleigh.

Nous avons dit que les ondes électromagnétiques stationnaires existant dans une enceinte sont assimilables pour le calcul des répartitions de l'énergie à des oscillateurs linéaires. Donc, d'après l'hypothèse des quanta, une telle onde stationnaire de fréquence ν peut avoir une énergie égale à $nh\nu$, n étant un nombre entier, résultat que l'hypothèse des photons rendra plus intuitif. La loi de distribution canonique de Gibbs-Boltzmann reste ici applicable, car on peut la démontrer sans supposer que les variations d'énergie soient continues. La probabilité pour que, dans l'état d'équilibre thermique à température T , une onde stationnaire de fréquence ν , assimilable à un oscillateur linéaire de même fréquence, ait une énergie égale à $nh\nu$ est donc proportionnelle à $e^{-\frac{nh\nu}{kT}}$ et l'énergie moyenne d'une telle onde stationnaire sera

$$\begin{aligned} \bar{E} &= \frac{\sum_0^{\infty} n h \nu e^{-\frac{nh\nu}{kT}}}{\sum_0^{\infty} e^{-\frac{nh\nu}{kT}}} = - \frac{\partial}{\partial \left(\frac{1}{kT} \right)} \log \sum_0^{\infty} e^{-\frac{nh\nu}{kT}} \\ &= \frac{\partial}{\partial \left(\frac{1}{kT} \right)} \log \left(1 - e^{-\frac{h\nu}{kT}} \right) = \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1}, \end{aligned}$$

expression qui se réduit bien à kT pour les petites valeurs de $\frac{\nu}{T}$. En multipliant par le nombre dn_ν des ondes stationnaires de l'inter-

valle (ν , $\nu + d\nu$) donné par la formule de Jeans, on obtient aisément la formule de Planck

$$\rho(\nu, T) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{kT}} - 1}.$$

C'est la célèbre *loi de répartition spectrale de Planck* qui est bien vérifiée par l'expérience.

Il est évident qu'elle est en accord avec la loi de répartition de Wien. Elle satisfait aussi à la loi de Stefan, car on a

$$\rho(T) = \int_0^\infty \rho(\nu, T) d\nu = \frac{8\pi h}{c^3} \int_0^\infty \frac{\nu^3 d\nu}{e^{\frac{h\nu}{kT}} - 1} = \frac{8\pi k^4}{c^3 h^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} T^4 = aT^4,$$

avec

$$a = \frac{8\pi k^4}{c^3 h^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{8\pi^5 k^4}{15 c^3 h^3}.$$

Lorsque que $h\nu$ est petit, la loi de Planck se confond avec la loi de Rayleigh parce qu'alors $e^{\frac{h\nu}{kT}} - 1 \simeq \frac{h\nu}{kT}$. Au contraire, pour les grandes valeurs de $\frac{h\nu}{kT}$, on peut remplacer $\frac{1}{e^{\frac{h\nu}{kT}} - 1}$ par $e^{-\frac{h\nu}{kT}}$ et l'on a la loi empirique de Wien

$$\rho(\nu, T) = \frac{8\pi h\nu^3}{c^3} e^{-\frac{h\nu}{kT}}$$

Écrite en longueurs d'onde, la loi de Planck prend la forme

$$\rho(\lambda, T) = \frac{c}{\lambda^3} \rho(\nu, T) = \frac{8\pi hc}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1}.$$

Pour déterminer la longueur λ_m qui correspond au maximum, il faut, nous l'avons vu, résoudre l'équation

$$\lambda_m T \mathcal{F}'(\lambda_m T) = 5 \mathcal{F}(\lambda_m T)$$

et l'on a ici $\mathcal{F}(\lambda T) = \frac{\text{const.}}{e^{\frac{hc}{\lambda kT}} - 1}$. En posant $x = \frac{hc}{\lambda k_m T}$, on trouve l'équation transcendante

$$e^{-x} + \frac{x}{5} - 1 = 0$$

dont la racine est $x = 4,965$. On a donc

$$\lambda_m T = b = \frac{hc}{4,965k}.$$

Des deux équations obtenues qui donnent a et b , nous pouvons, en introduisant les valeurs expérimentales $a = 7,64 \cdot 10^{-15}$ et $b = 0,294$, tirer les valeurs des constantes k et h en unités C. G. S. avec l'échelle des températures absolues. On trouve ainsi

$$k = 1,37 \cdot 10^{-16}, \quad h = 6,525 \cdot 10^{-27} \text{ erg. s.}$$

Pour la constante de Boltzmann k , on retrouve la valeur déjà connue par ailleurs (*voir* chap. précédent), mais on obtient du même coup la valeur de la nouvelle constante h , la fameuse « constante de Planck ». Il est remarquable que cette valeur de h , obtenue par Planck dès ses premières études de 1900, ait été sensiblement confirmée par toutes les méthodes très nombreuses et très diverses qui ont depuis lors permis de mesurer cette constante fondamentale de la Nature.

L'hypothèse des quanta de Planck a été au début accueillie avec un certain scepticisme. Elle entraînait, en effet, des conséquences très graves. Par exemple, si l'émission de la lumière par quanta pouvait à la rigueur se concilier avec la structure continue qu'on attribuait à la lumière depuis Fresnel, l'absorption par quanta semblait nécessiter un retour vers une conception corpusculaire de la lumière analogue à celle de Newton, ce qui soulevait d'énormes difficultés pour l'interprétation des interférences. Effrayé par les conséquences de ses propres idées, Planck a essayé de conserver l'émission par quanta sans postuler l'absorption par quanta. Sa théorie (connue naguère sous le nom de deuxième version de la théorie des Quanta) était ingénieuse et a eu une certaine vogue vers 1910. Mais elle créait une dissymétrie inacceptable entre l'émission et l'absorption des radiations et elle a été condamnée par le développement ultérieur des théories. Il a fallu en revenir, notamment avec la théorie de l'atome de Bohr, à la première version de la théorie des Quanta, version plus radicale qui admet à la fois l'émission et l'absorption par quanta et qui nous a obligé à réviser les idées classiques sur la structure continue du rayonnement.

7. La théorie quantique des chaleurs spécifiques (Einstein, Debye). — Confirmée par la découverte de la loi exacte du rayonnement noir, l'hypothèse des Quanta a assez rapidement conduit à améliorer la théorie des chaleurs spécifiques. Nous avons déjà dit pourquoi il devait en

être ainsi. Nous allons exposer rapidement la théorie (un peu simpliste) des chaleurs spécifiques des solides développée par Einstein vers 1910. théorie qui montre bien l'origine quantique des écarts par rapport à la loi de Dulong et Petit.

D'après la loi de Dulong et Petit justifiée par la Mécanique statistique classique, la chaleur spécifique atomique d'un corps solide homogène doit être voisine de 6 calories. Or, tous les corps à température très basse et certains corps solides (notamment des corps très durs comme le diamant), dès la température ordinaire se refusent à suivre cette loi et ont une chaleur spécifique atomique très inférieure à la valeur prévue. Nous avons expliqué pourquoi l'hypothèse des Quanta devait pouvoir interpréter cet échec de la Mécanique statistique classique. Faisons avec Einstein l'hypothèse un peu schématique que les atomes d'un corps solide homogène soient liés à leur position d'équilibre par des forces proportionnelles aux elongations, de sorte qu'ils soient tous susceptibles de vibrer autour de cette position d'équilibre avec une même fréquence ν . Cette hypothèse, assurément insuffisante, nous montrera cependant le sens des phénomènes. Chaque atome a dans l'état d'équilibre thermique à la température T une énergie moyenne

$$\bar{E} = \frac{3h\nu}{e^{\frac{h\nu}{kT}} - 1}$$

parce que l'atome a trois degrés de liberté. L'énergie d'un atome gramme du corps solide est

$$E_A = \frac{3\mathcal{N}h\nu}{e^{\frac{h\nu}{kT}} - 1} \quad (\mathcal{N} = \text{nombre d'Avogadro}).$$

La chaleur spécifique atomique à volume constant se calcule aussitôt

$$c_A = \frac{dE_A}{dT} = 3R \left(\frac{h\nu}{kT} \right)^2 \frac{e^{\frac{h\nu}{kT}}}{\left(e^{\frac{h\nu}{kT}} - 1 \right)^2}.$$

Si le quantum $h\nu$ est très inférieur à kT , on retrouve $c_A = 3R$, c'est-à-dire la loi de Dulong et Petit, comme on devait s'y attendre.

Si le quantum $h\nu$ est grand devant kT , on trouve

$$c_A = 3R \left(\frac{h\nu}{kT} \right)^2 e^{-\frac{h\nu}{kT}};$$

c_A est donc alors très inférieur à $3R$ et tend vers zéro quand T tend vers zéro. On s'explique ainsi qu'à très basse température les corps solides aient une chaleur spécifique plus faible que ne le voudrait la loi du Dulong et Petit. De plus, pour les corps solides très durs, les atomes doivent être liés à leur position d'équilibre par des forces très énergiques et, par suite, ν doit être élevé. On conçoit alors que, pour ces corps, le quotient $\frac{h\nu}{kT}$ puisse être grand, même aux températures usuelles et que la chaleur spécifique à ces températures soit déjà inférieure à $3R$.

La théorie précédente est très schématique. Un peu améliorée par F. A. Lindemann, elle a surtout été développée par les travaux de Debye et ceux de Born et Karman qui ont considéré, non plus une seule fréquence ν , mais l'ensemble des fréquences des ondes élastiques susceptibles de se propager dans un corps solide cristallisé. On obtient ainsi une théorie plus satisfaisante et plus susceptible d'être comparée à la réalité. Nous ne l'étudierons pas ici, la théorie primitive d'Einstein suffisant pour donner une première idée des applications de la théorie des Quanta à l'interprétation des chaleurs spécifiques.

BIBLIOGRAPHIE.

- [1] MAX PLANCK, *Wärmestrahlung*, J. A. Barth, Leipzig, 1921.
- [2] *Comptes rendus du premier Conseil de Physique Solvay*, tenu à Bruxelles en octobre 1911, Gauthier-Villars, Paris, 1912.
- [3] H. A. LORENTZ, *Les théories statistiques en Thermodynamique*, Teubner, Leipzig, 1916.
- [4] LÉON BRILLOUIN, *La théorie des Quanta et l'Atome de Bohr*, Presses Universitaires, Conférences-Rapports, Paris, 1922.
- [5] LÉON BRILLOUIN, *Les statistiques quantiques* (2 vol.), Presses Universitaires, Conférences-Rapports, Paris, 1930.

CHAPITRE VI.

LA STRUCTURE CORPUSCULAIRE DE LA LUMIÈRE. LES PHOTONS.

1. **Aperçu historique sur les théories de la Lumière.** — L'histoire des théories de la Lumière est l'une des branches les plus curieuses de l'histoire des Sciences. Deux grandes conceptions de la Lumière, la conception corpusculaire et la conception ondulatoire, se sont disputées la préférence des savants, et cette rivalité a abouti à l'heure actuelle à une synthèse d'un type très nouveau, faisant appel à des conceptions très originales dont il est essentiel de bien comprendre la nature si l'on veut s'initier aux théories quantiques contemporaines.

Les philosophes de l'Antiquité semblent avoir penché vers une conception corpusculaire de la lumière : Lucrèce, dans son poème *De Natura rerum*, imaginait que les corps projettent dans l'espace à des vitesses « inimaginables » des sortes de pellicules ultralégères ou « simulacres » qui se détacheraient de leur surface et reproduiraient les détails de celle-ci, de sorte que ces pellicules recueillies par notre œil pourraient y reconstituer l'aspect des corps qui les a émises. C'était là une théorie ingénieuse, mais sans base physique réelle. Il en fut de même pour toutes les conceptions de la lumière développées jusqu'au xvii^e siècle, époque où commence, avec les recherches de Descartes, de Snell et de Fermat, l'étude scientifique de la lumière. Descartes paraît avoir été plutôt favorable à l'hypothèse corpusculaire, mais les textes qu'il a laissés sur ce sujet sont obscurs et ont été diversement interprétés. Puis, tandis que le Danois Rømer mesure pour la première fois en 1676 la vitesse finie de la lumière, un esprit très vigoureux, le Hollandais Christian Huyghens, entreprend des travaux théoriques sur la lumière qui sont restés justement célèbres et développe des méthodes qui subsistent encore dans notre enseignement actuel. Huyghens émet pour la première fois, sous une forme scientifique, la conception

ondulatoire de la lumière. Il montre qu'avec elle, on peut interpréter les lois de la réflexion et de la réfraction connues depuis les travaux de Snell et de Descartes, ainsi que le phénomène, alors récemment découvert par Bartholin, de la double réfraction du spath d'Islande, mais il ne peut donner aucune explication précise du phénomène fondamental de la propagation rectiligne de la lumière si facile à comprendre avec l'hypothèse corpusculaire. L'œuvre de Huyghens était très en avance sur son temps et fut vite oubliée grâce surtout à l'influence de son contemporain Newton qui jouissait dans le monde savant d'une immense réputation, par suite de ses découvertes en Mathématiques, en Mécanique et en Astronomie. Newton a fait de belles recherches expérimentales en Optique : tout en se défendant de faire des hypothèses, il a toujours plutôt soutenu l'hypothèse corpusculaire. Néanmoins, comme il avait découvert le phénomène d'interférences qui porte son nom (anneaux de Newton), il avait senti la nécessité de compléter son point de vue corpusculaire en faisant intervenir des éléments périodiques : selon lui, les corpuscules de lumière passeraient alternativement par des accès de facile réflexion et des accès de facile transmission, ces accès étant sans doute dus à l'influence de perturbations périodiques qui accompagneraient le corpuscule lors de son passage à travers les milieux matériels. Newton obtenait ainsi une interprétation des anneaux qu'il avait découverts et il définissait même une « longueur d'accès » qui est étroitement apparentée à la longueur d'onde des théories ondulatoires. Cette tentative un peu bâtarde n'a pas été très fructueuse : elle est cependant très curieuse au point de vue de l'histoire des Sciences, car elle est une sorte de pressentiment de la Mécanique ondulatoire.

Après Newton, pendant le XVIII^e siècle, on adopta généralement l'hypothèse corpusculaire : quelques savants, tels que Euler, se prononcèrent cependant d'une façon un peu platonique en faveur de l'hypothèse inverse. Au début du siècle dernier, il y eut un changement complet. Le savant anglais Thomas Young découvre les phénomènes d'interférences. Augustin Fresnel les étudie ensuite ainsi que les phénomènes de diffraction et montre que seule la conception ondulatoire les interprète réellement : il lève la grosse objection qu'on pouvait faire à cette hypothèse en montrant qu'elle peut aussi expliquer la propagation rectiligne de la lumière. D'abord combattu par de grandes autorités scientifiques de son temps, Fresnel finit par imposer ses vues (1816-1820). Peu après, il interprète les phénomènes de double réfraction et ceux de polarisation (étudiés à cette époque par

Malus, Arago et Fresnel lui-même), en introduisant l'hypothèse de la transversalité des vibrations lumineuses et complète ainsi son œuvre admirable. Après sa mort (1827), l'hypothèse ondulatoire triomphe de tous ses antagonistes. Vers 1850, Fizeau et Foucault mesurent avec précision la vitesse de propagation de la lumière dans le vide et dans divers milieux matériels : ils montrent notamment que la vitesse de la lumière dans l'eau, où l'indice de réfraction n est supérieur à 1, est plus petite que dans le vide conformément à la formule $V = \frac{c}{n}$ de la théorie ondulatoire et en contradiction avec la formule $V = nc$ fournie par la théorie corpusculaire de Newton.

Ensuite vint Maxwell qui a donné de la théorie ondulatoire de Fresnel l'interprétation électromagnétique que nous avons étudiée au chapitre I.

Arrêtons ici cet aperçu historique. Pendant toute la fin du siècle dernier, les physiciens ont accumulé un nombre énorme de vérifications extrêmement précises des théories de l'Optique ondulatoire. Après cette longue et minutieuse mise à l'épreuve de l'Optique ondulatoire, il paraissait inconcevable qu'on puisse être amené à remettre en question sa validité. C'est cependant l'inconcevable qui s'est réalisé après la découverte de l'effet photoélectrique.

2. L'effet photoélectrique. — Une plaque de métal irradiée par de la lumière de courte longueur d'onde émet de l'électricité négative : tel est l'effet photoélectrique observé par Hertz en 1887. L'étude de ce phénomène a d'abord montré que cette émission d'électricité négative avait lieu sous la forme de projections d'électrons pour lesquels on a pu retrouver la valeur du rapport $\frac{e}{m}$ de la charge à la masse connue par ailleurs.

Or, le phénomène photoélectrique obéit à des lois qui paraissent déconcertantes parce que rien ne pouvait les faire prévoir. La fréquence de la radiation incidente y joue un rôle essentiel et tout à fait inattendu. Pour que le phénomène apparaisse, il faut d'abord que cette fréquence dépasse une certaine valeur dépendant du corps irradié et dite « fréquence seuil ». Lorsque la fréquence est supérieure à la fréquence seuil, les électrons expulsés par l'action de la lumière (photoélectrons) sortent de la matière avec une énergie cinétique d'autant plus grande que la fréquence dépasse davantage la fréquence seuil. La loi est

linéaire, de sorte que si ν désigne la fréquence, ν_0 la fréquence seuil et T l'énergie cinétique des photoélectrons, on a

$$T = k(\nu - \nu_0).$$

Ce qui est extraordinaire dans cette loi si simple, c'est que la fréquence, et la fréquence seule, ν intervienne, l'intensité n'y jouant aucun rôle : toutes les conceptions classiques auraient fait prévoir le contraire. C'est seulement le nombre des photoélectrons éjectés par seconde qui croît proportionnellement à l'intensité : tout se passe, remarque fort importante, comme si les processus photoélectriques élémentaires dépendaient uniquement de la fréquence et aucunement de l'intensité, celle-ci n'intervenant que dans la détermination de la probabilité des processus élémentaires. Il est impossible d'imaginer un processus conforme aux idées classiques sur la constitution de la lumière qui conduise à un résultat semblable, d'autant plus que l'émission photoélectrique commence dès que la substance est éclairée, circonstance qui écarte l'idée d'une accumulation progressive d'énergie puisée peu à peu par la matière dans une onde lumineuse où l'énergie serait uniformément répartie.

Réfléchissant à ce problème difficile, M. Einstein a proposé en 1905 de le résoudre en reprenant sous une forme nouvelle la vieille conception corpusculaire de la lumière. Il imagine que dans une onde lumineuse de fréquence ν l'énergie, au lieu d'être répartie uniformément comme le veulent les théories ondulatoires classiques, est pelotonnée sous forme de grains d'énergie de valeur $h\nu$, l'intensité de l'onde lumineuse étant proportionnelle au nombre de ces grains par unité de volume ou, si l'on préfère, à la probabilité de trouver un de ces grains dans une unité de volume. Cette conception est suggérée par les idées de Planck sur les quanta qui, on le sait, introduisent une proportionnalité entre l'énergie et la fréquence, la constante de proportionnalité étant la constante h . Si l'on admet l'existence de ces quanta d'énergie ou « photons » dans tout rayonnement monochromatique, les lois essentielles de l'effet photoélectrique en découlent immédiatement. En effet, si ω est l'énergie qu'un électron doit dépenser pour sortir du métal irradié, l'absorption par cet électron d'un quantum d'énergie $h\nu$ du rayonnement incident pourra provoquer sa sortie hors du métal avec l'énergie cinétique $h\nu - \omega$. Pour que l'expulsion photoélectrique puisse se produire, il faut donc que la fréquence ν soit supérieure à la fréquence $\nu_0 = \frac{\omega}{h}$ qui est la fréquence seuil décelée par l'expérience.

De plus, si ν est supérieure à ν_0 , l'énergie cinétique du photoélectron sera

$$T = h\nu - w = h(\nu - \nu_0).$$

C'est la relation d'Einstein pour l'effet photoélectrique, relation qui montre que la constante k de la loi empirique doit être égale à la constante h de Planck : ce fait a été bien vérifié d'abord pour la lumière, puis pour les radiations X et γ .

En ce qui concerne la lumière, l'étude quantitative la plus précise de l'effet photoélectrique est due à Millikan (1916). Il éclairait une surface propre de lithium ou de sodium avec les radiations monochromatiques d'un arc au mercure. L'énergie cinétique des photoélectrons était mesurée en déterminant la différence de potentiel minimum nécessaire pour les empêcher d'atteindre une électrode auxiliaire. Il fallait d'ailleurs tenir compte de la différence de potentiel de contact entre le métal éclairé et l'électrode auxiliaire, laquelle dépend de l'état des surfaces, etc. Quand on tient compte de ces corrections, on obtient pour représenter T en fonction de ν des droites dont le coefficient angulaire donne bien pour la constante de Planck une valeur voisine de celle que l'on connaît par ailleurs.

La vérification de la relation photoélectrique d'Einstein est donc bonne dans le domaine de la lumière, mais elle y est rendue délicate par le fait que les $h\nu$ sont petits, que les w sont de l'ordre des $h\nu$ et dépendent de circonstances difficiles à contrôler exactement comme l'état de propreté des surfaces métalliques, etc. La vérification de la loi d'Einstein est beaucoup plus facile et plus nette dans le domaine des hautes fréquences (rayons X et γ). Tout d'abord, ici, les quanta $h\nu$ sont très grands : ensuite, les photoélectrons expulsés sont, pour la plupart, des électrons arrachés à la structure profonde des atomes irradiés et non des électrons libres dans le métal et, par suite, le travail de sortie w est pour eux sensiblement égal au travail d'arrachement de l'électron hors de l'atome, travail devant lequel le travail de sortie hors du métal est négligeable (1 V devant 50 000 V, par exemple). D'ailleurs, le travail nécessaire pour arracher les diverses catégories d'électrons hors des atomes est fourni avec une grande précision par la spectrographie des rayons X. Il est donc facile de vérifier la loi d'Einstein. C'est ce qui a été fait pour les rayons X par Maurice de Broglie en 1921 et ensuite pour les rayons γ par Ellis et Jean Thibaud.

On s'est servi pour cela d'un « spectrographe magnétique » présentant la disposition suivante (*fig. 17*) :

La radiation incidente tombe à 45° sur un film d'un corps choisi comme radiateur photoélectrique. Les électrons qui s'en échappent pénètrent dans une chambre bien évacuée où règne un champ magnétique uniforme H normal au plan de la figure qui impose aux électrons une trajectoire circulaire, puis ils viennent frapper une plaque photographique en y produisant une impression. On constate alors sur la plaque photographique l'enregistrement de raies dont chacune correspond à l'impact d'un flot d'électrons. La position de ces raies sur la plaque permet de mesurer le rayon du cercle décrit par les électrons du flot correspondant et d'en déduire par le calcul l'énergie cinétique de ces électrons. Pour chaque radiation incidente monochromatique de fréquence ν , on décèle ainsi l'existence de flots d'électrons correspondant aux énergies cinétiques .

$$h\nu - \omega_1, \dots, h\nu - \omega_i, \dots,$$

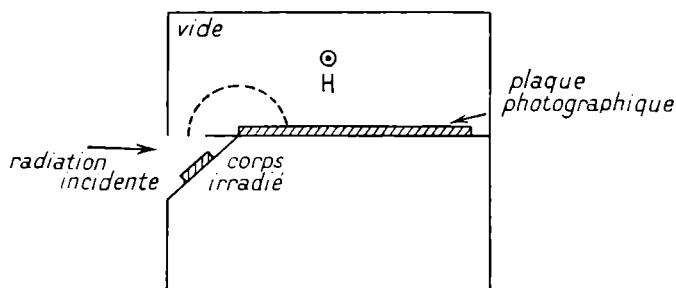


Fig. 17.

où les ω_i sont les travaux de sortie hors de l'atome des diverses catégories d'électrons intra-atomiques pour le corps choisi comme radiateur photoélectrique. La vérification ainsi obtenue pour la relation d'Einstein est excellente et très précise.

Il est très intéressant de noter que, pour les rayons X durs et les rayons γ , les quanta $h\nu$ et, par suite, les énergies cinétiques des photoélectrons sont grands, de sorte que la vitesse de ces photoélectrons devient de l'ordre de la vitesse de la lumière. Pour obtenir une bonne concordance avec les résultats de l'expérience, il est alors nécessaire, on le constate, d'employer pour calculer l'énergie cinétique des électrons à partir du rayon du cercle qu'ils décrivent dans le champ magnétique, non pas les formules de la Mécanique classique, mais celles de la Mécanique relativiste avec variation de la masse avec la vitesse. On obtient ainsi, en passant, une vérification remarquable de la Dynamique relativiste.

3. L'hypothèse des photons et ses succès. — L'hypothèse d'Einstein revient à supposer qu'au moins dans certaines circonstances la lumière se comporte comme si elle était formée de particules qu'on nomme aujourd'hui « photons ». Le photon lié à une onde plane monochromatique de fréquence ν est défini par la relation

$$W = h\nu.$$

Cela suffit déjà à prouver que la théorie des Photons ne constitue pas une théorie corpusculaire de la Lumière ayant un caractère autonome, puisque dans cette formule fondamentale figure la fréquence ν à laquelle seule une théorie introduisant une idée de périodicité peut donner un sens. Il faudra donc chercher à faire une synthèse de la théorie ondulatoire qui permet de définir ν et de cette nouvelle théorie corpusculaire.

Einstein a montré, notamment par des considérations sur l'équilibre entre les atomes matériels et le rayonnement dans une enceinte à température uniforme, que les photons devaient posséder une quantité de mouvement égale à

$$\mathbf{p} = \frac{h\nu}{c} \mathbf{n} = \frac{h}{\lambda} \mathbf{n},$$

où \mathbf{n} est un vecteur unité porté dans le sens de la propagation.

Cette formule peut s'obtenir aussi à partir des formules de la Dynamique relativiste dont l'application au photon est évidemment nécessaire. Pour un corpuscule de masse propre μ_0 , ces formules donnent en effet

$$W = \frac{\mu_0 c^2}{\sqrt{1-\beta^2}}, \quad \mathbf{p} = \frac{\mu_0 \mathbf{v}}{\sqrt{1-\beta^2}}, \quad \mathbf{p} = \frac{W}{c^2} \mathbf{v}.$$

Supposons que nous envisageons des corpuscules de masses propres μ_0 de plus en plus petites en faisant en même temps croître la vitesse de façon que W garde une valeur constante. A la limite pour $\mu_0 \rightarrow 0$ et $v \rightarrow c$, nous obtiendrons

$$\mathbf{p} \simeq \frac{W}{c^2} \mathbf{c} = \frac{W}{c} \mathbf{n}.$$

Donc, si l'on admet que $W = h\nu$, on trouve bien pour l'impulsion du photon la formule d'Einstein. Le photon nous apparaît ainsi comme une particule de masse propre évanouissante se déplaçant sensiblement avec la vitesse c ou, si l'on croit devoir passer à la limite, comme une

particule de masse propre nulle se déplaçant exactement avec la vitesse c .

Nous avons vu comment la vérification de la loi photoélectrique d'Einstein apportait une confirmation de la relation $W = h\nu$. Nous allons parler maintenant d'autres confirmations. L'une d'elles est la loi qui donne la limite du spectre continu des rayons X. Dans un tube à rayons X, l'anticathode est frappée par un flux d'électrons qui ont été accélérés par une même chute de potentiel V et ont tous une même énergie cinétique $T = eV$. L'anticathode émet alors un spectre de raies de rayons X dont la théorie relève de la théorie de l'atome et un spectre continu où sont représentées toutes les fréquences jusqu'à une fréquence maximum ν_m à laquelle correspond une longueur d'onde minimum $\lambda_m = \frac{c}{\nu_m}$, comme l'indique la figure 18.

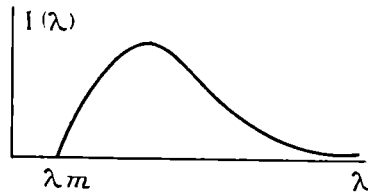


Fig. 18.

L'expérience prouve que l'on a (loi de Duane et Hunt)

$$\nu_m = \frac{T}{h} = \frac{eV}{h}.$$

Cette loi s'interprète immédiatement en admettant que, lors du ralentissement d'un électron dans la matière de l'anticathode, les rayons X sont émis par photons. Si, en effet, quand l'électron perd une fraction ΔT de son énergie cinétique, il y a émission d'un photon de fréquence $\nu = \frac{\Delta T}{h}$, la fréquence maxima présente dans le spectre continu sera émise dans le cas où un électron perd d'un seul coup la totalité de son énergie cinétique et cette fréquence aura pour valeur $\nu_m = \frac{T}{h}$, ce qui est bien la loi de Duane-Hunt.

La relation $W = h\nu$ a été aussi confirmée par le succès de la « loi des fréquences » de Bohr. Dans sa théorie de l'atome que nous étudierons dans le prochain chapitre, M. Bohr a admis que les atomes et autres édifices matériels de l'échelle atomique ne peuvent avoir que certains états énergétiques stables, les états stationnaires ou états quantifiés,

dont la détermination doit se faire suivant certaines règles de quantification où intervient la constante h . Soient W_1, \dots, W_j, \dots les énergies des états quantifiés d'un système atomique. M. Bohr a admis que le système quantifié peut passer brusquement de l'état d'énergie W_j à un état quantifié d'énergie moindre $W_k < W_j$ avec émission d'une radiation de fréquence

$$\nu_{jk} = \frac{W_j - W_k}{h}$$

Ceci se comprend immédiatement si l'on admet que l'énergie perdue par le système matériel est émise vers l'extérieur sous la forme d'un photon. De même, si le système se trouve initialement dans l'état quantifié d'énergie W_k et s'il est irradié par un rayonnement ayant la fréquence ν_{jk} , il peut absorber un photon d'énergie $h\nu_{jk}$ et passer ainsi dans l'état d'énergie supérieure W_j . On voit donc bien la liaison étroite qui existe entre la loi des fréquences de Bohr et l'idée de photon.

L'effet Raman apporte une autre vérification indirecte de la théorie des Photons. En 1928, le physicien hindou Venkata Raman, en éclairant des substances telles que le benzène par une radiation monochromatique de fréquence ν du spectre visible, a montré que la lumière diffusée contient, en plus de la fréquence ν elle-même, des fréquences de la forme $\nu - \nu_{jk}$, où les ν_{jk} sont des fréquences infrarouges figurant dans le spectre des molécules du corps diffuseur. On est ensuite arrivé à montrer que la radiation diffusée contenait, mais avec des intensités beaucoup plus faibles, des fréquences de la forme $\nu + \nu_{jk}$.

Tout ceci est facile à expliquer avec les idées quantiques et l'hypothèse du photon.

Soient encore W_j et $W_k < W_j$ l'énergie de deux états quantifiés de la molécule du corps diffuseur : cette molécule peut donc émettre ou absorber des photons de fréquence $\nu_{jk} = \frac{W_j - W_k}{h}$, et la fréquence ν_{jk} figure dans son spectre d'émission (infrarouge). Supposons maintenant que cette substance soit irradiée par de la lumière visible de fréquence ν , c'est-à-dire reçoive des photons d'énergie $h\nu$. Une molécule se trouvant dans l'état d'énergie W_j pourra alors, au moment de la diffusion, céder l'énergie $W_j - W_k = h\nu_{jk}$ au photon en passant dans l'état d'énergie W_k et le photon diffusé dont l'énergie sera $h(\nu + \nu_{jk})$ correspondra à la fréquence $\nu + \nu_{jk}$. De même, si la molécule diffusante se trouve initialement dans l'état d'énergie W_k , elle pourra emprunter au photon l'énergie $h\nu_{jk}$ pour passer dans l'état d'énergie supérieure W_j et le photon diffusé n'aura plus que la fréquence $\nu - \nu_{jk}$. Maintenant, dans la

substance diffusante, il y a normalement beaucoup plus de molécules dans l'état d'énergie inférieure W_k que dans l'état d'énergie supérieure W_l , comme cela se déduit de la loi de Gibbs-Boltzmann en $e^{-\frac{W}{kT}}$. D'où une intensité beaucoup plus forte pour la raie diffusée de fréquence $\nu - \nu_{ik}$ que pour la raie diffusée de fréquence $\nu + \nu_{ik}$. Cette conclusion est d'ailleurs en accord avec la loi générale dite « règle de Stokes », d'après laquelle *dans les phénomènes d'interaction entre matière et lumière il y a une tendance à l'abaissement des fréquences*.

Toutes les particularités de l'effet Raman se trouvent ainsi expliquées par l'interprétation quantique. On peut aussi donner une interprétation classique de ce phénomène, mais cette interprétation est incomplète et, en particulier, elle conduit à attribuer à la raie de fréquence $\nu + \nu_{ik}$ une intensité plus grande qu'à la raie de fréquence $\nu - \nu_{ik}$, contrairement à l'expérience.

4. L'effet Compton. — Nous venons de résumer quelques preuves à l'appui de la relation $W = h\nu$. Nous allons maintenant voir que l'effet Compton a permis non seulement de la vérifier une fois de plus, mais aussi de vérifier la relation $p = \frac{h\nu}{c}$.

L'effet découvert en 1923 par le physicien américain H. A. Compton est un phénomène de diffusion de radiation. Il est néanmoins très différent de l'effet Raman. L'effet Raman se produit pour la lumière; le changement de fréquence correspondant dépend de la nature du diffuseur et est indépendant de l'azimut de la diffusion. L'effet Compton, au contraire, se produit seulement pour les rayons X et γ ; le changement de fréquence correspondant est indépendant de la nature du radiateur, mais il varie avec l'azimut d'observation.

On obtient une théorie très simple et très satisfaisante de l'effet Compton en se plaçant au point de vue de la théorie des photons (Compton, Debye). On admet alors que les photons X ou γ , en traversant la matière, y rencontrent des électrons libres qui sont au repos ou sensiblement au repos. Cette rencontre s'accompagne en général d'une interaction, d'un « choc », avec échange d'énergie et de quantité de mouvement. Après le choc, le photon a perdu une partie de son énergie et est diffusé dans une direction généralement différente de sa direction initiale, avec une fréquence inférieure à sa fréquence initiale; en même temps l'électron, par effet de recul, est projeté avec une certaine vitesse.

Sur la figure 19, on a supposé l'électron initialement au repos en O et le photon incident possédant la quantité de mouvement $\frac{h\nu_0}{c}$. Après le choc, l'électron part dans la direction φ avec l'impulsion $\frac{m_0\beta c}{\sqrt{1-\beta^2}}$ et le photon dans la direction θ avec l'impulsion $\frac{h\nu_\theta}{c}$. La conservation de l'énergie et celle des deux composantes de la quantité de mouvement donnent les équations

$$h\nu_0 + m_0c^2 = h\nu_\theta + \frac{m_0c^2}{\sqrt{1-\beta^2}}, \quad \frac{h\nu_0}{c} \cos \theta + \frac{m_0\beta c}{\sqrt{1-\beta^2}} \cos \varphi = \frac{h\nu_0}{c},$$

$$\frac{h\nu_0}{c} \sin \theta = \frac{m_0\beta c}{\sqrt{1-\beta^2}} \sin \varphi.$$

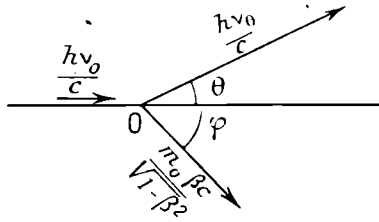


Fig. 19.

En éliminant entre ces trois relations, d'abord β et φ , puis β et ν_0 , on en tire

$$(1) \quad \nu_0 = \frac{\nu_\theta}{1 + \alpha(1 - \cos \theta)} = \frac{\nu_\theta}{1 + 2\alpha \sin^2 \frac{\theta}{2}},$$

$$(2) \quad 1 - \cos \theta = \frac{2}{1 + \operatorname{tg}^2 \varphi (1 + \alpha)^2} \quad \text{ou} \quad \operatorname{cotg} \varphi = (1 + \alpha) \operatorname{tg} \frac{\theta}{2},$$

avec l'abréviation

$$\alpha = \frac{h\nu_0}{m_0c^2}.$$

La relation (1) donne la fréquence du photon en fonction de l'angle θ de diffusion.

La variation de fréquence croît avec l'angle θ et avec la fréquence de la radiation incidente. La variation correspondante de la longueur d'onde est

$$\delta\lambda = \lambda_\theta - \lambda_0 = 2\alpha\lambda_0 \sin^2 \frac{\theta}{2} = 2 \frac{h}{m_0c} \sin^2 \frac{\theta}{2}.$$

Elle présente ceci de remarquable qu'elle est indépendante de la

longueur d'onde initiale λ_0 . Ces formules sont remarquablement vérifiées par l'expérience.

Quant à la formule (2), elle donne la relation existant entre l'angle de diffusion θ et l'angle φ de mise en mouvement de l'électron de recul. Dans une chambre de Wilson, on observe sur le passage d'un faisceau de rayons X une chevelure de courtes trajectoires électroniques : ce sont les trajectoires des électrons de recul (fish tracts) et la tangente à l'origine de l'une de ces trajectoires donne son angle φ (*fig. 20*).

L'angle θ est plus difficile à déterminer parce que le photon X diffusé ne laisse pas de trace dans la chambre de Wilson. Heureusement, il arrive que sur certains clichés on aperçoive à la fois un fish tract en O et une longue trajectoire électronique débutant brusquement en P en

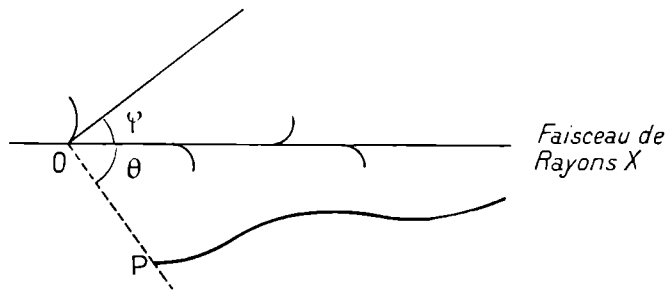


Fig. 20.

dehors du faisceau de rayons X. On en conclut qu'un effet Compton s'est produit en O et que le photon a été diffusé dans la direction OP et a produit en P un effet photoélectrique. L'angle θ se trouve ainsi déterminé et l'on peut vérifier la relation (2). Cette vérification faite par Compton et Simon montre que, dans le phénomène d'interaction entre photon et électron, il y a bien conservation de l'énergie et de la quantité de mouvement. Cette conclusion, corroborée par des expériences de Bothe et Geiger faites à l'aide d'un compteur à pointe, a conduit à abandonner l'hypothèse, un instant soutenue par Bohr, Slater et Kramers, suivant laquelle la conservation de l'énergie et de la quantité de mouvement aurait lieu seulement en moyenne. Il semble bien établi que cette conservation a lieu pour chaque phénomène individuel.

5. Difficultés soulevées par la théorie des Photons. — La théorie des « quanta de lumière » d'Einstein, aujourd'hui nommée théorie des Photons, a, dès son apparition en 1905, soulevé de graves objections.

Non seulement, nous l'avons vu, elle ne se suffit pas à elle-même puisqu'elle définit l'énergie du photon à l'aide de la fréquence ν qui est une notion extraite de la théorie ondulatoire, mais elle ne peut évidemment pas expliquer les vérifications innombrables et extrêmement précises de la théorie ondulatoire, notamment dans le domaine des interférences et de la diffraction. On peut bien imaginer évidemment qu'il y ait dans la lumière une association de corpuscules et d'ondes, mais comment préciser cette association ?

Une première idée serait de supposer que le photon est formé par l'ensemble d'un train d'ondes. Mais il existe des trains d'ondes lumineuses qui sont cohérents et ont une longueur de plusieurs mètres. On ne peut supposer qu'un tel train d'ondes forme un photon, car cela reviendrait à dire qu'un photon, particule insécable et bien localisée, a une longueur de plusieurs mètres. Lorentz avait présenté cette objection sous une autre forme en remarquant que dans une lunette astronomique recevant la lumière très faible venant d'une étoile éloignée, il se produit des effets de diffraction que l'on peut prévoir très exactement à l'aide de la théorie ondulatoire de la lumière : dans cette prévision, on suppose (ce qui est naturel en théorie ondulatoire) que l'onde incidente couvre toute la surface de l'objectif de la lunette, surface qui peut être de l'ordre du mètre carré dans les grandes lunettes. Or, la lumière venant de l'étoile étant très faible, les photons doivent arriver un par un sur la lunette et, pour pouvoir interpréter l'existence des phénomènes de diffraction, il semble qu'il faille supposer le photon assez large pour couvrir toute la surface de l'objectif : le photon aurait alors des dimensions transversales qui seraient de l'ordre du mètre, ce qui paraît encore incompatible avec ses propriétés d'insécabilité et de localisation.

Ne pouvant assimiler le photon à l'ensemble d'un train d'ondes, on pourrait chercher à interpréter les interférences et la diffraction en supposant qu'elles proviennent, non d'une propagation ondulatoire, mais d'une interaction entre les photons. Les faisceaux lumineux usuellement employés dans les laboratoires sont assez intenses pour contenir un grand nombre de photons : on pourrait tenter (ce ne serait pas facile !) d'interpréter les phénomènes d'interférences et de diffraction par des interactions entre ces photons. Mais il y a à cela une objection fondamentale. Les phénomènes d'interférence et de diffraction devraient alors dépendre de l'intensité de la lumière et disparaître pour une intensité très faible : or il n'en est rien. Nous avons déjà remarqué pour la lunette astronomique que, même dans le

cas des photons arrivant un par un d'une étoile éloignée, les phénomènes de diffraction se produisent. Des expériences précises ont été faites sur ce point capital par Taylor d'abord, puis par Dempster et Batho. Ils ont employé des appareils d'interférences éclairés par une lumière si faible que les photons devaient en moyenne arriver un par un sur l'appareil avec de grands intervalles de temps. En utilisant un enregistrement photographique, ils pouvaient prolonger l'expérience pendant des heures, de sorte que finalement l'énergie lumineuse recueillie au total par l'appareil devenait notable. Or ils ont trouvé de cette façon que les franges d'interférence étaient en fin de compte exactement les mêmes que si l'on avait opéré avec une lumière intense.

Il faut renoncer à expliquer les interférences par des interactions entre photon : il faut revenir à l'interprétation ondulatoire en admettant que chaque photon est associé à un train d'ondes, sans qu'on puisse cependant lui enlever son caractère ponctuel et l'assimiler à l'ensemble étendu du train d'ondes. Pour sortir de cette impasse, il a fallu adopter une interprétation d'un type nouveau qui tient compte entièrement du double aspect onde-corpuscule de la lumière, interprétation qui a été ensuite généralisée par la Mécanique ondulatoire au cas de toutes les particules matérielles. Nous allons insister sur cette interprétation qui est essentielle pour une bonne compréhension des théories actuelles de la Physique quantique.

6. Caractère probabiliste de la synthèse onde-corpuscule. — Nous commencerons par analyser le concept de corpuscule employé par les physiciens en discernant deux formes. On peut commencer par concevoir le corpuscule comme une unité susceptible de se manifester localement par des actions où il intervient tout entier ; on peut ensuite aller plus loin et imaginer le corpuscule décrivant au cours du temps une trajectoire dans l'espace avec une vitesse bien déterminée à chaque instant.

Le développement de la Physique quantique conduit à considérer le premier point de vue comme beaucoup plus essentiel que le second : ce qui caractérise essentiellement le corpuscule, c'est sa faculté de manifester sa présence localement comme une unité physique indécomposable. Quand on envisage l'évolution du corpuscule au cours du temps, il est souvent possible de lui attribuer *grosso modo* une trajectoire et une vitesse, mais cette propriété est beaucoup moins essentielle que la première et dans beaucoup de cas disparaît complètement : la raison en est que pour déterminer trajectoire et vitesse,

il faut à chaque instant repérer le corpuscule et ce repérage nécessite de faire entrer le corpuscule en interaction avec quelque dispositif de mesure, ce qui, dans certains cas, trouble complètement le mouvement du dit corpuscule.

Quand nous étudierons les incertitudes d'Heisenberg, nous verrons, en nous appuyant sur l'existence des quanta, qu'il est impossible de mesurer simultanément avec une précision absolue la position et la vitesse d'un corpuscule, ce qui exclut toute détermination complète de la trajectoire.

D'une façon générale, il faut abandonner l'idée que la conception du corpuscule entraîne celle d'une trajectoire et d'une vitesse bien déterminées, pour ne retenir que celle d'unité physique insécable pouvant se manifester localement en un point. Ainsi, pour le photon, dans les phénomènes d'optique géométrique, on peut considérer le rayon lumineux comme la trajectoire du photon, mais c'est à condition de ne pas vouloir trop bien définir le rayon, car, en étranglant un rayon lumineux pour mieux le délimiter, on fait apparaître des phénomènes de diffraction et la notion même de rayon disparaît. Dans les champs d'interférence, on ne peut plus définir un rayon-trajectoire du photon et cependant le photon peut toujours y manifester sa présence par un effet photoélectrique localisé, ce qui montre la supériorité de la première définition du corpuscule sur la seconde.

Ne gardant donc plus du corpuscule que l'idée d'unité physique insécable et susceptible de produire un effet localisé, nous pouvons alors associer à cette idée celle d'onde grâce à des conceptions qui font intervenir les probabilités. Dans une région où se trouve un photon, il pourra y avoir une action locale de ce photon et l'onde lumineuse représentera les probabilités pour que ce photon se localise ici ou là : d'une façon plus précise, la probabilité de localisation du photon en un point sera mesurée par l'intensité (carré de l'amplitude) de l'onde lumineuse en ce point. L'onde lumineuse nous apparaît alors comme une sorte de champ de probabilité dont l'extension provient justement du fait que le photon ne peut pas être localisé à chaque instant en un point, mais est susceptible de manifester sa présence par une action énergétique ponctuelle dans toute une région étendue de l'espace. Cette conception toute nouvelle des rapports entre l'onde (ici l'onde lumineuse) et le corpuscule (ici le photon) va prendre toute sa généralité dans le cadre de la Mécanique ondulatoire.

Avec cette conception nouvelle, on interprète facilement le résultat de l'expérience de Taylor. Chaque photon arrivant sur l'appareil

d'interférences est accompagné d'un train d'ondes lumineuses, champ de probabilité représentant ses localisations possibles. Pénétrant dans l'appareil, l'onde lumineuse interfère et la probabilité pour que le photon produise un effet photoélectrique en un point de la plaque photographique d'enregistrement est proportionnelle à l'intensité de l'onde lumineuse en ce point. La même chose se reproduit à l'arrivée de chaque photon. Finalement, au bout d'un temps très long, quand seront arrivés dans l'appareil un grand nombre de photons, les effets photoélectriques produits aux différents points de la plaque photographique par les photons successifs seront répartis proportionnellement au carré de l'amplitude de l'onde lumineuse dans le champ d'interférences. Les probabilités relatives des divers phénomènes élémentaires se traduisent en fin de compte par une répartition *statistique*. Ainsi se dessinent sur la plaque d'enregistrement les franges claires et les franges obscures prévues par la théorie classique de Fresnel, mais les franges claires sont tracées point par point par les arrivées successives égrenées dans le temps des divers photons et non pas par l'arrivée continue d'une énergie radiante continûment répartie. Les apparences finalement observables sont donc les mêmes que celles que prévoyait la théorie classique, mais la façon dont elles sont obtenues est essentiellement différente. Dans le phénomène statistique finalement enregistré, l'onde lumineuse apparaît en quelque sorte comme étant la réalité physique, alors que cependant elle n'est qu'un champ de probabilité pour les manifestations discontinues et localisées des divers photons.

Cette interprétation assez subtile paraît bien être la seule qui permette de réunir dans un schéma théorique cohérent l'existence des photons et les vérifications des prévisions de la théorie ondulatoire et d'expliquer pourquoi on peut obtenir les mêmes phénomènes d'interférences avec une lumière intense et une pose courte ou avec une lumière très faible et une très longue pose. Il importe de bien réfléchir à ces idées si l'on veut bien comprendre le sens des théories quantiques actuelles.

BIBLIOGRAPHIE.

- [1] CHRISTIAN HUYGHENS, *Traité de la Lumière*, Gauthier-Villars, Paris, 1920.
- [2] ISAAC NEWTON, *Traité d'optique*, 1704.
- [3] AUGUSTIN FRESNEL, *Oeuvres complètes*, édition Verdet, 1866.
- [4] ALBERT EINSTEIN, *Ann. Physik*, t. 17, 1905, p. 132.
- [5] MAURICE ET LOUIS DE BROGLIE, *Introduction à la Physique des rayons X et γ* , Gauthier-Villars, Paris, 1928.

CHAPITRE VII.

LA THÉORIE QUANTIQUE DE L'ATOME DE BOHR-SOMMERFELD.

1. La formule de Balmer, les termes spectraux et le principe de combinaison. — La série de raies spectrales la plus anciennement connue dans le spectre de l'hydrogène est la série de Balmer. Elle est formée de quatre raies principales, qui sont en réalité de petits doublets, dont voici les noms et les longueurs d'onde :

$$H_{\alpha} : 6\,563 \text{ \AA}, \quad H_{\beta} : 4\,861 \text{ \AA}, \quad H_{\gamma} : 4\,340 \text{ \AA}, \quad H_{\delta} : 4\,102 \text{ \AA}.$$

Il y a 70 ans, Balmer est parvenu à trouver une formule empirique qui donne les fréquences de ces raies. Cette formule est la suivante :

$$\nu_m = R \left[\frac{1}{4} - \frac{1}{m^2} \right] \quad (m = 3, 4, 5, 6),$$

ν_3 est la fréquence de H_{α} , ν_4 celle de H_{β} , etc. R est une constante dite « *constante de Rydberg* », très sensiblement égale à $3,29201 \cdot 10^{14}$ en C. G. S.

D'autres séries de raies découvertes postérieurement dans le spectre non visible de l'hydrogène suivent des lois analogues. Telles sont notamment la série ultraviolette de Lyman et la série infrarouge de Paschen dont les fréquences sont respectivement données par

$$\nu_m = R \left[1 - \frac{1}{m^2} \right] \quad (m = 2, 3, 4, \dots),$$

$$\nu_m = R \left[\frac{1}{9} - \frac{1}{m^2} \right] \quad (m = 4, 5, \dots).$$

On voit que toutes ces formules rentrent dans le type général

$$(1) \quad \nu = R \left[\frac{1}{n^2} - \frac{1}{m^2} \right] \quad (m > n).$$

Pour la série de Lyman, $n = 1$; pour la série de Balmer, $n = 2$; pour la série de Paschen, $n = 3$, etc.

De cette constatation, on peut s'élever à une loi générale qui s'est montrée être exacte pour toutes les raies spectrales de tous les corps. C'est le « *principe de combinaison de Ritz* » dont voici l'énoncé :

La fréquence de toute raie spectrale est égale à la différence de deux termes spectraux caractéristiques du corps considéré.

Ou encore sous une autre forme :

Pour tout corps émetteur, on peut dresser un tableau de nombres, dits « termes spectraux », tels que la fréquence de toute raie spectrale du corps soit donnée par la différence de deux termes spectraux.

La formule (1) nous montre alors que pour l'hydrogène, les termes spectraux ont, au moins en valeur absolue, la forme $\frac{R}{n^2}$, avec $n = 1, 2, 3, \dots$

Une étude expérimentale plus approfondie des raies de la série de Balmer a d'ailleurs montré qu'elles sont formées de raies très voisines. L'interprétation de cette « structure fine » du spectre de l'hydrogène a été, nous le verrons, le but de la théorie de Sommerfeld. Ajoutons que, pour les corps autres que l'hydrogène, il existe toujours des termes spectraux, mais que leur forme est plus compliquée que la forme $\frac{R}{n^2}$. Rydberg et Ritz ont donné, il y a longtemps, des formules empiriques pour représenter ces termes spectraux.

2. La théorie de l'atome quantifié de Bohr. — En 1913, M. Bohr est parvenu à interpréter les termes spectraux de l'hydrogène et il a ainsi fondé toute la théorie moderne de l'atome.

A cette époque, à la suite des expériences de Rutherford sur la déviation subie par les particules α quand elles traversent la matière, les physiciens venaient de se rallier à un modèle planétaire de l'atome. Selon cette vue, l'atome d'un corps simple dont le numéro d'ordre dans la série de Mendéléeff est Z serait muni d'un noyau central portant la charge électrique $+Ze$, où e est la charge élémentaire $4,8 \cdot 10^{-10}$ u. e. s. Autour de ce noyau, graviteraient Z électrons de charge $-e$, de sorte que l'ensemble de l'atome serait électriquement neutre. M. Bohr a eu l'idée de soumettre ce modèle au calcul en lui appliquant les lois de quanta introduites avec succès par Planck dans l'étude du rayonnement noir.

Bohr a admis un premier postulat, *le postulat des états stationnaires quantifiés*, d'après lequel un électron ne peut décrire dans l'atome autour du noyau central que certaines trajectoires déterminées par une règle de quantification inspirée des travaux de Planck. A ces mouvements quantifiés, seuls stables, de l'électron, correspondent pour l'atome des « états stationnaires » durant lesquels, contrairement aux prévisions de la théorie électromagnétique classique, aucun rayonnement n'est émis. L'émission des raies spectrales ne peut donc avoir lieu que lors du passage brusque de l'atome d'un état stationnaire à un autre état stationnaire de moindre énergie. Mais quelle sera la fréquence de la raie émise pendant un tel passage ?

Pour répondre à cette question, Bohr introduit un deuxième postulat, *la loi des fréquences*, en admettant que l'énergie perdue par l'atome au moment de la transition est rayonnée sous la forme d'un seul quantum de lumière de valeur $h\nu$, c'est-à-dire d'un seul photon. Si donc E_1 et E_2 désignent les énergies de l'atome dans l'état stationnaire initial et dans l'état stationnaire final, la fréquence ν_{12} émise lors de la transition $1 \rightarrow 2$ sera

$$(2) \quad \nu_{12} = \frac{E_1 - E_2}{h}.$$

Cette simple formule explique immédiatement le sens du principe de combinaison de Ritz et montre que les termes spectraux d'un atome sont égaux aux énergies de ses états stationnaires divisées par la constante de Planck.

La réversibilité des processus élémentaires exige d'ailleurs que, si un atome se trouvant dans l'état d'énergie E_2 est frappé par un rayonnement de fréquence ν_{12} , il puisse absorber un photon de ce rayonnement pour passer à l'état d'énergie supérieure E_1 , ce processus d'absorption étant ainsi exactement l'inverse du processus d'émission.

Le problème essentiel qui se posait à Bohr était donc de déterminer les énergies des états stationnaires. Pour cela, il a admis dans son travail initial que l'électron se comporte comme une charge ponctuelle obéissant aux lois de la Dynamique de Newton, mais il restreint le nombre des mouvements possibles en introduisant les règles de quantification de Planck. A l'époque où Bohr écrivait, la question se présentait de la façon suivante. Planck, ayant reconnu que la constante h , qui a les dimensions ML^2T^{-1} d'une Action au sens de la Mécanique, constituait une sorte d'unité d'action, avait remplacé son énoncé primitif de quantification de l'énergie, valable seulement pour

un oscillateur harmonique ⁽¹⁾, par un énoncé de quantification de l'action qu'il n'avait su d'ailleurs donner que dans le cas d'un mouvement périodique à un seul degré de liberté. Pour ce cas, la méthode de quantification de Planck était la suivante : p étant le moment de Lagrange conjugué de l'unique variable q , on devait écrire

$$\oint p \, dq = nh \quad (n \text{ entier}),$$

l'intégrale étant étendue à un cycle entier du mouvement. C'est cette formule que Bohr a tout naturellement cherché à employer pour déterminer les états stables quantifiés de l'atome.

Il a envisagé les trajectoires quantifiées de l'unique électron-planète de l'atome d'hydrogène ($Z=1$) autour du noyau (proton) de charge $+e$. Pour n'introduire qu'une seule variable q , il s'est astreint à n'envisager que les orbites circulaires de rayon r sur lesquelles la position de l'électron est déterminée par une seule variable, l'azimut θ . La condition de Planck donne alors

$$m r^2 \dot{\theta} = n \frac{h}{2\pi} \quad \left(\dot{\theta} = \frac{d\theta}{dt} \right).$$

Elle exprime que le moment de la quantité de mouvement de l'électron sur l'orbite stable doit être un multiple entier de $\frac{h}{2\pi}$. D'autre part, les lois de la Dynamique classique nous donnent la relation

$$m r \dot{\theta}^2 = \frac{e^2}{r^2}.$$

On trouve aisément pour l'énergie du $n^{\text{ième}}$ mouvement circulaire quantifié

$$E_n = \frac{1}{2} m r^2 \dot{\theta}^2 - \frac{e^2}{r} = -\frac{1}{2} m r^2 \dot{\theta}^2 = -\frac{2\pi^2 m e^4}{n^2 h^2}.$$

Donc, si l'on s'en tient aux mouvements circulaires, les termes spectraux de l'hydrogène doivent être de la forme

$$(3) \quad \frac{E_n}{h} = -\frac{2\pi^2 m e^4}{n^2 h^3}.$$

D'après la formule (2), les raies de l'hydrogène doivent avoir des fréquences données par la relation générale

$$\nu_{nn'} = \frac{2\pi^2 m e^4}{h^3} \left(\frac{1}{n^2} - \frac{1}{n'^2} \right) \quad (n' > n).$$

(1) Voir le début du chapitre suivant.

On retombe sur la formule empirique (1) en posant

$$R = \frac{2\pi^2 m e^4}{h^3}.$$

Or, le calcul de R d'après cette relation donne une valeur sensiblement égale à la valeur expérimentale de la constante de Rydberg, ce qui fut un grand succès pour les idées de Bohr.

On peut recommencer le même calcul en supposant que l'on a affaire à un atome de nombre atomique Z ionisé $Z - 1$ fois et comportant, par suite, un seul électron tournant autour d'un noyau de charge Ze (atome hydrogénoïde). On trouve alors à la place de (3)

$$\frac{E_n}{h} = - \frac{2\pi^2 m e^4}{n^2 h^3} Z^2 = - \frac{R}{n^2} Z^2,$$

c'est-à-dire que les termes spectraux sont multipliés par le carré du nombre atomique. Le cas le plus simple est celui de Phélium pour lequel $Z = 2$: les termes spectraux et les fréquences sont alors quadruplés. Les nombres expérimentaux indiquent toutefois que la constante R n'a pas tout à fait la même valeur pour H et pour He^+ . Mais M. Bohr a pu rendre compte brillamment de cette divergence en tenant compte de la réaction de l'électron sur le noyau, plus importante dans le cas du noyau léger de H que dans le cas du noyau quatre fois plus lourd de He .

3. Succès et insuffisances de la théorie de Bohr. — La théorie de l'atome de Bohr nous a fait comprendre la véritable structure des édifices atomiques et nous a appris que les quanta γ jouent un rôle essentiel. Les notions d'états stationnaires, de transition brusque d'un état stationnaire à un autre, notions si étrangères à la Physique classique, se sont dès lors montrées indispensables pour l'interprétation des phénomènes atomiques.

Il est devenu certain que les systèmes atomiques possèdent des niveaux discrets d'énergie et que les fréquences des raies spectrales s'obtiennent par la relation du quantum à partir des différences entre ces niveaux d'énergie. Cette idée a donné la clef de l'interprétation, non seulement de tous les spectres optiques, mais aussi des spectres de rayons X. La seule différence entre les deux cas est que dans le cas optique l'électron qui subit une transition est un électron périphérique, tandis que dans le cas des rayons X, c'est un électron de la structure profonde de l'atome. Un électron périphérique (électron optique) peut.

par suite d'une action extérieure (choc), être porté vers un niveau d'énergie normalement inoccupé : son retour par transition quantique vers des niveaux d'énergie plus faible donne alors lieu à l'émission de raies spectrales du spectre optique. En ce qui concerne l'interprétation des raies X, on a admis dès le début de la théorie de Bohr, que les niveaux internes de l'atome présentaient un phénomène de « saturation », c'est-à-dire qu'il ne pouvait pas y avoir plus d'un certain nombre maximum d'électrons par niveau. Ce principe est extrêmement important et nous aurons à y revenir lorsque nous parlerons du principe de Pauli : il est lié à la périodicité des propriétés physico-chimiques des éléments quand on s'élève dans la série de Mendéléeff (Z croissant). Au fur et à mesure que des électrons s'ajoutent à la structure de l'atome, il se constitue successivement, par suite de la saturation des niveaux, des couches successives d'électrons et, les propriétés des éléments dépendant principalement du nombre des électrons sur la couche périphérique, on conçoit que la constitution successive des couches provoque une périodicité des propriétés. Ces idées développées notamment par Kossel sont aujourd'hui bien connues et nous renvoyons à leur sujet aux traités de Physique.

Une conséquence intéressante de la saturation des niveaux dans l'atome est l'absence pour les rayons X du phénomène de renversement des raies si classique dans l'Optique ordinaire. C'est un principe bien connu de l'Optique lumineuse que si un corps est susceptible d'émettre une raie, il est aussi capable de l'absorber. Ceci s'explique immédiatement avec les idées de Bohr car, les raies optiques étant émises lors des transitions électroniques entre niveaux périphériques inoccupés, la transition est en principe possible dans les deux sens, de sorte que si l'émission d'une raie est possible, l'absorption l'est aussi. Il en est différemment pour les raies X : leur émission est provoquée par le fait que, dans l'anticathode du tube émetteur, l'arrivée des électrons du faisceau cathodique provoque l'arrachement de certains électrons de la structure interne de l'atome. Alors il se crée « une place vide » sur un niveau normalement occupé : un électron appartenant à un niveau plus extérieur peut, par une transition quantique, venir occuper cette place vide et cette transition s'accompagne de l'émission d'une raie X. Ces conceptions ont permis d'interpréter tous les détails de la spectrographie X apportant ainsi à la théorie de Bohr une très remarquable confirmation. Mais il est maintenant évident que, si l'on place sur le trajet d'un faisceau de rayons X une lame d'un corps capable d'émettre une des raies spectrales contenues dans le faisceau, il n'y a pas

absorption de la raie. En effet, dans les atomes de la lame interposée, tous les niveaux internes sont occupés au maximum et la transition d'absorption ne peut pas pour cette raison se produire. On voit sur cet exemple que les schémas tirés de la théorie de Bohr ont une grande puissance d'explication.

Ils rendent également très bien compte des phénomènes d'excitation et d'ionisation des atomes. Si des particules de même énergie bombardent un corps, elles pourront, si leur énergie est suffisante, faire passer l'atome d'un niveau d'énergie à un niveau d'énergie plus élevé, écartant un électron de son niveau normal et provoquant un état d'excitation de l'atome qui sera ainsi mis à même d'émettre certaines raies spectrales par retour de l'électron vers un niveau d'énergie inférieure. Si l'on bombarde le corps avec un faisceau de particules incidentes dont on fait croître progressivement l'énergie, on verra apparaître des effets d'excitation du corps chaque fois que l'on atteindra l'énergie correspondant à l'une des transitions quantiques d'excitation possibles. Si les particules incidentes sont des particules électrisées qui ont été accélérées par une différence de potentiel, on déterminera ainsi des « potentiels de résonance » correspondant en électron-volts à des différences d'énergie entre niveaux de Bohr. Il pourra même arriver que les particules incidentes aient assez d'énergie pour arracher complètement un des électrons de l'atome bombardé, mettant cet atome en état d'ionisation : les potentiels accélérateurs correspondants sont des « potentiels d'ionisation » qui mesurent en électron-volts les énergies de liaison des divers électrons de l'atome. L'étude des potentiels de résonance et des potentiels d'ionisation par choc poursuivie par de nombreux physiciens a apporté une nouvelle confirmation des idées de Bohr et en a montré la fécondité.

La théorie de l'atome de Bohr a donc été une révélation pour ceux qui étudiaient la structure de l'atome et les phénomènes qui en dérivent. Elle restait cependant visiblement insuffisante.

Du point de vue théorique, elle juxtaposait d'une manière très peu homogène des images et des calculs empruntés aux théories classiques, aux formules de la théorie des Quanta et aux conceptions si nouvelles d'états stationnaires et de transition quantique. Comme la théorie de Bohr déterminait la trajectoire des électrons à l'aide des équations de la Mécanique classique, quitte à restreindre, par l'application de règles quantiques, le nombre des trajectoires possibles, on pouvait croire qu'elle n'innovait pas d'une façon essentielle, du moins en ce qui concernait les principes. Mais à la réflexion, elle apparaissait comme

tout à fait révolutionnaire. Bohr lui-même insistait sur le fait que les états stationnaires sont en quelque sorte suspendus en dehors du temps et que les transitions quantiques ne devaient pas être conçues comme des mouvements de transition entre deux états de mouvement stables, mais devaient être regardées comme impossibles à décrire dans le cadre de l'espace et du temps. Bref, on sentait que la théorie de Bohr avait un caractère bâtarde et provisoire. C'est le développement des nouvelles Mécaniques qui a permis de mieux comprendre sa nature profonde.

En dehors de cette insuffisance, la théorie de Bohr présentait aussi une insuffisance pratique puisqu'elle ne permettait le calcul des états stationnaires que dans le cas, évidemment trop particulier, des trajectoires circulaires. Il en résultait que les calculs de Bohr ne fournissaient qu'une première approximation et ne pouvaient rendre compte de la complexité des structures fines spectrales. Même dans le cas de l'hydrogène, ils ne permettaient pas de rendre compte du fait que les raies de la série de Balmer sont en réalité des doublets. Il fallait donc chercher à s'affranchir de l'obligation de ne considérer que des mouvements à un seul degré de liberté. Nous allons voir maintenant comment, dès 1916, on y est parvenu.

4. **La théorie de Sommerfeld.** — Soit un système à n degrés de liberté dont la configuration est définie par n variables de Lagrange q_1, \dots, q_n . Si toutes les variables admettent la même période T , c'est-à-dire si à des intervalles de temps T elles reprennent toutes la même valeur, le système reprend périodiquement la même configuration : le mouvement est périodique. Si chaque variable q_i est périodique, mais chacune avec sa période propre T_i de variation, les périodes T_i étant incommensurables entre elles, le système est « quasi-périodique ». Les conditions de quanta n'ont de sens que pour les systèmes périodiques ou quasi-périodiques. Pour tous les systèmes de ce genre que l'on a eu à quantifier dans l'ancienne théorie des Quanta, il a toujours été possible de choisir les variables de façon qu'elles forment un système de « variables séparées », c'est-à-dire que chacun des moments de Lagrange p_i puisse s'exprimer en fonction de la seule coordonnée q_i correspondante. Les variables étant ainsi choisies, Wilson et Sommerfeld ont montré qu'on obtenait des conditions de quanta satisfaisantes en posant

$$\oint p_i dq_i = n_i h \quad (n_i \text{ entier; } i = 1, 2, \dots, n),$$

l'intégrale étant prise pour une période entière T_i de la variable q_i .

En appliquant ces nouvelles conditions de quanta à l'atome d'hydrogène de Bohr, on devait pouvoir quantifier les trajectoires elliptiques képlériennes qui, dans le champ coulombien du noyau en $\frac{1}{r^2}$, doivent pouvoir exister pour l'électron-planète et ne plus se borner à la considération exclusive des trajectoires circulaires. Il suffisait d'appliquer les conditions de quanta de Wilson-Sommerfeld aux deux coordonnées r et θ , rayon-vecteur et azimut, qui varient périodiquement le long de l'ellipse képlérienne et l'on pouvait espérer obtenir ainsi les structures fines observées dans le spectre de l'hydrogène.

Cet espoir fut déçu. Sommerfeld a fait le calcul en introduisant les deux nombres quantiques n_1 et n_2 correspondant aux deux variables $q_1 = \theta$ et $q_2 = r$, nombres quantiques qui sont des nombres entiers pouvant varier de 0 à ∞ , et il a trouvé pour les énergies des états stationnaires l'expression

$$E = -\frac{2\pi^2 m e^4}{(n_1 + n_2)^2 h^2} = -\frac{2\pi^2 m e^4}{n^2 h^2}$$

en posant $n = n_1 + n_2$. Il est alors évident que l'on obtient ainsi les mêmes niveaux d'énergie que dans la théorie primitive de Bohr et que, les nombres quantiques n_1 et n_2 n'intervenant que par leur somme, leur introduction n'a permis aucun progrès pour la prévision des structures fines.

C'est alors que Sommerfeld a eu l'idée très ingénieuse de refaire les calculs en utilisant, à la place des formules de la Dynamique classique celles de la Dynamique relativiste. Il s'est, en effet, aperçu que, pour les trajectoires internes de l'électron dans un atome de Bohr, les vitesses devenaient comparables à c , ce qui logiquement conduit à envisager l'emploi de la Dynamique relativiste.

La fonction de Lagrange sera ici

$$\mathcal{L} = -m_0 c^2 \sqrt{1 - \beta^2} - U = -m_0 c^2 \sqrt{1 - \beta^2} + \frac{e^2}{r}, \quad \text{avec} \quad \beta = \frac{1}{c} \sqrt{\dot{r}^2 + r^2 \dot{\theta}^2}$$

et les moments de Lagrange seront

$$p_r = \frac{\partial \mathcal{L}}{\partial \dot{r}} = \frac{m_0 \dot{r}}{\sqrt{1 - \beta^2}}, \quad p_\theta = \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = \frac{m_0 r^2 \dot{\theta}}{\sqrt{1 - \beta^2}},$$

ce qui conduit à écrire pour les conditions de Wilson-Sommerfeld.

$$(4) \quad \int_0^{2\pi} \frac{m_0 r^2 \dot{\theta}}{\sqrt{1 - \beta^2}} d\theta = n_1 h, \quad \oint \frac{m_0 \dot{r}}{\sqrt{1 - \beta^2}} dr = n_2 h.$$

L'étude détaillée de la trajectoire, que nous ne ferons pas ici, montre que l'électron décrit une ellipse avec rotation lente du périhélie (trajectoire en rosette); autrement dit, la trajectoire réelle est à chaque instant tangente à une ellipse osculatrice qui tourne lentement dans son plan. Le rayon vecteur oscille donc entre deux valeurs extrêmes r_1 et r_2 , mais le temps qu'il met à décrire le cycle $r_1 \rightarrow r_2 \rightarrow r_1$ (période de la variable r) est un peu plus long que le temps mis par l'azimut θ pour augmenter de 2π (période de la variable θ), de sorte que le mouvement n'est plus strictement périodique, mais seulement quasi-périodique.

Ceci posé, d'après la dynamique relativiste, l'expression hamiltonienne de l'énergie est (1)

$$W = c \sqrt{m_0^2 c^2 + p_r^2} + \frac{p_\theta^2}{r^2} - \frac{e^2}{r}.$$

Le moment de Lagrange p_θ est le moment de la quantité de mouvement de l'électron par rapport au noyau : il est constant en vertu du théorème des aires qui est encore valable en Dynamique relativiste. La première condition (4) donne donc

$$p_\theta = \frac{m_0 r^2 \dot{\theta}}{\sqrt{1-\beta^2}} = n_1 \frac{h}{2\pi}.$$

Portant cette valeur de p_θ dans l'expression de W , on trouve

$$p_r = \pm \sqrt{A + 2 \frac{B}{r} + \frac{C}{r^2}},$$

(1) Il suffit, pour obtenir ce résultat, d'effectuer un calcul analogue à celui du chapitre III, paragraphe 6, calcul qui consiste à éliminer $\dot{\theta}$ entre les expressions

$$W = \frac{m_0 c^2}{\sqrt{1-\beta^2}} - \frac{e^2}{r}, \quad p_\theta = \frac{m_0 r^2 \dot{\theta}}{\sqrt{1-\beta^2}}, \quad p_r = \frac{m_0 \dot{r}}{\sqrt{1-\beta^2}}.$$

En calculant $\frac{p_\theta^2}{r^2} + p_r^2$, on trouve

$$\frac{p_\theta^2}{r^2} + p_r^2 = \frac{m_0^2 (\dot{r}^2 + r^2 \dot{\theta}^2)}{1-\beta^2} = \frac{m_0^2 \beta^2 c^2}{1-\beta^2},$$

d'où

$$\frac{1}{\sqrt{1-\beta^2}} = \frac{\sqrt{m_0^2 c^2 + p_r^2 + \frac{p_\theta^2}{r^2}}}{m_0 c},$$

ce qui conduit à l'expression indiquée pour W .

avec les définitions

$$\begin{aligned} A &= \frac{E^2}{c^2} + 2m_0 E = m_0^2 c^2 \left[\left(1 + \frac{E}{m_0 c^2} \right)^2 - 1 \right], \\ B &= \frac{E e^2}{c^2} + m_0 e^2, \\ C &= \frac{e^4}{c^2} - p_0^2 = \frac{e^4}{c^2} - \frac{n_1^2 h^2}{4\pi^2}, \end{aligned}$$

$E = W - m_0 c^2$ étant la somme de l'énergie potentielle et de l'énergie cinétique.

Sommerfeld introduit alors dans son calcul ce qu'il a nommé la « constante de structure fine » qui est un nombre sans dimensions

$$\alpha = \frac{2\pi e^2}{hc} \quad \left(\alpha = 7,29 \cdot 10^{-2} \simeq \frac{1}{137} \right).$$

On peut donc écrire

$$C = -\frac{n_1^2 h^2}{4\pi^2} \left(1 - \frac{\alpha^2}{n_1^2} \right).$$

Or, l'application du théorème des résidus permet d'établir que l'on a

$$\oint \sqrt{A + \frac{2B}{r} + \frac{C}{r^2}} dr = -2\pi i \left(\sqrt{C} - \frac{B}{\sqrt{A}} \right).$$

En égalant le second membre à $n_2 h$ conformément à la seconde condition (4) et en remplaçant A, B, C par leurs valeurs, Sommerfeld obtient, après un calcul simple, la formule rigoureuse

$$1 + \frac{E}{m_0 c^2} = \left[1 + \frac{\alpha^2}{\left[n_2 + \sqrt{n_1^2 - \alpha^2} \right]^2} \right]^{-\frac{1}{2}}$$

qui donne l'énergie de l'état quantifié de nombres quantiques n_1 et n_2 .

La quantité α^2 étant très petite, une première approximation consiste à la négliger devant l'unité. On retombe alors sur la formule de la théorie non relativiste et l'on ne trouve pas de structure fine. Une approximation meilleure consistera à garder les termes en α^2 et, en posant $n_1 + n_2 = n$, on trouvera alors

$$E = -\frac{2\pi^2 m_0 e^4}{n^2 h^2} \left[1 + \frac{\alpha^2}{n^2} \left(\frac{1}{4} + \frac{n_2}{n_1} \right) \right] \quad (n_1 \text{ et } n_2 = 0, 1, 2, \dots).$$

Le dernier terme du crochet explique l'existence d'une structure fine parce qu'il dépend séparément de n_1 et de n_2 et non plus seulement de leur somme.

On désigne souvent le nombre $n_1 + n_2 = n$ sous le nom de « nombre quantique total » et l'on nomme « nombre quantique azimutal » le nombre $k = n_1$. La donnée des nombres quantiques n et k suffit à déterminer un niveau d'énergie puisque le « nombre quantique radial » n_2 est égal à $n - k$. En introduisant la constante de Rydberg, on obtient pour l'expression de l'énergie du niveau n, k

$$E_{n,k} = - \frac{R h}{n^2} \left[1 + \frac{\alpha^2}{n^2} \left(\frac{n}{k} - \frac{3}{4} \right) \right].$$

Dans l'ancienne théorie des Quanta, on admettait que le nombre azimutal k ne pouvait jamais prendre la valeur zéro parce qu'elle aurait correspondu à une trajectoire rectiligne passant par le noyau. D'après la formule obtenue, chaque orbite stationnaire possède une énergie $E_{n,k}$ qui ne dépend plus seulement de n , mais aussi de k ; seulement, comme α^2 est très petit devant l'unité, les divers termes spectraux correspondant à une même valeur de n sont très voisins et l'on obtient bien ainsi une structure fine des raies prévues par la théorie primitive de Bohr.

Chaque terme spectral de Bohr correspondant à une valeur donnée de n se décompose en n termes voisins, car pour n fixé il y a n valeurs possibles de k , savoir 1, 2, ..., n . Il convient de remarquer que l'écartement des termes spectraux voisins est d'autant plus petit que n est plus élevé à cause de la présence de n^2 au dénominateur du terme en α^2 .

Considérons la série de Balmer. En première approximation ses raies sont données par

$$\nu = R \left[\frac{1}{2^2} - \frac{1}{n^2} \right] \quad (n = 3, 4, \dots).$$

En seconde approximation, on doit d'après Sommerfeld remplacer le terme spectral $\frac{R}{2^2}$ par

$$\frac{R}{2^2} \left[1 + \frac{\alpha^2}{2^2} \left(\frac{2}{k} - \frac{3}{4} \right) \right]$$

et le terme spectral $\frac{R}{n^2}$ par

$$\frac{R}{n^2} \left[1 + \frac{\alpha^2}{n^2} \left(\frac{n}{k} - \frac{3}{4} \right) \right].$$

Il y a dans la série de Balmer une structure fine à écartement constant provenant du dédoublement $k = 1, 2$ du premier terme spectral fixe et une autre structure fine à écartement décroissant quand on s'élève dans la série, due, celle-ci, à la complexité du second terme spectral variable.

Cette deuxième structure fine est pratiquement inobservable parce qu'elle est trop fine. La première correspond à une décomposition de chaque raie prévue par la formule de Balmer en un doublet à écartement constant pour toute la série est égal à

$$\Delta\nu_{II} = \frac{R}{2^2} \left[1 + \frac{\alpha^2}{2^2} \left(\frac{2}{1} - \frac{3}{4} \right) \right] - \frac{R}{2^2} \left[1 + \frac{\alpha^2}{2^2} \left(\frac{2}{2} - \frac{3}{4} \right) \right] = \frac{R\alpha^2}{16}.$$

Numériquement, en remplaçant les fréquences par les nombres d'onde, qui sont les fréquences réelles divisées par c , on obtient

$$\Delta\nu_{II} = 0,365\text{cm}^{-1}.$$

Cette valeur est en assez bon accord avec les données de l'expérience. On obtient donc une interprétation quantitative de l'existence des doublets dans la série de Balmer par l'introduction dans la théorie de Bohr de la Dynamique relativiste.

Si l'on reprend les calculs de Sommerfeld en supposant que l'on ait affaire à un atome de nombre atomique Z ionisé ($Z-1$) fois et ne gardant, par suite, qu'un électron périphérique, on trouve pour l'état stationnaire caractérisé par les nombres quantiques n et k

$$E_{nk} = - \frac{RZ^2 h}{n^2} \left[1 + \frac{\alpha^2 Z^2}{n^2} \left(\frac{n}{k} - \frac{3}{4} \right) \right].$$

L'écartement des doublets va donc se trouver multiplié par Z^4 . Pour l'hélium ionisé ($Z=2$), l'écartement des doublets dans la série qui correspond à la série de Balmer va donc être 16 fois plus grand que celui de la série de Balmer elle-même. On conçoit donc que l'étude des doublets dans le spectre de He^+ ait pu fournir une vérification précise de la théorie de Sommerfeld : cette vérification faite par Paschen a été très satisfaisante.

§. Succès et insuffisance de la théorie de Sommerfeld. — La théorie de Sommerfeld avait donc donné pour les spectres de H et de He^+ des résultats qui semblaient excellents. Par des calculs moins rigoureux, Sommerfeld a pu aussi en tirer une justification des formules empiriques proposées par Rydberg et par Ritz pour représenter les raies spectrales des éléments non hydrogénéoïdes.

Lorsque l'on cherche à étendre la théorie de Bohr à des atomes contenant plus d'un électron-planète, on se heurte à de grosses difficultés. Le problème dynamique devient compliqué et l'application des règles de Quantà devient incertaine. Cependant, l'analogie générale des spectres de tous les éléments, l'intervention dans les formules empiriques du

type Rydberg de la constante de Rydberg conduisait à penser que le schéma planétaire si utile dans le cas de l'atome d'hydrogène devait pouvoir aussi être utilisé de quelque manière dans le cas des autres éléments. Pour le faire, on a commencé par admettre une hypothèse assez grossière : dans l'atome de nombre atomique Z , on a considéré $Z - 1$ des électrons-planètes comme tournant au voisinage du noyau et constituant autour de lui une sorte de « carcasse électronique », tandis que le $Z^{\text{ème}}$ électron-planète, dit « électron optique », aurait son orbite à l'extérieur de la carcasse électronique. Les transitions de l'électron optique d'un état stationnaire à un autre détermineraient l'émission des raies du spectre optique de l'élément. Grâce à l'hypothèse de la carcasse électronique, on peut admettre que l'action du noyau de charge Ze sur l'électron optique est à une unité près compensée par celle de la carcasse électronique, de sorte que l'électron optique se meut à peu près comme s'il était soumis au champ d'une charge $+e$. On est ainsi ramené au problème de l'atome d'hydrogène avec un seul nombre quantique n . En deuxième approximation, on cherchera avec Sommerfeld à tenir compte de ce que la charge du noyau et celle de la carcasse ne se compensent pas exactement à une unité près : la trajectoire de l'électron optique qui peut même pénétrer dans la carcasse électronique ne se ferme plus et il faut introduire un deuxième nombre quantique k . On trouvera tous ces développements théoriques exposés en détail dans les livres consacrés à l'ancienne théorie des Quanta, notamment dans ceux de Sommerfeld et de Léon Brillouin [1], [2].

D'autres calculs, également assez approximatifs, ont permis à Sommerfeld d'étendre sa formule de structure fine au cas des spectres de rayons X. Dès le premier travail de Bohr, on avait cherché à rendre compte des spectres X en considérant l'atome de nombre atomique Z comme contenant Z électrons tournant sur des cercles coplanaires (cercle K, cercle L, etc.). Pour tenir compte grossièrement de la répulsion mutuelle des électrons, on avait supposé qu'elle se traduisait par un simple « effet d'écran » ayant pour conséquence une diminution apparente de la charge du noyau. Ainsi l'électron K serait soumis à une force $(Z - k) \frac{e^2}{r^2}$, l'électron L à une force $(Z - l) \frac{e^2}{r^2}$, etc., k, l, \dots étant appelés « nombres d'écran ». Les calculs de Bohr pour l'atome d'hydrogène peuvent alors s'appliquer ici sans difficulté et l'on trouve les termes spectraux

$$- \frac{R h (Z - k)^2}{1^2}, \quad - \frac{R h (Z - l)^2}{2^2}, \quad \dots,$$

ce qui permet de retrouver approximativement la loi de Moseley suivant laquelle les termes spectraux des rayons X varient en gros comme Z^2 . Mais cette première approximation est manifestement insuffisante parce qu'elle conduit à prévoir seulement *un* niveau L, *un* niveau M, etc., alors qu'on sait que tous les niveaux (sauf le niveau K) sont multiples.

En introduisant sa théorie de la structure fine et en supposant hardiment qu'on peut isoler par la pensée les diverses trajectoires des électrons de la carcasse électronique, Sommerfeld est parvenu pour le terme spectral X caractérisé par les nombres quantiques n et k à l'expression

$$(5) \quad \frac{E_{nk}}{h} = - \frac{R(Z - z_{nk})^2}{n^2} \left[1 + z^2 \frac{(Z - z_{nk})^2}{n^2} \left(\frac{n}{k} - \frac{3}{4} \right) \right],$$

où z_{nk} est le nombre d'écran relatif à la trajectoire caractérisée par n et k . Comme pour n donné, il y a n valeurs possibles pour k , la théorie de Sommerfeld prévoit *un* niveau K, *deux* niveaux L, *trois* niveaux M, etc. En réalité, la structure fine des raies X est plus riche que cela et il y a *trois* niveaux L, *cinq* niveaux M, etc. La théorie de Sommerfeld plus complète que la théorie primitive est donc encore trop étroite.

Néanmoins, malgré cette insuffisance évidente, elle a paru remporter un très grand succès pour l'explication quantitative des doublets dits « réguliers ». Dans les séries L, par exemple, ces doublets proviennent de la combinaison des termes spectraux L_{II} et L_{III} avec un même terme M, N, etc. Les deux raies du doublet ont donc des fréquences de la forme $\nu_{LII} - \nu_i$ et $\nu_{LIII} - \nu_i$ et l'écartement du doublet est $\nu_{LII} - \nu_{LIII}$. Dans sa théorie, Sommerfeld laissait de côté comme provisoirement inexpliqué le niveau L_I et il attribuait aux niveaux L_{II} et L_{III} les nombres quantiques $n = 2, k = 1$ et $n = 2, k = 2$ respectivement. La formule (5) lui donnait alors

$$\delta\nu_L = \nu_{LII} - \nu_{LIII} = R z^2 \frac{(Z - z_L)^2}{2^4} \left(\frac{2}{1} - \frac{2}{2} \right) = \frac{\pi^4 m_0 e^8}{2 c^2 h^5} (Z - z_L)^2.$$

L'écartement des doublets L doit donc varier comme $(Z - z_L)^4$ ou, comme pour les atomes par trop légers z_L est beaucoup plus petit que Z , comme Z^4 : c'est bien la loi constatée expérimentalement. Si l'on pose $Z = 1, z_L = 0$, on retombe sur l'écartement $\delta\nu_{II}$ des doublets de la série de Balmer. On a donc

$$\delta\nu_L = \delta\nu_{II}(Z - z_L)^4,$$

formule qui est bien vérifiée numériquement à condition d'adopter pour z_L la valeur raisonnable 3,5. Les doublets des séries M et N sont aussi numériquement bien prévus par l'application des formules de

Sommerfeld. Cette prévision exacte des doublets réguliers a paru tout d'abord un très grand succès de la théorie de Sommerfeld.

D'autres succès de cette théorie furent la prévision exacte de l'effet Zeeman normal et de l'effet Stark. Nous renvoyons sur ces points aux Ouvrages cités.

Parlons maintenant des insuffisances de la théorie de Sommerfeld. Au point de vue théorique général, elle souffre des mêmes insuffisances que la théorie dont elle n'est que le prolongement. La réalité physique des trajectoires en rosette calculées à l'aide d'un mélange assez bâtarde de règles classiques et de conceptions quantiques est douteuse. D'autres insuffisances apparaissent quand on examine les résultats mêmes de la théorie. Le grand mérite de la théorie de Sommerfeld, c'est d'avoir rendu compte d'une partie des structures fines que la tentative initiale de Bohr ignorait complètement. Mais les structures fines réelles sont encore plus complexes que ne le prévoient les raisonnements de Sommerfeld : la chose est particulièrement visible pour les spectres X où les niveaux sont beaucoup trop nombreux pour pouvoir être numérotés à l'aide des deux seuls nombres quantiques n et k . Pour établir une classification complète des niveaux dans le domaine optique et dans le domaine X, Sommerfeld lui-même a été obligé d'introduire empiriquement un troisième nombre quantique j dont la signification théorique restait très obscure. On n'a pu comprendre réellement l'origine de la multiplicité réelle des niveaux et le sens véritable du nombre quantique j que le jour où l'on a attribué à l'électron une propriété nouvelle jusqu'alors méconnue : *le spin*.

Une autre difficulté avait surgi entre temps. En introduisant la dynamique relativiste, Sommerfeld était parvenu à prévoir l'existence et la variation en fonction de Z des doublets réguliers, ce qui paraissait un très beau succès. Mais les formules de Sommerfeld prévoyaient l'existence des différences de fréquence, origines de ces doublets, entre les niveaux de même n dont le k diffère d'une unité, comme le montre la formule (5). Or quand on établit la classification rationnelle des niveaux en introduisant le nombre j , on s'aperçoit que les niveaux origines des doublets réguliers, (tels, par exemple, que les niveaux L_{II} et L_{III}) ont mêmes nombres n et k et que c'est leur nombre j qui diffère d'une unité. Le succès des formules de Sommerfeld était donc remis en cause, les doublets réguliers ne se trouvant plus être placés là où la théorie les prévoyait. La théorie de l'électron à spin de Dirac a montré qu'en introduisant à la fois le spin et la Relativité dans le cadre de la Mécanique ondulatoire, on obtient une formule analogue à celle de Sommerfeld,

mais où le nombre j remplace le nombre k , de sorte que tout rentre dans l'ordre (*voir* chap. XIV, § 6).

Nous étudierons ces dernières questions d'une façon plus approfondie dans le chapitre XIV consacré au spin et à la théorie de Dirac.

BIBLIOGRAPHIE.

- [1] A. SOMMERFELD, *La constitution de l'atome et les raies spectrales*. Blanchard, Paris, 1923.
- [2] LÉON BILLOUIN, *L'atome de Bohr*, Presses universitaires, Conférences-Rapports, Paris, 1931.
- [3] EUGÈNE BLOCH, *L'ancienne et la nouvelle théorie des Quanta*. Hermann, Paris, 1930.
- [4] LOUIS DE BROGLIE, *L'électron magnétique*, Hermann. Paris, 1934.

CHAPITRE VIII.

LE PRINCIPE DE CORRESPONDANCE.

1. **Objet du principe de correspondance.** — L'introduction des Quanta dans la théorie de l'atome avait considérablement amélioré nos interprétations des phénomènes atomiques en ce qui concerne les échanges d'énergie et le calcul des fréquences spectrales. Elle avait même permis de retrouver certaines formules classiques concernant les fréquences où la constante h se trouve éliminée, telles que les formules de l'effet Zeeman normal. Mais les physiciens se sont trouvés dans un grand embarras quand il leur a fallu transposer dans le cadre des théories quantiques les questions d'intensité et de polarisation et les théories comme celles de la diffusion et de la dispersion, en un mot l'interprétation des phénomènes où les propriétés ondulatoires de la lumière jouent un rôle essentiel. Tous ces phénomènes qui, dans leur ensemble, s'interprétaient d'une façon très satisfaisante avec les idées classiques, devenaient très difficiles à rattacher aux idées quantiques. C'est pour tenter de faire cette pénible jonction que M. Bohr a développé, quelques années après sa théorie de l'atome, son principe de correspondance (1916). Ce principe était en réalité une sorte de fil directeur ayant pour rôle de guider les physiciens dans le difficile travail qui consiste à raccorder les anciennes théories ondulatoires avec les conceptions quantiques.

L'idée qui a dirigé Bohr dans sa découverte du principe de correspondance est que les théories classiques, donnant des résultats satisfaisants pour les phénomènes à notre échelle, doivent avoir le caractère d'approximations statistiques valables pour les phénomènes qui mettent en jeu un grand nombre de quanta. Il en est bien ainsi, on peut le constater, pour les fréquences des raies spectrales, car, si l'on considère des états quantiques correspondant à un très grand nombre de quanta, les fréquences d'émission par application de la loi des fréquences de Bohr coïncident avec celles que la théorie classique prévoit en décomposant le mouvement

de l'électron en composantes harmoniques. Ce résultat constitue ce que l'on peut appeler « le théorème de correspondance pour les fréquences ». Il était donc assez naturel de généraliser cette coïncidence en supposant que, pour les transitions entre états à grand nombre de quanta, les prévisions de la théorie classique relatives aux intensités et aux polarisations étaient aussi sensiblement exactes. Hardiment, Bohr a été plus loin : il a admis que les prévisions classiques pouvaient encore être utilisées pour évaluer, au moins grossièrement, les intensités et les polarisations même pour les raies émises par transition entre états à petits nombres de quanta.

Il importe de bien comprendre dans quel sens on peut dire que les intensités calculées par la théorie classique sont exactes dans le cas des grands nombres quantiques. Au point de vue quantique, chaque acte d'émission donne naissance à un seul quantum $h\nu$ et cela n'a pas de sens de parler de l'intensité du rayonnement ainsi émis. Mais, s'il y a un très grand nombre N d'atomes semblables présents dans le même état initial, il y aura une certaine fraction P de ces atomes qui, pendant l'unité de temps, subira la transition avec émission du quantum $h\nu$, de sorte que l'énergie émise par unité de temps par l'ensemble des atomes sera $PNh\nu$ et c'est cette énergie que le principe de correspondance permet d'évaluer par l'emploi des formules classiques. En fin de compte, les formules classiques serviront donc à déterminer « les probabilités de transition P ». C'est donc seulement au point de vue statistique que l'électromagnétisme classique conserve une certaine valeur et ce serait une erreur de dire que, d'après le principe de correspondance, l'atome émet suivant les lois de l'électromagnétisme. Pour parler correctement, il faut dire : « Les ensembles d'atomes émettent statistiquement, dans le domaine des grands nombres quantiques, suivant les lois de l'électromagnétisme ».

Avant d'aborder l'étude détaillée du principe de correspondance, il est nécessaire que nous disions quelques mots des variables angulaires et de leur intervention dans l'ancienne théorie des Quanta.

2. Les variables angulaires. — Rappelons comment Planck avait introduit la notion de quantum. Il envisageait des oscillateurs linéaires harmoniques dont la fréquence d'oscillation ν est indépendante de l'amplitude et, adoptant un énoncé de « quanta d'énergie », il avait admis que les seuls mouvements stables de l'oscillateur sont ceux dont l'énergie est de la forme

$$(1) \quad E = nh\nu \quad (n \text{ entier}).$$

Cet énoncé a l'inconvénient de ne pouvoir s'appliquer qu'à un oscillateur quasi-élastique dont la fréquence est indépendante de l'amplitude. Planck a alors cherché un énoncé plus général qui soit applicable à un système périodique quelconque à un degré de liberté et il a adopté le suivant :

$$(2) \quad \oint p dq = nh,$$

où l'intégrale est étendue à une période de la variable q . Cette nouvelle règle de quantification a l'avantage de bien faire ressortir que la constante h a le caractère d'une unité d'action, car elle exprime que l'intégrale d'Action Maupertuisienne étendue à un cycle complet du mouvement périodique est égale à un multiple entier de la constante h . Appliqué au cas particulier de l'oscillateur linéaire harmonique, la formule (2) permet d'ailleurs de retrouver aisément l'énoncé (1).

La formule de quantification (2), après avoir été utilisée par Bohr dans sa théorie primitive des orbites circulaires, a été généralisée par Wilson et Sommerfeld pour les mouvements définis par plus d'une variable. Les mouvements de ce genre que l'on a à quantifier sont toujours des mouvements quasi-périodiques où chacune des variables a son cycle propre de variation. Il y a alors toujours au moins une manière de choisir les variables q_i de façon que chacun des moments p_i puisse s'exprimer à l'aide de la seule variable q_i qui lui est conjuguée (système de variables séparées). Chacune des intégrales $\oint p_i dq_i$ étendue à un cycle entier de variations de q_i a alors un sens bien défini et, d'après Wilson et Sommerfeld, on doit quantifier en posant

$$(3) \quad \oint p_i dq_i = n_i h \quad (n_i \text{ entier}).$$

On obtient ainsi autant de conditions de quantification qu'il y a de variables q_i .

Cette méthode de quantification donne des résultats satisfaisants et univoques quand il n'y a pas de « dégénérescence », c'est-à-dire quand le nombre des dimensions de la trajectoire n'est pas inférieur au nombre des variables q_i . La forme des orbites quantifiées est alors entièrement déterminée par le choix unique des variables séparées.

Il n'en est pas de même dans le cas où il y a dégénérescence, parce qu'alors il se trouve qu'il y a plusieurs choix possibles de variables séparées. Selon le système de variables séparées que l'on choisit, la forme des orbites quantifiées diffère, ce qui n'est pas satisfaisant. On

vérifie cependant, que les énergies des orbites quantifiées restent toujours les mêmes, comme si seules les énergies, et non les formes, des orbites quantifiées avaient un sens physique. Dans ces cas de dégénérescence, la méthode des variables séparées de Wilson et Sommerfeld semble donc donner un nombre trop grand de conditions quantiques et il paraissait désirable de trouver une autre méthode ne donnant pas des conditions trop strictes dans les cas de dégénérescence. Ce but a été atteint par l'emploi des « variables angulaires » qui ont joué, par ailleurs, un rôle essentiel dans le développement mathématique du principe de correspondance.

Considérons un système à f degrés de liberté décrit par f coordonnées q_k et f moments conjugués de Lagrange p_k . Entre ces variables canoniques, nous avons les équations de Hamilton

$$\frac{dq_k}{dt} = + \frac{\partial H}{\partial p_k}, \quad \frac{dp_k}{dt} = - \frac{\partial H}{\partial q_k} \quad (k = 1, 2, \dots, f),$$

où H est la fonction hamiltonienne qui exprime l'énergie du système en fonction des q_k et des p_k .

On démontre en Mécanique analytique ⁽¹⁾ le théorème suivant très souvent utilisé en Mécanique céleste :

Si l'on parvient à définir de nouvelles variables

$$Q_i = f_i(q_k, p_k) \quad \text{et} \quad P_i = g_i(q_k, p_k)$$

à l'aide de l'ensemble des anciennes variables canoniques q_k et p_k et si l'expression

$$\sum_i P_i dQ_i - \sum_k p_k dq_k$$

est la différentielle exacte d'une fonction des q_k et des p_k , les Q_i et les P_i sont encore des variables canoniques, c'est-à-dire qu'elles satisfont à des équations de la forme

$$\frac{dQ_i}{dt} = \frac{\partial \mathcal{E}}{\partial P_i} \quad \frac{dP_i}{dt} = - \frac{\partial \mathcal{E}}{\partial Q_i}.$$

De plus, si la fonction hamiltonienne primitive $H(q_k p_k)$ ne dépend pas explicitement du temps (ce qui est le cas pour les systèmes conser-

(1) On trouvera la démonstration de cette proposition ainsi que de toutes celles qui suivent à la fin du tome II du livre de Sommerfeld [1].

catifs et, en particulier, pour les systèmes quantifiables), la fonction \mathcal{H} des Q_i et des P_i s'obtient simplement en remplaçant dans $H(q_k, p_k)$ les q_k et les p_k par leurs expressions en fonction des Q_i et des P_i .

Le passage des variables q_k, p_k aux variables Q_i, P_i qui conserve la forme des équations de Hamilton est nommé « changement de variables canoniques ».

Considérons alors un système quasi-périodique et soient q_1, \dots, q_f un jeu de variables séparées pour ce système. On peut démontrer qu'il est possible de trouver un changement de variables canoniques tel que les nouveaux moments P_i soient précisément égaux aux intégrales cycliques de Wilson-Sommerfeld (ou à des combinaisons linéaires de ces intégrales). Ces P_i sont alors des constantes, les $\frac{dP_i}{dt}$ sont nuls et, d'après les équations de Hamilton, les $\frac{dH}{dQ_i}$ sont nuls aussi. Donc la fonction hamiltonienne, exprimée à l'aide des P_i que nous nommerons J_i et des coordonnées Q_i que nous nommerons dorénavant ω_i , se trouve ne pas dépendre des ω_i . Comme, d'autre part, on a

$$\frac{dQ_i}{dt} = \frac{d\omega_i}{dt} = \frac{\partial \mathcal{H}}{\partial J_i}$$

et que les $\frac{\partial H}{\partial J_i}$ sont des constantes, les ω_i sont des fonctions linéaires du temps de la forme

$$\omega_i = \nu_i t + \delta_i,$$

où les constantes ν_i sont définies par

$$\nu_i = \frac{\partial H}{\partial J_i},$$

relation fondamentale pour le principe de correspondance.

On démontre encore que les variables q_k sont des fonctions périodiques des variables ω_i .

Bref, il est possible de trouver un changement de variables canoniques faisant passer des variables séparées primitives q_k et p_k à de nouvelles variables canoniques ω_i et J_i telles que :

- 1° la fonction hamiltonienne exprimée avec ces variables ne dépend que des J_i ;
- 2° les J_i sont des constantes;
- 3° les coordonnées ω_i sont des fonctions linéaires du temps $\omega_i = \nu_i t + \delta_i$;
- 4° les q_k sont des fonctions périodiques des ω_i .

Les ω_i sont appelées des « *variables angulaires* », parce qu'elles croissent proportionnellement au temps comme un angle dans un mouvement de rotation uniforme. Les quantités J_i sont appelées les « *variables d'action* », nom dont l'origine est évidente.

Il faut remarquer que les ω_i et les J_i ne sont pas entièrement déterminées par les conditions 1^o, 2^o, 3^o et 4^o ci-dessus, car si l'on a trouvé un jeu de ω_i et de J_i qui satisfasse à ces conditions, on en trouvera un autre en posant

$$\omega'_i = \alpha_i \omega_i, \quad J'_j = \frac{J_j}{\alpha_j},$$

les α_i étant des constantes arbitraires. Pour achever de déterminer les ω_i , on pose par convention que la période de tous les q_k par rapport aux ω_i doit être égale à l'unité.

Ceci posé, on prend comme conditions de quantification dans la méthode des variables angulaires les formules

$$(4) \quad J_i = n_i h \quad (n_i \text{ entier}).$$

Quand il n'y a pas de dégénérescence, le nombre des variables angulaires est égal au nombre des variables séparées et chaque J_i est égal à l'une des intégrales cycliques de Wilson-Sommerfeld. Il y a donc alors coïncidence entre les conditions (3) et (4), c'est-à-dire entre les deux méthodes des variables séparées et des variables angulaires.

Mais il n'en est plus de même dans les cas de dégénérescence. On trouve alors un nombre de variables angulaires inférieur à celui des variables séparées et certaines au moins des variables J_i sont des combinaisons linéaires des intégrales cycliques de Wilson-Sommerfeld. Le nombre des conditions (4) est donc inférieur à celui des conditions (3), c'est-à-dire que la méthode des variables angulaires détermine moins strictement que la méthode des variables séparées la forme des trajectoires et les difficultés soulevées par l'application des conditions (3) de Wilson-Sommerfeld n'existent plus avec les conditions (4), ce qui donne une nette supériorité à la méthode des variables angulaires.

Puisque tous les q_k (et, par suite, la configuration du système) sont des fonctions périodiques de période 1 des variables angulaires ω_i , une grandeur quelconque qui dépend de la configuration du système (par exemple l'une des composantes de son moment électrique) pourra se développer en série multiple de Fourier sous la forme

$$\sum_{\tau_1, \dots, \tau_n} c_{\tau_1, \dots, \tau_n}(J_1, \dots, J_n) e^{2\pi i(\tau_1 \omega_1 + \dots + \tau_n \omega_n)},$$

les τ_i étant des nombres entiers positifs ou négatifs et n étant le nombre des variables angulaires ($n \leq f$). Si la grandeur doit être réelle d'après sa signification physique, on aura

$$c_{-\tau_1, \dots, -\tau_n} = c_{\tau_1, \dots, \tau_n}^*$$

Les ω_i étant des fonctions linéaires du temps, nous écrirons le développement précédent sous la forme

$$\sum_{\tau_1, \dots, \tau_n} d_{\tau_1, \dots, \tau_n}(J_1, \dots, J_n) e^{2\pi i(\tau\nu)t},$$

ou, en abrégé

$$\sum_{\tau} d_{\tau}(J) e^{2\pi i(\tau\nu)t},$$

avec la définition

$$(\tau\nu) = \tau_1 \nu_1 + \dots + \tau_n \nu_n.$$

3. Exemples de quantification par la méthode des variables angulaires. — Nous allons donner deux exemples de quantification par la méthode des variables angulaires.

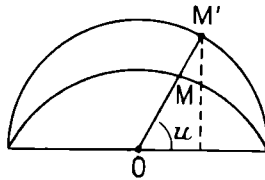


Fig. 21.

D'abord l'atome d'hydrogène. Dans cet atome, l'unique électron décrit une orbite képlérienne sous l'action de la force de Coulomb : dans la théorie des ellipses képlériennes, classique en Mécanique rationnelle, on montre qu'il existe un angle, l'anomalie moyenne, qui croît proportionnellement au temps. Cet angle θ est relié à l'anomalie excentrique u par la relation bien connue

$$\theta = u - e \sin u,$$

où e est l'excentricité de l'ellipse.

En posant $\omega = \frac{\theta}{2\pi}$, on obtient une variable angulaire augmentant proportionnellement au temps et telle que le système reprenne la même configuration quand ω augmente d'une unité. Il y a ici dégénérescence et nous avons une seule variable angulaire au lieu des deux variables

séparées r et θ . La variable d'action J conjuguée de ω est égale à la somme des intégrales de Wilson-Sommerfeld

$$\oint p_\theta d\theta + \oint p_r dr.$$

Quand on cherche à évaluer l'énergie du mouvement de l'électron en fonction de ω et de J , on trouve

$$H(J) = - \frac{2\pi^2 m e^4}{J^2}$$

qui est bien constante et indépendante de ω . Il suffit alors de poser $J = n\hbar$ pour retrouver la formule des états stationnaires de Bohr sous la forme

$$E_n = - \frac{2\pi^2 m e^4}{n^2 \hbar^2}.$$

Un deuxième exemple nous sera fourni par la théorie de l'effet Zeeman pour un atome d'hydrogène. L'atome est placé dans un champ magnétique uniforme et constant \mathbf{H} . Un théorème célèbre dû à Sir Joseph Larmor nous apprend que pour les valeurs des champs usuellement réalisables, le mouvement de l'électron peut s'obtenir de la façon simple suivante : « Imaginons un système de référence qui tourne d'un mouvement uniforme autour de la direction du champ \mathbf{H} avec la vitesse angulaire (dite de Larmor)

$$\omega = \frac{eH}{2m_0c},$$

où e et m_0 sont la charge et la masse de l'électron, alors le mouvement de l'électron par rapport à ce système de référence sera le même que si le champ \mathbf{H} n'existait pas et si le système ne tournait pas ». Autrement dit, pour les valeurs pas trop fortes du champ, il y a compensation entre l'effet de la rotation et l'action du champ.

La trajectoire de l'électron sera donc une certaine ellipse képlérienne tournant d'un mouvement uniforme de vitesse angulaire ω autour de \mathbf{H} .

Dans le système tournant, nous pouvons donc toujours définir une anomalie moyenne que nous nommerons maintenant θ_1 et qui aura toujours pour moment conjugué la variable d'action

$$J_1 = \oint p_{\theta_1} d\theta_1 + \oint p_r dr,$$

où θ_1 est l'azimut dans le système tournant.

Mais maintenant pour connaître la position de l'électron, il faut

encore nous donner un angle θ_2 qui fixe la position du système tournant par rapport au système de référence de l'observateur et, comme cet angle augmente avec la vitesse angulaire ω , on obtient une seconde variable angulaire en posant

$$\omega_2 = \frac{\theta_2}{2\pi} = \frac{\omega}{2\pi} t + \delta_2 = \nu_2 t + \delta_2, \quad \text{avec} \quad \nu_2 = \frac{eH}{4\pi m_0 c},$$

ν_2 est la fréquence de Larmor. Le moment conjugué J_2 est le produit par 2π du moment de la quantité de mouvement de l'électron autour de la direction \mathbf{H} qui est constant.

Pour obtenir la fonction hamiltonienne $H(J_1, J_2)$, nous remarquerons qu'elle contient d'abord le terme $-\frac{2\pi m_0 e^4}{J_1^2}$ correspondant au mouvement relatif dans le système tournant, plus un terme dû à la rotation du système de Larmor qui dépendra de J_2 . Nous calculerons ce terme en remarquant que sa dérivée par rapport à J_2 doit nous donner ν_2 , de sorte que finalement

$$H(J_1, J_2) = -\frac{2\pi^2 m_0 e^4}{J_1^2} + \frac{eH}{4\pi m_0 c} J_2,$$

En posant $J_1 = nh$ et $J_2 = mh$, nous obtiendrons pour les énergies quantifiées

$$E_{n,m} = -\frac{2\pi m_0 e^4}{n^2 h^2} + m \frac{ehH}{4\pi m_0 c},$$

d'où, pour les fréquences émises,

$$\begin{aligned} \nu_{nm, n'm'} &= \frac{E_{nm} - E_{n'm'}}{h} = \frac{2\pi^2 m_0 e^4}{h^3} \left(\frac{1}{n'^2} - \frac{1}{n^2} \right) + (m - m') \frac{eH}{4\pi m_0 c} \\ &= \nu_{nn'}^{(0)} + \delta m \frac{eH}{4\pi m_0 c}, \end{aligned}$$

où $\nu_{nn'}^{(0)}$ est la fréquence émise en l'absence de champ et δm la variation pendant la transition du nombre quantique m . En admettant la « règle de sélection » suivant laquelle le nombre m peut varier de 0, +1 ou -1, on retrouve l'effet Zeeman normal. Remarquons que dans ce calcul, la constante h s'élimine, ce qui permet de retrouver un résultat de la Physique classique.

4. Le principe de correspondance. — Nous avons exposé au paragraphe 1 de ce chapitre l'idée de base du principe de correspondance. Nous pouvons la retrouver en partant de la remarque suivante : l'écart entre les théories classiques et les théories quantiques réside dans le fait

que la constante h a une valeur finie et cet écart s'évanouirait si h tendait vers zéro. Donc dans les phénomènes mettant en jeu un nombre énorme de quanta, on doit retrouver asymptotiquement la théorie classique puisqu'alors tout se passe comme si h était infiniment petit.

Nous allons d'abord prouver qu'il en est bien ainsi en ce qui concerne les fréquences en démontrant le théorème de correspondance pour les fréquences dont nous avons déjà parlé.

Considérons un système défini par des variables angulaires $\omega_1, \dots, \omega_i, \dots$ et un certain mouvement défini par des valeurs constantes des variables d'action J_1, \dots, J_i, \dots . Comme toutes les grandeurs attachées au système, le moment électrique $\vec{\mathcal{E}}$ peut s'exprimer sous la forme

$$\vec{\mathcal{E}} = \sum_{\tau} \mathbf{p}_{\tau}(J) e^{2\pi i(\tau\nu)t},$$

où τ représente l'ensemble des nombres entiers $\tau_1, \dots, \tau_i, \dots$ J l'ensemble des variables d'action et où $(\tau\nu) = \sum_i \tau_i \nu_i$. Enfin on a

$$\nu_k = \frac{\partial H}{\partial J_k}.$$

Supposons maintenant que le mouvement du système soit un mouvement quantifié à un grand nombre de quanta, c'est-à-dire que l'on ait $J_k = n_k h$ pour tous les J_k avec tous les n_k grands devant l'unité. D'après la conception classique de l'émission du rayonnement, il y aurait *émission simultanée* par le système de toutes les fréquences de la forme $(\tau\nu) = \sum_i \tau_i \nu_i$ figurant dans le développement de $\vec{\mathcal{E}}$, l'intensité de l'émission de la fréquence $(\tau\nu)$ étant déterminée par la valeur du \mathbf{p}_{τ} correspondant.

Dans la conception quantique, au contraire, il n'y a pas de rayonnement tant que dure l'état initial, mais il y a possibilité de transition de l'état initial défini par les $J_k = n_k h$ à d'autres états quantifiés d'énergie moindre et le passage à un état final défini par $J_k = n'_k h$ donne lieu à l'émission d'une radiation de fréquence

$$\nu_{nn'} = \frac{E_n - E_{n'}}{h},$$

d'après la loi de Bohr.

Or, d'après l'idée de base de la méthode de correspondance, si tous les n_k sont grands et si tous les $|\delta n_k| = |n_k - n'_k|$ sont petits devant

les n_k , on doit retomber sur les résultats de la théorie classique. Comme les n_k varient alors peu en valeur relative dans la transition $n \rightarrow n'$ on peut écrire pour la variation correspondante de l'énergie

$$E_n - E_{n'} = \sum_k \frac{\partial H}{\partial J_k} \delta J_k = \sum_k \frac{\partial H}{\partial J_k} h \delta n_k$$

et, par suite,

$$\nu_{nn'} = \sum_k \frac{\partial H}{\partial J_k} \delta n_k = \sum_k \delta n_k \nu_k.$$

Nous pouvons donc faire correspondre à chaque terme de fréquence (ν) du développement classique de $\vec{\mathcal{E}}$ la transition quantique

$$(n_1, n_2, \dots, n_l, \dots) \rightarrow (n_1 - \tau_1, n_2 - \tau_2, \dots, n_l - \tau_l, \dots)$$

puisqu, selon la règle de Bohr, cette transition donne précisément lieu à l'émission de la fréquence $(\tau\nu)$.

Si au début nous avons affaire à un ensemble d'atomes dans l'état d'énergie E_n , ces atomes subiront dans le domaine des grands nombres quantiques les diverses transitions possibles, de sorte qu'un observateur verra dans le rayonnement global émis par l'ensemble d'atomes toutes les fréquences prévues par la théorie classique. Statistiquement, le résultat classique sera donc vérifié malgré la différence profonde des processus d'émission imaginés par la théorie classique et par la théorie quantique. Nous avons ainsi démontré que les fréquences émises suivant la loi de Bohr coïncident asymptotiquement pour les grands nombres quantiques avec les fréquences prévues par la théorie classique à condition d'établir une certaine correspondance entre les transitions quantiques et les termes du développement classique du moment électrique en série de Fourier.

Vérifions le théorème de correspondance pour les fréquences dans le cas de l'atome d'hydrogène. Comme nous l'avons vu, nous avons alors

$$H(J) = -\frac{2\pi^2 m e^4}{J^2}, \quad \nu = \frac{\partial H}{\partial J} = \frac{4\pi^2 m e^4}{J^3},$$

ν étant la fréquence de rotation de l'électron sur son orbite. La fréquence émise lors de la transition $J + \tau h \rightarrow J$ est donnée par

$$\frac{E(J + \tau h) - E(J)}{h} = \frac{2\pi^2 m e^4}{h} \left[\frac{1}{J^2} - \frac{1}{(J + \tau h)^2} \right] \approx \tau \frac{4\pi^2 m e^4}{J^3} = \tau\nu$$

et le théorème est vérifié.

Le résultat précédent concernant les fréquences donnant une base solide à la méthode de correspondance, M. Bohr l'a étendu aux questions d'intensité et de polarisation. Il considère un nombre énorme d'atomes tous orientés de même et se trouvant dans le même état quantique initial défini par les valeurs $J_k = n_k h$ avec des n_k très grands. Soit

$$(5) \quad \mathcal{P}_x = \sum_{\tau} p_{\tau}^{(x)}(J) e^{2\pi i(\tau\nu)t}$$

le développement de Fourier de l'une des composantes rectangulaires du moment électrique de chacun de ces atomes. Au point de vue classique, chacun de ces atomes devrait émettre toutes les fréquences de la forme

$$\sum_k \tau_k \nu_k = \sum_k \tau_k \frac{\partial H}{\partial J_k} \quad (\text{avec des } \tau_k \text{ entiers}),$$

l'intensité de la composante parallèle à l'axe des x de la radiation de fréquence $(\tau\nu)$ étant déterminée par la valeur de $p_{\tau}^{(x)}$. Plus précisément, la quantité d'énergie rayonnée par unité de temps sous forme de radiation de fréquence $(\tau\nu)$ avec vibration électrique parallèle à Ox est en moyenne par atome égale à

$$\frac{64\pi^4(\tau\nu)^4}{3c^3} |p_{\tau}^{(x)}|^2$$

d'après la formule (14) du chapitre I. Un observateur qui analyserait, par exemple, à l'aide d'un réseau, le rayonnement de l'ensemble des atomes, y verrait donc des composantes spectrales ayant toutes les fréquences de la forme $\sum_k \tau_k \nu_k$, les intensités et les polarisations de chacune d'elles se trouvant définies par la loi qui vient d'être rappelée, appliquée aux trois axes x, y, z .

Dans la conception quantique, les choses se passent différemment puisque chacun des N atomes de l'ensemble ne peut émettre qu'une seule des fréquences $(\tau\nu)$ en effectuant la transition $n \rightarrow n - \tau$ correspondante. Il faudra donc ici introduire des probabilités de transition. On dira, par exemple, que tous les atomes de l'ensemble ont la probabilité $P_{n, n-\tau}^{(x)}$ d'effectuer la transition $n \rightarrow n - \tau$ avec émission de la fréquence $(\tau\nu)$ et vibration électrique parallèle à l'axe des x . Dès lors, N étant par hypothèse très grand, il y aura par seconde $NP_{n, n-\tau}^{(x)}$ atomes qui subiront la transition $n \rightarrow n - \tau$ dans ces conditions. L'observateur qui observe le rayonnement émis par l'ensemble d'atomes sans pouvoir

préciser le rôle de chaque atome, trouvera donc dans le rayonnement des composantes ayant toutes les fréquences $(\tau\nu)$, l'énergie émise en moyenne par seconde avec la fréquence $(\tau\nu)$ et une vibration électrique parallèle à Ox étant égale à $NP_{n,n-\tau}^{(x)} h(\tau\nu)$. Mais ici intervient l'idée de correspondance qui conduit à postuler que, dans le domaine des grands nombres quantiques, les faits observés doivent être en accord avec les résultats classiques et que, par suite, on doit avoir

$$NP_{n,n-\tau}^{(x)} h(\tau\nu) = N \frac{64 \pi^4 (\tau\nu)^4}{3 c^3} |p_{\tau}^{(x)}|^2,$$

d'où l'on tire l'expression de $P_{n,n-\tau}^{(x)}$

$$P_{n,n-\tau}^{(x)} = \frac{64 \pi^4 (\tau\nu)^4}{3 h c^3} |p_{\tau}^{(x)}|^2$$

avec naturellement des formules analogues en ce qui concerne les axes des y et des z .

Par ce principe, la théorie de correspondance détermine, à l'aide des coefficients du développement classique de $\vec{\mathcal{X}}$ en série de Fourier, les probabilités de transition dans le domaine des grands nombres quantiques défini par les conditions

$$(6) \quad n_k \gg 1, \quad n'_k \gg 1, \quad |\delta n_k| = |n_k - n'_k| \ll n_k.$$

Mais M. Bohr a été hardiment beaucoup plus loin. Il a étendu par induction le même principe en dehors du domaine des grands nombres quantiques.

Considérons un ensemble de très nombreux atomes dans un état quantique initial quelconque défini par un jeu de nombres quantiques sur lesquels nous ne faisons plus maintenant les hypothèses (6). Dans cet état initial, nous avons toujours un développement de Fourier pour \mathcal{X}_x de la forme (5) et à chacun des termes de ce développement, nous pouvons encore faire correspondre la transition quantique $n \rightarrow n - \tau$. La fréquence émise lors de cette transition est donnée par la règle de Bohr et elle n'est plus égale à $(\tau\nu)$ puisque le théorème de correspondance a été démontré sous les conditions (6) que nous ne supposons plus vérifiées. Soit alors $\nu_{n,n-\tau}$ la fréquence émise. Bohr suppose alors que la probabilité de transition $P_{n,n-\tau}^{(x)}$ est encore donnée approximativement par la relation

$$P_{n,n-\tau}^{(x)} = \frac{64 \pi^4 \nu_{n,n-\tau}^4}{3 h c^3} |p_{\tau}^{(x)}|^2.$$

Cette expression devrait permettre, d'après Bohr, une évaluation au moins approximative des intensités et des polarisations.

Remarquons qu'ainsi énoncé, le principe de Bohr n'était pas très satisfaisant. En effet, si l'on considère le mouvement quantifié initial, l'image classique fournit bien pour les $p_{\tau}^{(x)}$ des valeurs bien déterminées. Mais pourquoi choisir les $p_{\tau}^{(x)}$ de l'état initial plutôt que ceux de l'état final pour évaluer les probabilités de transition entre ces deux états ? Dans le domaine des très grands nombres quantiques, la difficulté ne se présentait pas car, les $p_{\tau}^{(x)}$ étant des fonctions continues des J_k et les J_k variant très peu en valeur relative pour les transitions qui satisfont aux conditions (6), les $p_{\tau}^{(x)}$ de l'état final diffèrent très peu de ceux de l'état initial, de sorte que l'indétermination s'évanouit. Mais quand les conditions (6) ne sont pas réalisées, y a-t-il lieu de prendre les $p_{\tau}^{(x)}$ de l'état initial, ceux de l'état final ou encore une sorte de moyenne de $p_{\tau}^{(x)}$ prise sur l'ensemble des états non quantifiés intermédiaires entre l'état initial et l'état final ? Bohr a fait une intéressante tentative dans son rapport au Troisième Conseil de Physique Solvay [2] pour préciser la manière dont une telle moyenne pourrait être définie, mais malgré ses efforts, le principe de correspondance pour les intensités a gardé dans l'ancienne théorie des Quanta un caractère un peu flou, et seule, la nouvelle Mécanique a pu, comme nous le verrons au chapitre XII, en donner un énoncé tout à fait précis.

Bien qu'imparfait, le principe de correspondance a cependant rendu dans le cadre de l'ancienne théorie des Quanta des services éminents en permettant de faire de nombreuses prévisions à peu près exactes pour les intensités et les polarisations des raies spectrales et en fournissant une démonstration des « règles de sélection » sur laquelle nous reviendrons plus loin. Guidés par l'idée fructueuse de correspondance, les théoriciens de l'École de Copenhague, collaborateurs ou élèves de Bohr, sont parvenus vers 1920 à développer des théories parfois un peu bâtardes, mais qui ont été en leur temps très utiles parce qu'elles ont préparé la voie aux théories actuelles. C'est ainsi que MM. Kramers et Heisenberg ont développé la première théorie quantique de la dispersion et ont pu donner la formule qui porte leur nom et dont nous aurons à reparler. Ce sont ces études orientées par le principe de correspondance qui ont mis à cette époque le jeune Werner Heisenberg sur la voie de la Mécanique quantique développée sous la forme de théorie des Matrices qui a été très importante pour l'essor de la nouvelle Mécanique et sur laquelle nous aurons l'occasion de revenir.

5. **Démonstration des règles de sélection par le principe de correspondance.** — Il faut d'abord définir ce que l'on appelle « règles de sélection ». La théorie de Bohr prévoit pour chaque sorte d'atome une suite indéfinie d'états stationnaires quantifiés dont les énergies divisées par h fournissent les termes spectraux à partir desquels on obtient par différence les fréquences des raies qu'en principe l'atome peut émettre. Mais toutes les raies ainsi prévues figurent-elles réellement dans le spectre de l'atome ? La réponse est : non, en général. Le fait qu'une fréquence spectrale est donnée par la loi de Bohr est une condition nécessaire, mais non suffisante pour l'existence de la raie correspondante. Il y a donc lieu de compléter la loi des fréquences de Bohr par des « règles de sélection » indiquant quelles sont, parmi les fréquences

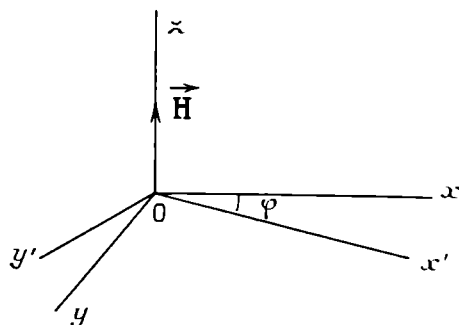


Fig. 22.

prévues par la loi de Bohr, celles qui figurent réellement dans le spectre de l'atome considéré.

Le principe de correspondance nous fournit un moyen de trouver ces règles de sélection, car il nous permet d'évaluer les intensités des raies spectrales prévues par la règle de Bohr et, s'il se trouve que l'intensité de certaines de ces raies soit nulle, elles seront absentes du spectre étudié. Par cette voie, on a pu arriver à des règles de sélection qui se trouvent être identiques à celles que Rubinovicz avait antérieurement énoncées en partant de considérations sur le moment de la quantité de mouvement des ondes lumineuses sphériques.

Pour bien faire comprendre sur un exemple comment on peut démontrer les règles de sélection à l'aide du principe de correspondance, reprenons le cas de l'atome d'hydrogène placé dans un champ magnétique uniforme. En ce cas, dans le système de référence de Larmor qui tourne autour du champ magnétique \mathbf{H} avec la vitesse de rotation uniforme $\omega = \frac{eH}{2m_0c}$, l'électron a le même mouvement que si le champ n'existait pas et si le système ne tournait pas.

Dans le système tournant, le moment électrique sera donc le même que si le système ne tournait pas : il aura la valeur $\vec{\mathcal{P}}^{(0)}$ qu'il aurait alors en l'absence de champ. Mais si nous prenons des axes $Oxyz$ tels que Oz coïncide avec la direction de \mathbf{H} , l'axe Ox' du système tournant fera avec Ox un angle $\varphi = \omega t$ et nous aurons à chaque instant

$$\mathcal{P}_x = \mathcal{P}_x^{(0)} \cos \varphi - \mathcal{P}_y^{(0)} \sin \varphi, \quad \mathcal{P}_y = \mathcal{P}_x^{(0)} \sin \varphi + \mathcal{P}_y^{(0)} \cos \varphi, \quad \mathcal{P}_z = \mathcal{P}_z^{(0)}.$$

Écrivons \mathcal{P}_x , \mathcal{P}_y , \mathcal{P}_z sous forme complexe en posant

$$\mathcal{P}^{(0)} = \left| \vec{\mathcal{P}}^{(0)} \right| e^{2\pi i \nu_1 t},$$

ν_1 étant la fréquence introduite au paragraphe 3. Nous pourrions écrire

$$\mathcal{P}_x + i\mathcal{P}_y = (\mathcal{P}_x^{(0)} + i\mathcal{P}_y^{(0)}) e^{i\varphi}, \quad \mathcal{P}_x - i\mathcal{P}_y = (\mathcal{P}_x^{(0)} - i\mathcal{P}_y^{(0)}) e^{-i\varphi}, \quad \mathcal{P}_z = \mathcal{P}_z^{(0)}.$$

En introduisant la variable angulaire

$$\omega_2 = \frac{\varphi}{2\pi} = \frac{\omega}{2\pi} t$$

à laquelle correspond la fréquence de Larmor $\nu_2 = \frac{eH}{2\pi m_0 c}$, nous obtenons

$$\begin{aligned} \mathcal{P}_x + i\mathcal{P}_y &= (\mathcal{P}_x^{(0)} + i\mathcal{P}_y^{(0)}) e^{2\pi i \nu_1 t}, \\ (\mathcal{P}_x - i\mathcal{P}_y) &= (\mathcal{P}_x^{(0)} - i\mathcal{P}_y^{(0)}) e^{-2\pi i \nu_1 t}, \quad \mathcal{P}_z = \mathcal{P}_z^{(0)}. \end{aligned}$$

Or, à la variable angulaire ω_2 correspond la variable d'action J_2 égale au produit par 2π du moment cinétique de l'électron autour de Oz et le nombre quantique que nous avons appelé m (nombre quantique magnétique). Les formules précédentes montrent alors, d'après la correspondance établie entre les transitions quantiques et les coefficients du développement de Fourier du moment électrique que, $\vec{\mathcal{P}}^{(0)}$ étant évidemment indépendant de ω_2 , les seules transitions associées à un moment électrique non nul sont celles pour lesquelles $\delta m = 0, \pm 1$.

Aux transitions pour lesquelles $\delta m = +1$, correspond un moment électrique tel que

$$\mathcal{P}_x + i\mathcal{P}_y \neq 0 \quad \text{avec} \quad \mathcal{P}_x - i\mathcal{P}_y = \mathcal{P}_z = 0$$

et la première de ces équations montre que

$$\mathcal{P}_y = \mathcal{P}_x e^{-i\frac{\pi}{2}},$$

c'est-à-dire que nous avons affaire à l'émission d'une onde circulaire

gauche autour de Oz . Aux transitions $\delta m = -1$ correspond un moment électrique tel que

$$\mathcal{X}_x + i\mathcal{X}_y = 0, \quad \mathcal{X}_x - i\mathcal{X}_y \neq 0, \quad \mathcal{X}_z = 0, \quad \text{d'où} \quad \mathcal{X}_y = \mathcal{X}_x e^{i\frac{\pi}{2}},$$

c'est-à-dire que nous avons affaire à une onde circulaire *droite* autour de Oz . Enfin, aux transitions $\delta m = 0$, correspond un moment électrique dirigé suivant Oz , car alors

$$\mathcal{X}_x = \mathcal{X}_y = 0 \quad \text{et} \quad \mathcal{X}_z \neq 0,$$

c'est-à-dire que nous avons affaire à une onde rectilignement polarisée avec vibration électrique parallèle à Oz .

Si nous portons ces résultats dans la formule précédemment obtenue pour les fréquences modifiées par l'effet Zeeman, savoir

$$\nu_{n.m.n'm'} = \nu_{n,n'}^{(0)} + \delta m \frac{eH}{4\pi m_0 c},$$

nous retrouvons immédiatement la théorie de l'effet Zeeman normal.

Un autre exemple de règle de sélection nous sera fourni par l'atome de la théorie de Sommerfeld. La variation relativiste de la masse de l'électron avec sa vitesse a pour effet d'imprimer à l'orbite électronique képlérienne un mouvement de rotation uniforme dans son plan autour du noyau. Si l'on prend le plan du mouvement en rosette comme plan xOy et l'axe des z normal à ce plan et passant par le noyau, il existe un système de référence tournant autour de Oz avec la vitesse du périhélie par rapport auquel la trajectoire est rigoureusement l'orbite képlérienne qui serait décrite dans le système de l'observateur en l'absence de la variation de la masse avec la vitesse. Dans ce système tournant, le moment électrique $\mathcal{X}^{(0)}$ est contenu dans le plan de la trajectoire : donc $\mathcal{X}_z^{(0)} = 0$ constamment. On obtient alors, comme précédemment,

$$\mathcal{X}_x + i\mathcal{X}_y = (\mathcal{X}_x^{(0)} + i\mathcal{X}_y^{(0)}) e^{i\varphi}, \quad \mathcal{X}_x - i\mathcal{X}_y = (\mathcal{X}_x^{(0)} - i\mathcal{X}_y^{(0)}) e^{-i\varphi}.$$

Si l'on prend comme variation angulaire $\omega_2 = \frac{\dot{\varphi}}{2\pi}$, où φ est l'angle de rotation lié au périhélie, on trouve

$$\mathcal{X}_x + i\mathcal{X}_y = (\mathcal{X}_x^{(0)} + i\mathcal{X}_y^{(0)}) e^{2\pi i \nu_2 t}, \quad \mathcal{X}_x - i\mathcal{X}_y = (\mathcal{X}_x^{(0)} - i\mathcal{X}_y^{(0)}) e^{-2\pi i \nu_2 t}, \quad \mathcal{X}_z = 0.$$

La variable J_2 conjuguée de ω_2 est égale au produit par 2π du moment de quantité de mouvement de l'électron autour du noyau et il lui correspond le nombre quantique azimutal k . On trouve donc la règle de

sélection $\partial k = \pm 1$, les deux possibilités correspondant à des émissions circulaires de sens inverse. La possibilité $\partial k = 0$ se trouve ici éliminée parce que $\mathfrak{A}_z = 0$. La règle $\partial k = \pm 1$ a rendu des services pour classer les raies spectrales dans le schéma de Sommerfeld.

Des règles de sélection analogues pouvaient être obtenues chaque fois que l'on parvenait à définir dans un système atomique un système de référence tournant d'un mouvement uniforme autour d'une direction D et par rapport auquel le mouvement du système était indépendant de la rotation. Nous n'insisterons pas sur ces généralisations.

BIBLIOGRAPHIE.

- [1] A. SOMMERFELD, *La constitution de l'atome et les raies spectrales*. Compléments du t. II, Blanchard, Paris, 1923.
- [2] *Comptes-rendus du Troisième Conseil de Physique Solvay, Atomes et Électrons*, Gauthier-Villars, Paris, 1923, p. 228.
- [3] LOUIS DE BROGLIE, *Le principe de correspondance et les interactions entre la matière et le rayonnement*, Hermann, Paris, 1938.

CHAPITRE IX.

IDÉES DE BASE ET ÉQUATIONS FONDAMENTALES DE LA MÉCANIQUE ONDULATOIRE.

1. **Point de départ.** — L'idée de base de la Mécanique ondulatoire a été la suivante : puisque l'expérience a révélé que, pour la lumière, il existe un aspect corpusculaire et un aspect ondulatoire des phénomènes reliés entre eux par la relation $W = h\nu$ où figure la constante h de Planck, on peut penser qu'il pourrait exister aussi pour les particules de matière un aspect corpusculaire et un aspect ondulatoire (ce dernier longtemps méconnu) qui seraient reliés entre eux par une relation où figurerait la constante de Planck.

Pour chercher à développer cette idée, il faut chercher à associer un élément périodique à l'idée de corpuscule. Imaginons un corpuscule qui se meut d'un mouvement rectiligne et uniforme en l'absence de tout champ : nous ne précisons pas où se trouve le corpuscule, nous n'envisageons que son état de mouvement. Ce mouvement a lieu dans une certaine direction que nous prendrons comme axe des z et il y est dynamiquement défini par les deux grandeurs énergie et quantité de mouvement dont les expressions relativistes sont

$$(1) \quad W = \frac{m_0 c^2}{\sqrt{1 - \beta^2}}, \quad \mathbf{p} = \frac{m_0 \mathbf{v}}{\sqrt{1 - \beta^2}} \quad \left(\beta = \frac{v}{c} \right)$$

dont on tire

$$p = |\mathbf{p}| = \frac{W}{c^2} v,$$

Le mouvement est ainsi défini pour un certain observateur galiléen A lié à un système de référence $Oxyz$ et dont le temps propre est t .

Soit maintenant un autre observateur galiléen B lié à un système de

référence qui se déplace avec la vitesse $v = \beta c$ dans la direction Oz . Nous pouvons dire que B est entraîné par le mouvement du corpuscule. Nous supposons que B a choisi un axe des z_0 qui glisse sur Oz et des axes des x_0 et des y_0 qui sont parallèles aux axes Ox et Oy . Alors les coordonnées x, y, z, t et x_0, y_0, z_0, t_0 sont reliées par les formules de la transformation simple de Lorentz. Pour B, le corpuscule est au repos et il pose

$$(2) \quad W_0 = m_0 c^2, \quad \mathbf{p}_0 = 0.$$

Pour introduire un élément périodique, nous allons d'abord chercher à le définir dans le système propre du corpuscule, c'est-à-dire dans le système de l'observateur B. Comme dans ce système tout est au repos, il est naturel d'introduire dans ce système un élément périodique stationnaire, analogue à une onde stationnaire. Désignant cet élément par Ψ_0 , nous écrirons donc

$$\Psi_0 = A e^{2\pi i \nu_0 t_0}$$

en employant la forme complexe de la représentation d'une onde stationnaire. Au cours du temps t_0 , Ψ_0 oscille avec la fréquence ν_0 . Nous admettons que A est une constante, de sorte que Ψ_0 a la même valeur en tout point du système B à l'instant t_0 . Nous pouvons nous représenter la répartition des valeurs de Ψ_0 dans le système propre en imaginant une infinité de petites horloges disposées en tous les points du système propre, synchronisées entre elles par l'observateur B et possédant la période $T_0 = \frac{1}{\nu_0}$. Si l'aiguille de l'horloge a pour longueur A, Ψ_0 sera en somme l'affixe de l'extrémité de l'aiguille dans le plan du cadran considéré comme le plan d'une variable complexe.

Quelle valeur allons-nous attribuer à la fréquence ν_0 ? Nous devons la définir à partir d'une grandeur caractérisant le corpuscule dans son système propre et cette grandeur ne peut être que W_0 . La relation qui se présente alors tout naturellement à l'esprit quand on connaît les théories quantiques est la relation

$$\nu_0 = \frac{W_0}{h},$$

analogue à la relation d'Einstein pour les photons.

Comment alors va se manifester l'élément périodique ainsi introduit pour l'observateur qui voit passer le corpuscule avec la vitesse βc ? Une remarque que nous avons faite à propos du ralentissement des

horloges ⁽¹⁾ nous fournit tout de suite la solution : les phases de l'élément périodique introduit sont, pour l'observateur A, réparties dans l'espace comme celles d'une onde de fréquence ν se propageant le long de l'axe des z avec la vitesse \mathcal{V} , ν et \mathcal{V} ayant les valeurs :

$$\nu = \frac{\nu_0}{\sqrt{1-\beta^2}}, \quad \mathcal{V} = \frac{c^2}{v} = \frac{c}{\beta}.$$

En effet en remplaçant t_0 par $\frac{t - \frac{\beta z}{c}}{\sqrt{1-\beta^2}}$ dans l'expression de Ψ_0 , on trouve

$$\Psi = A e^{2\pi i \frac{\nu_0}{\sqrt{1-\beta^2}} \left(t - \frac{\beta z}{c} \right)} = A e^{2\pi i \nu \left(t - \frac{z}{\mathcal{V}} \right)}.$$

La longueur d'onde de cette onde est

$$\lambda = \frac{\mathcal{V}}{\nu}.$$

En tenant compte des expressions (1) et (2) de W et de W_0 , on obtient aisément

$$W = \frac{W_0}{\sqrt{1-\beta^2}} = \frac{h\nu_0}{\sqrt{1-\beta^2}} = h\nu.$$

de sorte que la relation du quantum « énergie = h . fréquence », qui a été postulée dans le système B, se trouve aussi être exacte dans le système A et plus généralement dans tout système galiléen. De plus, on a

$$\lambda = \frac{\mathcal{V}}{\nu} = \frac{c^2}{v} \frac{h}{h\nu} = \frac{c^2}{v} \frac{h}{W} = \frac{h}{p}.$$

La formule fondamentale $\lambda = \frac{h}{p}$ prend, pour les faibles vitesses, quand on peut se contenter de la Mécanique newtonienne et poser $p = mc$, la forme bien connue

$$\lambda = \frac{h}{mv}.$$

Ces expressions de la longueur d'onde ont été vérifiées par les expériences de diffraction des électrons par les cristaux, ainsi que par les expériences analogues faites avec certains noyaux d'atomes et avec les neutrons.

(1) Voir chapitre II, fin du paragraphe 4.

Il résulte des formules précédentes que nous pouvons écrire

$$\Psi = A e^{\frac{2\pi i}{h} (Wt - \rho z)}$$

ou, plus généralement, pour un mouvement rectiligne uniforme qui n'a pas lieu dans la direction de l'axe des z

$$\Psi = A e^{\frac{2\pi i}{h} (Wt - \rho_x x - \rho_y y - \rho_z z)} = A e^{\frac{2\pi i}{h} \int (\mathbf{I} ds)},$$

où \mathbf{I} est le vecteur d'espace-temps « impulsion d'Univers » pour le corpuscule considéré (1). On voit alors que la phase de l'élément ondulatoire Ψ coïncide au facteur $\frac{2\pi i}{h}$ près avec l'action $\int (\mathbf{I} ds)$ du corpuscule. On comprend alors que le principe de moindre action pour le corpuscule n'est qu'une traduction du principe de Fermat (principe du temps minimum) pour l'onde Ψ qui lui est associée. On parvient ainsi à l'idée que la Mécanique ancienne (aussi bien sous la forme relativiste d'Einstein que sous la forme classique de Newton) ne doit être qu'une approximation ayant le même degré de validité que l'Optique géométrique (qui, on le sait, dérive tout entière du principe de Fermat). On est donc amené à concevoir la nécessité de construire une nouvelle Mécanique, une Mécanique ondulatoire, qui soit par rapport à la Mécanique ancienne ce qu'est l'Optique ondulatoire par rapport à l'Optique géométrique.

Telles sont les considérations qui ont été le point de départ de la Mécanique ondulatoire : elles sont imprégnées de Relativité. Néanmoins, la nouvelle Mécanique s'est développée ensuite sous une forme non relativiste en suivant une voie que nous allons maintenant examiner, quitte à revenir plus tard au point de vue relativiste.

2. La théorie de Jacobi en Mécanique analytique. — Nous voulons d'abord rappeler quelques résultats bien connus de la Mécanique rationnelle dans le cas où l'on envisage un corpuscule de masse donnée m soumis à un champ de force également donné qui dérive d'une fonction potentielle $V(x, y, z, t)$.

Dans l'ancienne Mécanique, on admet qu'un corpuscule, un point matériel, est un objet extrêmement petit qui occupe à chaque instant dans l'espace une position bien déterminée que l'on peut caractériser

(1) Voir chapitre III, paragraphe 4.

par l'emploi de trois coordonnées rectangulaires x, y, z . Au cours du temps t , la position du corpuscule varie en général et il décrit sous l'influence du champ de force une certaine courbe que l'on nomme la trajectoire. Au point de vue mathématique, la trajectoire du corpuscule et la manière dont elle est décrite sont connues si l'on obtient l'expression des coordonnées x, y, z en fonction du temps. Les équations dont partait la théorie classique pour déterminer le mouvement du corpuscule sont les équations classiques

$$m \frac{d^2 x}{dt^2} = - \frac{\partial V}{\partial x}, \quad m \frac{d^2 y}{dt^2} = - \frac{\partial V}{\partial y}, \quad m \frac{d^2 z}{dt^2} = - \frac{\partial V}{\partial z}.$$

Ces trois équations différentielles simultanées du second ordre ont des solutions qui sont parfaitement déterminées quand on connaît les valeurs initiales de $x, y, z, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}$, c'est-à-dire la position et la vitesse initiales du mobile.

Nous ne rappellerons pas les très nombreuses conséquences que les traités de Mécanique rationnelle tirent des équations précédentes et nous nous bornerons à énoncer un théorème de Mécanique analytique dû à Jacobi qui prépare le passage à la nouvelle Mécanique.

Dans le champ de force donné, il y a une infinité de mouvements possibles pour le corpuscule suivant sa position et sa vitesse initiales. Le théorème de Jacobi a pour but d'étudier l'ensemble de ces mouvements et de les classer. Voici l'énoncé de ce théorème.

Si l'on parvient à trouver une intégrale complète $S(x, y, z, t, \alpha, \beta, \gamma)$, c'est-à-dire une solution dépendant de trois constantes arbitraires non additives, de l'équation aux dérivées partielles du premier ordre

$$\frac{1}{2m} \left[\left(\frac{\partial S}{\partial x} \right)^2 + \left(\frac{\partial S}{\partial y} \right)^2 + \left(\frac{\partial S}{\partial z} \right)^2 \right] + V(x, y, z, t) = \frac{\partial S}{\partial t}$$

dite « équation de Jacobi », alors les équations

$$\frac{\partial S}{\partial \alpha} = a, \quad \frac{\partial S}{\partial \beta} = b, \quad \frac{\partial S}{\partial \gamma} = c,$$

- où a, b, c sont des constantes, représentent l'un des mouvements possibles du corpuscule dans le champ de force. De plus, les composantes de la quantité de mouvement du corpuscule, lorsqu'en exécutant

ce mouvement il occupe la position x, y, z à l'instant t , sont données par les relations

$$p_x = mv_x = -\frac{\partial S}{\partial x}, \quad p_y = mv_y = -\frac{\partial S}{\partial y}, \quad p_z = mv_z = -\frac{\partial S}{\partial z};$$

$$\mathbf{p} = -\text{grad } S.$$

Le grand intérêt du théorème de Jacobi est de permettre de classer les trajectoires. En effet, nous pouvons, connaissant l'intégrale complète $S(x, y, z, t, \alpha, \beta, \gamma)$, répartir les mouvements possibles en diverses classes, chaque classe correspondant à un même choix des « constantes primaires » α, β, γ . Chacune de ces classes comprend une infinité de mouvements possibles, chacun étant caractérisé par un certain choix des « constantes secondaires » a, b, c .

Nous allons maintenant fixer notre attention sur le cas particulièrement important des champs permanents où V ne dépend pas explicitement du temps. Nous savons qu'alors il y a conservation de l'énergie, c'est-à-dire que, durant tout le cours du mouvement, la somme $\frac{1}{2}mv^2 + V$ reste égale à une constante E . Il est naturel de faire jouer à E le rôle d'une des constantes primaires, par exemple γ . On peut alors trouver une intégrale complète de l'équation de Jacobi en posant

$$S(x, y, z, t, \alpha, \beta, E) = Et - S_1(x, y, z, \alpha, \beta, E),$$

où S_1 , la fonction de Jacobi raccourcie, ne dépend pas du temps et est une intégrale complète, dépendant de la constante E et de deux autres constantes arbitraires non additives α et β , de l'équation de Jacobi raccourcie

$$(3) \quad \frac{1}{2m} \left[\left(\frac{\partial S_1}{\partial x} \right)^2 + \left(\frac{\partial S_1}{\partial y} \right)^2 + \left(\frac{\partial S_1}{\partial z} \right)^2 \right] + V(x, y, z) = E.$$

Quand on a obtenu une intégrale complète $S_1(x, y, z, \alpha, \beta, E)$, il résulte du théorème de Jacobi que l'un des mouvements possibles du corpuscule est donné par

$$\frac{\partial S_1}{\partial \alpha} = a, \quad \frac{\partial S_1}{\partial \beta} = b, \quad \frac{\partial S}{\partial \gamma} = \frac{\partial S}{\partial E} = t = \frac{\partial S_1}{\partial E} = c$$

et la quantité de mouvement du corpuscule en chaque point de la trajectoire correspondante est donnée par

$$\mathbf{p} = \text{grad } S_1.$$

Les mouvements possibles du corpuscule dans le champ permanent.

se répartissent donc en classes correspondant à une même intégrale complète S_1 avec des valeurs déterminées de l'énergie E et des deux autres constantes primaires α et β . Chaque classe contient une triple infinité de mouvements, chacun d'eux étant caractérisé par une valeur donnée des trois constantes secondaires a, b, c .

Les équations

$$\frac{\partial S_1}{\partial x} = a \quad \text{et} \quad \frac{\partial S_1}{\partial \beta} = b$$

ne contiennent pas le temps : elles définissent une courbe de l'espace qui est la trajectoire du corpuscule. L'équation $t - \frac{\partial S_1}{\partial E} = c$ que l'on peut écrire $\frac{\partial S_1}{\partial E} = t - t_0$ donne le mouvement du mobile le long de la trajectoire : on l'appelle parfois l'équation de l'horaire. On peut donc dire que, dans le cas des champs permanents, l'étude de la trajectoire peut être séparée de celle du mouvement sur la trajectoire. Cette circonstance ne se présente pas dans le cas des champs variables avec le temps.

Un autre résultat important valable dans le cas des champs permanents est le suivant : les trajectoires d'une même classe correspondant à une même intégrale complète $S_1(x, y, z, \alpha, \beta, E)$ avec valeurs données de α, β, E sont orthogonales aux surfaces $S_1 = \text{const}$. Ceci se déduit immédiatement de la relation

$$\mathbf{p} = \text{grad } S_1.$$

Cette propriété des trajectoires d'être normales aux surfaces $S_1 = \text{const}$. conduit facilement au principe de moindre action de Maupertuis dont voici l'énoncé :

Dans un champ de force permanent, toute trajectoire passant par deux points A et B de l'espace est caractérisée par le fait que l'intégrale $\int_A^B \sqrt{2m[E - V(x, y, z)]} dl$ prise le long de la trajectoire de A en B est plus petite que la valeur de la même intégrale prise sur toute courbe infiniment voisine de la trajectoire et allant de A en B, la valeur de E étant maintenue constante dans cette variation.

Pour terminer, nous illustrerons par un exemple très simple ces considérations sur les mouvements possibles d'un corpuscule dans un

champ permanent. Nous supposons le champ nul : alors $V = 0$ et l'équation de Jacobi raccourcie

$$\frac{1}{2m} \left[\left(\frac{\partial S_1}{\partial x} \right)^2 + \left(\frac{\partial S_1}{\partial y} \right)^2 + \left(\frac{\partial S_1}{\partial z} \right)^2 \right] = E$$

admet l'intégrale complète

$$S_1 = \sqrt{2mE} (\alpha x + \beta y + \sqrt{1 - \alpha^2 - \beta^2} z).$$

Nous obtenons les trajectoires possibles en écrivant

$$\begin{aligned} \frac{\partial S_1}{\partial x} &= \sqrt{2mE} \left(x - \frac{\alpha z}{\sqrt{1 - \alpha^2 - \beta^2}} \right) = a, \\ \frac{\partial S_1}{\partial y} &= \sqrt{2mE} \left(y - \frac{\beta z}{\sqrt{1 - \alpha^2 - \beta^2}} \right) = b. \end{aligned}$$

Les trajectoires d'une même classe sont donc ici des droites de cosinus directeurs α , β , $\gamma = \sqrt{1 - \alpha^2 - \beta^2}$. Le mouvement sur ces droites est donné par

$$\frac{\partial S_1}{\partial E} = \sqrt{\frac{m}{2E}} (\alpha x + \beta y + \sqrt{1 - \alpha^2 - \beta^2} z) = t - t_0$$

ou

$$\alpha x + \beta y + \gamma z = \sqrt{\frac{2E}{m}} (t - t_0).$$

C'est un mouvement uniforme de vitesse $v = \sqrt{\frac{2E}{m}}$, ce qui devait être puisqu'ici nous avons $\frac{1}{2} m v^2 = E$. Les surfaces $S_1 = \text{const.}$ sont des plans parallèles dont la normale a pour cosinus directeurs α , β , γ et les trajectoires de la classe correspondante sont bien les normales à ces plans. Enfin, la relation $\mathbf{p} = \text{grad} S_1$ se vérifie aussi très aisément.

3. Rappel de notions fondamentales de la théorie des ondes. — Ayant sommairement rappelé quelques points de l'ancienne Mécanique, nous devons aussi, pour préparer le passage à la nouvelle Mécanique, rappeler quelques notions sur la propagation des ondes monochromatiques dans un milieu isotrope réfringent et dispersif. Cette propagation est régie par l'équation

$$\Delta \Psi = \frac{1}{\mathfrak{V}^2} \frac{\partial^2 \Psi}{\partial t^2},$$

où \mathfrak{V} est une fonction du point considéré (x, y, z) et de la fréquence ν

de l'onde monochromatique. L'onde monochromatique est représentée par une fonction de la forme

$$\Psi(x, y, z, t) = f(x, y, z) e^{2\pi i \nu t}$$

et, si l'on pose pour définir l'indice de réfraction n :

$$\frac{1}{\mathfrak{V}} = \frac{n(x, y, z, \nu)}{\mathfrak{V}_0},$$

où \mathfrak{V}_0 est la valeur de \mathfrak{V} dans un milieu de référence homogène et non dispersif (par exemple le vide dans la théorie de la lumière), on a

$$\Delta \Psi + \frac{4\pi^2 \nu^2 n^2}{\mathfrak{V}_0^2} \Psi = 0.$$

Rigoureusement, l'étude de la propagation doit se faire en cherchant les solutions de cette équation correspondant à des conditions traduisant le problème posé, mais il arrive assez fréquemment que l'on puisse résoudre le problème par une méthode approximative qui constitue « l'Optique géométrique ». Pour comprendre ce qu'est cette méthode, considérons d'abord le cas où l'indice n ne dépend pas de x, y, z (milieu réfringent homogène). Alors on obtiendra une solution rigoureuse de l'équation de propagation en considérant l'onde plane

$$\Psi = a e^{2\pi i \nu \left[t - \frac{n(\nu)}{\mathfrak{V}_0} (\alpha x + \beta y + \gamma z) \right]},$$

avec $\gamma^2 = 1 - \alpha^2 - \beta^2$. La constante a est appelée l'amplitude de l'onde plane; la fonction

$$\varphi = \nu \left[t - \frac{n(\nu)}{\mathfrak{V}_0} (\alpha x + \beta y + \gamma z) \right]$$

est sa « phase ». Les surfaces d'égale phase $\varphi = \text{const.}$ sont ici des plans parallèles. Au cours du temps, les valeurs de la phase et, par suite, celles de la fonction d'onde Ψ progressent dans la direction (α, β, γ) avec la vitesse $\mathfrak{V} = \frac{\mathfrak{V}_0}{n(\nu)}$. A un instant donné, on retrouve les mêmes valeurs de Ψ sur les plans d'égale phase parallèles et séparés les uns des autres par la distance constante $\lambda = \frac{\mathfrak{V}}{\nu}$ appelée « longueur d'onde ». En un point donné de l'espace, on retrouve les mêmes valeurs de Ψ à des intervalles de temps égaux à la période $T = \frac{1}{\nu}$. Les droites normales aux plans d'égale phase, dont les cosinus directeurs sont α, β, γ , sont par définition les « rayons » de la propagation d'ondes.

Tout ceci est bien connu. Mais considérons maintenant un milieu réfringent non homogène, c'est-à-dire où l'indice n varie d'un point à l'autre. Une onde monochromatique pourra s'écrire

$$\Psi = a(x, y, z) e^{2\pi i[\nu t - \varphi_1(x, y, z)]}.$$

Si, à l'intérieur du corps réfringent, la variation de l'indice est négligeable à l'échelle de la longueur d'onde, on peut voir aisément que φ_1 doit obéir approximativement à l'équation

$$(4) \quad \left(\frac{\partial \varphi_1}{\partial x}\right)^2 + \left(\frac{\partial \varphi_1}{\partial y}\right)^2 + \left(\frac{\partial \varphi_1}{\partial z}\right)^2 = \frac{\nu^2 n^2(x, y, z, \nu)}{\mathfrak{V}_0^2} = \frac{1}{\lambda^2(x, y, z)},$$

λ étant la longueur d'onde $\frac{\mathfrak{V}}{\nu}$ qui varie d'un point à l'autre. L'amplitude a varie alors lentement à grande échelle, mais reste presque constante à l'échelle de la longueur d'onde. Le point remarquable est que l'équation obtenue pour φ_1 , dite équation de l'Optique géométrique, permet de déterminer les variations rapides de la phase sans se préoccuper des variations lentes de l'amplitude.

Soit alors $\varphi_1(x, y, z, \alpha, \beta, \nu)$ une intégrale complète de l'équation de l'Optique géométrique : la fonction d'onde

$$\Psi = a e^{2\pi i[\nu t - \varphi_1(x, y, z, \alpha, \beta, \nu)]},$$

où a est une fonction lentement variable à grande échelle de x, y, z sera, avec l'hypothèse faite sur les variations de l'indice n , une solution *approximative* de l'équation des ondes dans le milieu réfringent et représentera donc une des propagations d'ondes possibles dans le milieu. Par définition, les courbes orthogonales aux surfaces d'égale phase $\varphi_1 = \text{const.}$ sont les « rayons » de la propagation d'ondes.

En appliquant le même raisonnement qui conduit en Mécanique analytique au principe de Maupertuis, on peut alors démontrer le « principe de Fermat » dont voici l'énoncé :

Si une courbe C joignant deux points A et B d'un milieu réfringent à indice lentement variable est un rayon d'une propagation d'ondes, sa forme est déterminée par le fait que l'intégrale

$$\int_A^B \frac{dl}{\lambda} = \int_A^B \frac{n \nu}{\mathfrak{V}_0} dl$$

prise le long de C de A en B est plus petite que la même intégrale prise le long de toute courbe infiniment voisine de C joignant aussi les points A et B.

L'Optique géométrique fondée sur l'équation (4) et le principe de Fermat qui en découle ne constituent qu'une approximation. Quand l'indice de réfraction du milieu varie sensiblement à l'échelle de la longueur d'onde (et également au voisinage immédiat des sources et foyers et des bords de l'ombre géométrique où l'amplitude varie toujours rapidement), l'Optique géométrique cesse d'être valable : il en est de même quand des obstacles rencontrés en cours de route par l'onde produisent des phénomènes d'interférences et de diffraction. Dans ces cas, on ne peut plus calculer les variations de la phase indépendamment de celles de l'amplitude et il faut, pour étudier les phénomènes, partir de l'équation rigoureuse de propagation. Remarquons d'ailleurs que, si la longueur d'onde était infiniment petite, l'Optique géométrique serait toujours valable.

4. Nouvelle manière d'accéder à la Mécanique ondulatoire. — La grande analogie de forme de l'équation de Jacobi raccourcie (3) et de l'Optique géométrique (4), analogie qui se traduit par l'identité de forme des principes de Maupertuis et de Fermat, avait frappé, il y a plus d'un siècle, l'esprit pénétrant du géomètre écossais Hamilton, mais c'est seulement le développement de la théorie des Quanta qui a permis de donner un sens physique à cette analogie formelle. C'est ce que nous allons maintenant indiquer et, par cette voie, nous allons retrouver des résultats tout à fait analogues à ceux que la méthode relativiste du paragraphe 1 nous avait donnés.

Commençons par comparer le mouvement d'un corpuscule en l'absence de champ ($V = 0$) avec la propagation d'une onde dans le milieu de référence ($n = 1$). Pour le corpuscule, la fonction S_1 peut être prise égale à l'intégrale complète de l'équation de Jacobi raccourcie

$$S_1 = \sqrt{2mE}(ax + \beta y + \gamma z)$$

avec

$$E = \frac{1}{2} m v^2 \quad \text{et} \quad \alpha^2 + \beta^2 + \gamma^2 = 1.$$

Pour l'onde, nous avons alors l'équation en φ_1

$$\left(\frac{d\varphi_1}{dx}\right)^2 + \left(\frac{d\varphi_1}{dy}\right)^2 + \left(\frac{d\varphi_1}{dz}\right)^2 = \frac{v^2}{\mathcal{Q}_0^2}$$

qui admet l'intégrale complète

$$\varphi_1 = \frac{v}{\mathcal{Q}_0}(\alpha x + \beta y + \gamma z) = \frac{1}{\lambda}(\alpha x + \beta y + \gamma z)$$

avec

$$\alpha^2 + \beta^2 + \gamma^2 = 1 \quad \text{et} \quad \lambda = \frac{\mathcal{Q}_0}{\nu}.$$

Les fonctions S et φ correspondant à S_1 et φ_1 sont

$$\begin{aligned} S &= Et - m\nu(\alpha x + \beta y + \gamma z), \\ \varphi &= \nu t - \frac{1}{\lambda}(\alpha x + \beta y + \gamma z), \end{aligned}$$

où l'on supposera que α, β, γ ont les mêmes valeurs.

Un des principes essentiels de la théorie des Quanta, qui s'est manifesté notamment dans la théorie des Quanta de lumière d'Einstein, c'est de toujours admettre une proportionnalité entre l'énergie d'un corpuscule et la fréquence d'une onde, proportionnalité de la forme $E = h\nu$, h étant la constante de Planck. Pour établir cette proportionnalité entre l'énergie d'un corpuscule dont S est la fonction de Jacobi et la fréquence d'une « onde associée » dont φ serait la phase, il faut poser

$$S = h\varphi,$$

car, en égalant les coefficients de t , on trouve bien alors $E = h\nu$. Mais on trouve du même coup en égalant les coefficients de $\alpha x + \beta y + \gamma z$

$$\lambda = \frac{h}{m\nu},$$

équation qui établit une liaison entre la quantité de mouvement du corpuscule et la longueur d'onde de l'onde associée.

Donc en posant $S = h\varphi$, nous avons associé au mouvement rectiligne et uniforme du corpuscule d'énergie E et de quantité de mouvement $m\mathbf{v} = \mathbf{p}$ la propagation dans la direction du mouvement d'une onde plane monochromatique

$$\Psi = a e^{\frac{2\pi i}{h} S} = a e^{\frac{2\pi i}{h} [Et - \mu(\alpha x + \beta y + \gamma z)]}$$

de fréquence $\nu = \frac{E}{h}$ et de longueur d'onde $\lambda = \frac{h}{p}$. Nous retrouvons donc exactement sous une forme non relativiste les conclusions obtenues au paragraphe 1.

Mais cette correspondance entre mouvement d'un corpuscule et propagation d'une onde associée n'est toujours obtenue que pour le cas très particulier du mouvement rectiligne et uniforme en l'absence de champ. Nous allons chercher à la généraliser au cas du mouvement d'un corpuscule sous l'action d'un champ permanent défini par l'énergie poten-

tielle $V(x, y, z)$. Pour le faire, nous comparerons le mouvement du corpuscule à la propagation d'une onde dans un milieu où l'indice $n(x, y, z, \nu)$ varie d'un point à un autre. La fonction S de Jacobi pour le corpuscule s'écrivant alors $Et - S_1$ et la fonction de phase φ de l'onde associée s'écrivant $\nu t - \varphi_1$, nous poserons encore

$$S = h\varphi,$$

d'où nous tirerons

$$E = h\nu, \quad S_1 = h\varphi_1.$$

Les fonctions S_1 et φ_1 satisfont aux équations (3) et (4)

$$\begin{aligned} \left(\frac{\partial S_1}{\partial x}\right)^2 + \left(\frac{\partial S_1}{\partial y}\right)^2 + \left(\frac{\partial S_1}{\partial z}\right)^2 &= 2m[E - V(x, y, z)], \\ \left(\frac{\partial \varphi_1}{\partial x}\right)^2 + \left(\frac{\partial \varphi_1}{\partial y}\right)^2 + \left(\frac{\partial \varphi_1}{\partial z}\right)^2 &= \frac{\nu^2 n^2}{\mathcal{V}_0^2}. \end{aligned}$$

Pour avoir $S_1 = h\varphi_1$, il faut donc poser

$$2m[E - V(x, y, z)] = h^2 \frac{\nu^2 n^2}{\mathcal{V}_0^2}.$$

Or $\frac{\mathcal{V}_0}{\nu}$ est la longueur d'onde correspondant à la fréquence ν en l'absence de champ; on a donc

$$\frac{\mathcal{V}_0}{\nu} = \frac{h}{\sqrt{2mE}}.$$

Par suite, en éliminant \mathcal{V}_0 entre les deux dernières relations, nous obtenons

$$n^2 = 1 - \frac{V(x, y, z)}{E} = 1 - \frac{V(x, y, z)}{h\nu},$$

relation qui a l'allure d'une formule de dispersion. Nous obtenons ainsi l'indice de réfraction variable que nous devons attribuer à la région de l'espace où règne le potentiel V quand nous envisageons la propagation de l'onde que la relation $S = h\varphi$ associe au mouvement du corpuscule. En chaque point, nous pouvons définir une longueur d'onde locale

$$\lambda(x, y, z) = \frac{\mathcal{V}_0}{n\nu} = \frac{h}{\sqrt{2m[E - V(x, y, z)]}} = \frac{h}{m\nu},$$

ν étant la vitesse du corpuscule en (x, y, z) .

Portons

$$\frac{n^2 \nu^2}{\mathcal{V}_0^2} = \frac{2m}{h^2} [E - V(x, y, z)]$$

dans l'équation des ondes

$$\Delta \Psi - 4\pi^2 \frac{n^2 \nu^2}{\mathcal{V}_0^2} \Psi = 0,$$

on obtient pour l'équation des ondes associées au corpuscule la fameuse « équation de Schrödinger »

$$(5) \quad \Delta \Psi + \frac{8\pi^2 m}{h^2} [E - V(x, y, z)] \Psi = 0.$$

5. Remarques importantes et généralisation au cas des champs non permanents. — Nous avons supposé jusqu'ici que le corpuscule obéissait aux lois de la Mécanique classique, ce qui permet de considérer la fonction S de Jacobi et nous sommes ainsi parvenus à la conclusion suivante : aux mouvements classiques d'énergie bien définie du corpuscule, correspond la propagation à l'*approximation de l'Optique géométrique* d'une onde monochromatique obéissant à l'équation (5).

Chaque fois que l'Optique géométrique est valable pour la propagation de l'onde Ψ régie par l'équation (5), nous pouvons poser

$$\Psi = a e^{\frac{2\pi i}{h} S} = a e^{\frac{2\pi i}{h} [Et - S_1(x, y, z)]}$$

et les trajectoires d'une même classe prévue par la Mécanique classique, normales aux surfaces $S_1 = \text{const.}$, ne sont pas autre chose que les rayons de la propagation d'onde associée qui sont normaux aux surfaces d'égale phase.

Il semble donc y avoir, dans les limites de l'Optique géométrique, une coïncidence parfaite entre les deux points de vue. Néanmoins, nous voyons déjà qu'une même onde Ψ correspond à tout l'ensemble des mouvements d'une même classe conçus à la façon classique. Du point de vue classique, le corpuscule qui est supposé bien localisé dans l'espace à tout instant décrit l'une des trajectoires possibles, tandis que le point de vue de la Mécanique ondulatoire nous oblige déjà, même dans les limites de l'Optique géométrique, à adopter une sorte de vue statistique qui envisage simultanément toutes les trajectoires d'une même classe.

Mais une divergence beaucoup plus grave des deux points de vue apparaît quand la propagation de l'onde Ψ associée au corpuscule, propagation régie par l'équation (5) ne peut plus être décrite par l'approximation de l'Optique géométrique. Comme l'idée essentielle de la nouvelle Mécanique est de donner le rôle essentiel à la propagation,

il paraît naturel en ce cas d'abandonner la description du mouvement corpusculaire donnée par la Mécanique classique comme constituant une approximation qui n'est plus valable. En d'autres termes, la Mécanique classique ne serait utilisable que quand l'Optique géométrique serait valable pour la propagation de l'onde associée : à la fin du prochain chapitre, nous précisons plus exactement comment cela a lieu. Pour que l'emploi de la Mécanique classique soit possible, il faut donc que l'indice n défini précédemment varie lentement à l'échelle de la longueur d'onde et il faut de plus, que l'onde ne rencontre pas sur son parcours d'obstacles provoquant des phénomènes d'interférences et de diffraction. D'après la formule $\lambda = \frac{h}{mv}$ et le fait que l'Optique géométrique serait toujours valable si la longueur d'onde était infiniment petite, on voit que la Mécanique classique serait toujours valable si h était infiniment petit. Les modifications que la nouvelle Mécanique a dû apporter aux conceptions classiques nous apparaissent donc comme intimement liées à la valeur finie de la constante de Planck.

Nous sommes arrivés à substituer aux équations de la Dynamique ponctuelle du point matériel dans un champ permanent l'équation de propagation d'une onde monochromatique. Mais, comme nous le verrons, on est amené en Mécanique ondulatoire, comme dans l'ancienne théorie des ondes, à considérer des trains d'ondes non monochromatiques que l'analyse de Fourier permet de représenter par une superposition d'ondes monochromatiques de différentes fréquences et de différentes directions de propagation. Il est donc nécessaire de chercher une forme de l'équation des ondes qui puisse rester valable dans le cas où l'onde Ψ n'est pas monochromatique. Naturellement, cette équation devra nous redonner l'équation (5) dans le cas d'un Ψ monochromatique.

L'équation

$$\Delta \Psi - \frac{8\pi^2 m}{h^2} V(x, y, z) \Psi = \frac{4\pi i m}{h} \frac{\partial \Psi}{\partial t}$$

satisfait à ces conditions, car pour une onde Ψ monochromatique de fréquence $\frac{E}{h}$, on a

$$\frac{\partial \Psi}{\partial t} = \frac{2\pi i}{h} E \Psi$$

et, en substituant dans l'équation précédente, on retrouve l'équation (5).

L'avantage qu'il y a à écrire l'équation des ondes sous cette forme plus générale est très grand. D'abord, toute onde Ψ formée par une

superposition quelconque d'ondes monochromatiques va être solution de l'équation nouvelle, ce qui permettra de représenter un train d'ondes par une solution Ψ . De plus, comme nous ne sommes plus obligés de considérer une onde monochromatique, le temps ne va plus jouer de rôle particulier et nous allons pouvoir nous affranchir de la condition que le champ soit permanent. Nous allons donc pouvoir supposer que la fonction V dépend à la fois des coordonnées et du temps et que la forme la plus générale de l'équation de propagation est

$$\Delta\Psi - \frac{8\pi^2 m}{h^2} V(x, y, z, t) \Psi = \frac{4\pi im}{h} \frac{\partial\Psi}{\partial t}.$$

Nous examinerons dans le prochain chapitre comment nous devons utiliser cette équation pour la prévision des phénomènes observables où le corpuscule manifeste sa présence.

6. Le théorème de la vitesse de groupe. — Rappelons d'abord que, dans la théorie générale des ondes, on appelle « groupe d'ondes » un train d'ondes formé par la superposition d'ondes planes monochromatiques de fréquences et de directions de propagation très voisines. On démontre qu'un tel groupe d'ondes occupant une région limitée de l'espace se transporte sensiblement en bloc d'un mouvement d'ensemble avec une vitesse constante U dite « vitesse de groupe ». D'après un théorème de lord Rayleigh, la vitesse de groupe est donnée par les formules

$$\frac{1}{U} = \frac{d\frac{1}{\lambda}}{d\nu} = \frac{d\frac{\nu}{\mathcal{V}}}{d\nu} = \frac{1}{\mathcal{V}_0} \frac{d(n\nu)}{d\nu},$$

\mathcal{V}_0 étant toujours la vitesse de propagation dans le milieu de référence.

Donnons une démonstration du théorème de Rayleigh. Le groupe d'ondes peut se représenter par la formule

$$\Psi = \int_{\nu_0 - \Delta\nu}^{\nu_0 + \Delta\nu} c(\nu) e^{2\pi i[\nu t - \varphi_1(\nu)]} d\nu,$$

ν_0 étant la fréquence « centrale » du groupe et φ_1 la fonction de phase qui dépend de x, y, z et de ν . En posant $\nu = \nu_0 + \eta$, nous pouvons écrire

$$\Psi = e^{2\pi i[\nu_0 t - \varphi_1(\nu_0)]} \int_{-\Delta\nu}^{+\Delta\nu} c(\eta) e^{2\pi i\eta \left[t - \left(\frac{\partial \varphi_1}{\partial \nu} \right)_{\nu_0} \right]} d\eta$$

en négligeant les termes de l'ordre de η^2 . Ceci peut aussi s'écrire

$$\Psi = F \left[t - \left(\frac{\partial \varphi_1}{\partial \nu} \right)_{\nu_0} \right] e^{2\pi i (\nu_0 t - \varphi_1(\nu_0)},$$

ce qui montre que l'onde Ψ est assimilable à une sorte d'onde monochromatique de fréquence ν_0 dont l'amplitude serait fonction de $t - \left(\frac{\partial \varphi_1}{\partial \nu} \right)_{\nu_0}$. Si nous nous déplaçons le long du rayon normal à la surface $\varphi_1 = \text{const.}$ d'une longueur dl en un temps dt , l'amplitude F restera la même si dl et dt sont liés par la relation

$$dt - \frac{\partial}{\partial l} \left(\frac{\partial \varphi_1}{\partial \nu} \right)_{\nu_0} dl = 0.$$

Comme $\frac{\partial \varphi_1}{\partial l}$ est égale à l'inverse de la longueur d'onde, on voit que les valeurs de l'amplitude se déplacent le long du rayon avec une vitesse $\frac{dl}{dt} = U$ telle que

$$U = \frac{\partial}{\partial \nu} \left(\frac{\partial \varphi_1}{\partial l} \right) = \frac{\partial \frac{1}{\lambda}}{\partial \nu}$$

et nous avons bien obtenu la formule de Rayleigh.

Nous voulons maintenant appliquer la théorie des groupes d'ondes à la Mécanique ondulatoire. Nous avons obtenu au paragraphe 4 les relations

$$\lambda = \frac{h}{\sqrt{2m[E - V(x, y, z)]}}, \quad \nu = \frac{E}{h}.$$

Nous en tirons

$$\frac{\partial \frac{1}{\lambda}}{\partial \nu} = \frac{\partial \sqrt{2m(E - V)}}{\partial E} = \frac{m}{\sqrt{2m(E - V)}} = \frac{m}{mc} = \frac{1}{c},$$

d'où, par la formule de Rayleigh,

$$U = c.$$

Nous voyons donc que *la vitesse de groupe des ondes Ψ associées à un corpuscule est égale à la vitesse du mouvement classique du corpuscule.*

Ce théorème est très important parce qu'il permet, ainsi que nous l'expliquerons en détail à la fin du prochain Chapitre, de faire le raccord

entre la nouvelle Mécanique et l'ancienne en considérant un petit groupe d'ondes qui glisse le long d'un rayon ⁽¹⁾.

BIBLIOGRAPHIE.

- [1] LOUIS DE BROGLIE, *Thèse de Doctorat*. Masson, Paris, 1924; *Ann. de Physique*, t. 3, p. 22.
- [2] ERWIN SCHRÖDINGER, *Mémoires sur la Mécanique ondulatoire*, Presses Universitaires, Paris, 1937.
- [3] LOUIS DE BROGLIE, *Introduction à l'étude de Mécanique ondulatoire*, Hermann, Paris, 1930.
- [4] LOUIS DE BROGLIE, *Ondes, corpuscules et Mécanique ondulatoire*, Albin Michel, Paris, 1945.
- [5] LOUIS DE BROGLIE, *Optique électronique et corpusculaire*, Hermann, Paris, 1950.
- [6] J. J. TRILLAT, *Les preuves expérimentales de la Mécanique ondulatoire (Actualités scientifiques)*, Hermann, Paris, fasc. 212).

⁽¹⁾ Sur les vérifications expérimentales de la Mécanique ondulatoire, on pourra consulter les Ouvrages cités en [3], [4] et [6] dans la Bibliographie.

CHAPITRE X.

LA SIGNIFICATION PHYSIQUE DE LA MÉCANIQUE ONDULATOIRE.

1. **Remarques générales.** — Nous avons obtenu les équations fondamentales de la Mécanique ondulatoire du point matériel. Nous devons maintenant chercher ce qu'elles signifient et ce qu'on peut en tirer.

En somme, le passage de la Mécanique ancienne à la Mécanique ondulatoire consiste à prendre comme base de l'étude du mouvement d'un corpuscule, à la place des équations de Newton, l'équation de propagation

$$(1) \quad \Delta\Psi - \frac{8\pi^2 m}{h^2} V(x, y, z, t)\Psi = \frac{4\pi im}{h} \frac{\partial\Psi}{\partial t},$$

du moins quand on peut négliger les corrections de relativité.

Tout état de nos connaissances sur un corpuscule sera donc représenté par une solution Ψ de l'équation (1). Exceptionnellement, cette onde Ψ pourra être une onde plane monochromatique qui, nous l'avons vu, correspond à un état de mouvement rectiligne et uniforme, mais en général la solution représentera un train d'ondes que l'analyse de Fourier permettra de représenter par une superposition d'ondes planes monochromatiques. Un cas particulier sera celui du groupe d'ondes formé par la superposition d'ondes planes monochromatiques de fréquences et de directions de propagation très voisines.

Au sujet de l'équation (1), une remarque s'impose : cette équation contenant le symbole $i = \sqrt{-1}$ est une équation complexe et ses solutions sont essentiellement complexes. La fonction Ψ , étant complexe de nature, ne peut représenter une vibration ayant un sens physique réel. En Optique classique, on considérait la vibration lumineuse comme la vibration d'un milieu et l'on devait donc la représenter par une ou plusieurs fonctions réelles. Souvent, il est vrai, on se servait dans les calculs d'expressions complexes dont les véritables fonctions d'onde étaient

les parties réelles, mais ce n'était là qu'un artifice mathématique ayant pour but de faciliter les calculs et, ceux-ci achevés, on ne retenait que la partie réelle des expressions complexes qui, seule, paraissait avoir un sens physique. Il ne va plus en être de même en Mécanique ondulatoire puisque la fonction Ψ est essentiellement complexe. Mais une grandeur complexe ne peut pas avoir de signification physique directe, c'est-à-dire représenter quelque chose de « mesurable », puisque le résultat d'une mesure est toujours un nombre réel. Nous sommes ainsi amenés à penser que la fonction d'onde Ψ est seulement une grandeur qui peut nous servir à former des quantités réelles qui, elles, auront un sens physique. Ce qui suit va confirmer cette idée. Évidemment, en passant des conceptions anciennes aux conceptions nouvelles, la théorie physique va faire un pas important dans le sens de l'abstraction. Certains désireraient peut-être revenir à des images plus concrètes, mais le développement de la Physique depuis vingt-cinq ans ne paraît nullement confirmer cette espérance.

Nous allons retrouver en Mécanique ondulatoire pour les particules matérielles les idées que nous avons exposées à la fin du Chapitre VI au sujet de la lumière et des photons. Le lien entre le corpuscule et l'onde va encore nous apparaître comme de nature probabiliste, l'onde étant essentiellement une onde de probabilité qui représente les localisations possibles du corpuscule. Chaque fois que l'Optique géométrique ne sera pas valable pour la propagation de l'onde, la notion de trajectoire liée à celle de rayon va disparaître et nous serons obligés, comme pour le photon, de considérer le corpuscule comme une entité susceptible de produire des effets localisés, mais à laquelle nous ne pouvons pas toujours attribuer une localisation bien définie dans l'espace à chaque instant.

Dans la nouvelle Mécanique, il s'agit essentiellement de savoir ce que l'on peut déduire de la connaissance de l'onde Ψ et de sa propagation relativement aux phénomènes corpusculaires observables. Dans l'ancienne conception des corpuscules, nous pouvions déterminer avec une précision en principe illimitée, en pratique limitée seulement par l'imperfection de nos observations, la position et la vitesse d'un corpuscule à un instant donné. Les équations classiques de la Dynamique nous permettaient alors de prévoir, avec une précision d'autant plus grande que l'observation initiale avait été plus précise, à quel endroit et avec quelle vitesse le corpuscule pourrait être observé à un instant ultérieur t . L'hypothèse que l'observation initiale était en principe susceptible d'une précision illimitée et qu'une prévision rigoureuse de l'obser-

vation à l'instant t était possible exprimait le déterminisme mécanique et était conforme au postulat d'une entière prévisibilité des phénomènes naturels. Nous verrons que le développement de la nouvelle Mécanique a conduit à des conceptions très différentes.

2. Principe de localisation ou principe d'interférences. — Quand on a cherché à interpréter la signification physique des ondes Ψ , on a tout de suite été amené à énoncer le principe suivant :

Le carré du module de la fonction complexe Ψ mesure en chaque point et à chaque instant la probabilité pour qu'une observation permette de localiser le corpuscule en ce point à cet instant.

La fonction Ψ peut toujours s'écrire $\Psi = a e^{i\varphi}$, où a et φ sont le module et l'argument de la quantité complexe Ψ et sont, par suite, des fonctions réelles des coordonnées et du temps. En employant le langage de la théorie classique des ondes, nous dirons que a est l'amplitude et a^2 l'intensité de l'onde Ψ . Si nous désignons par Ψ^* la quantité complexe conjuguée de Ψ , c'est-à-dire $a e^{-i\varphi}$, nous aurons

$$a^2 = \Psi\Psi^* = |\Psi|^2;$$

c'est cette quantité qui, d'après le principe énoncé, donne la probabilité de localisation du corpuscule.

Nous devons maintenant expliquer comment on a été amené à adopter le principe énoncé plus haut que l'on peut nommer « *principe des interférences* ». Dans la théorie classique de la lumière, l'intensité de l'onde lumineuse mesure en chaque point la quantité d'énergie lumineuse qu'on peut y recueillir. C'est cette règle qui permet la prévision correcte des phénomènes d'interférences et de diffraction. Mais on sait que l'étude des phénomènes élémentaires d'action entre la matière et le rayonnement a montré que tout se passe comme si toute radiation de fréquence ν était formée de corpuscules (photons) d'énergie $h\nu$. Pour concilier cette structure discontinue de la radiation avec le calcul classique des interférences, nous devons dire que l'intensité de l'onde lumineuse associée à un photon mesure en chaque point et à chaque instant la probabilité pour que le photon produise un effet observable en ce point et à cet instant. C'est bien là le contenu du principe des interférences et nous allons voir qu'il y a bien lieu de l'étendre des photons aux corpuscules matériels.

En effet, un résultat expérimental très important obtenu en 1927 a été la découverte de la diffraction des électrons par les cristaux. Si

l'association des ondes avec les corpuscules exposée dans le dernier chapitre est exacte, à tout électron de vitesse v doit être associée une onde dont la longueur d'onde, en laissant de côté les corrections relativistes, doit être égale à $\frac{h}{mv}$. Cette expression conduit à associer aux électrons usuellement employés dans l'expérience des longueurs d'onde de l'ordre de 10^{-8} à 10^{-9} cm. Cette longueur d'onde est de l'ordre de celle des rayons X : par suite, on pouvait espérer observer avec des électrons des phénomènes analogues à ceux découverts par Laue pour les rayons X. Un faisceau d'électrons de vitesse v étant associé à une onde de longueur d'onde $\frac{h}{mv}$, si ce faisceau tombe sur un cristal, l'onde subira une diffraction prévue par la théorie développée par Laue et Bragg et l'amplitude de l'onde sera maximum dans certaines directions privilégiées déterminées par des formules bien connues. C'est bien ce qu'ont établi les belles expériences de Davisson et Germer, puis celles de G. P. Thomson, de Rupp, de Ponte, de Kikuchi, etc. Ces expériences, complétées en 1940 par les expériences de Börsch mettant en évidence la diffraction des électrons par le bord d'un écran, prouvent que l'intensité de l'onde associée donne bien en chaque point et à chaque instant la probabilité pour que l'électron manifeste sa présence en ce point à cet instant. Le principe des interférences, valable pour les photons, est donc aussi exact pour les électrons et nous sommes fondés à l'adopter d'une façon générale.

Nous allons préciser l'énoncé du principe des interférences. Tout d'abord, l'onde Ψ , solution d'une équation aux dérivées partielles linéaire, n'est définie qu'à une constante multiplicative complexe près et nous pouvons choisir le module de cette constante de façon à avoir

$$(2) \quad \int \Psi^* \Psi \, d\tau = 1,$$

l'intégrale étant étendue à tout l'espace. Nous verrons tout à l'heure que, si cette condition est réalisée à un instant quelconque, elle reste ensuite toujours exacte en vertu même de l'équation de propagation. On dit alors que la fonction Ψ est « normée ».

La fonction d'onde étant normée, nous énoncerons le principe des interférences en disant :

Soit $d\tau$ un élément de volume de l'espace contenant le point de coordonnées x, y, z ; la probabilité pour qu'une observation faite à

L'instant t permette de localiser le corpuscule à l'intérieur de cet élément de volume est

$$\Psi(x, y, z, t) \Psi^*(x, y, z, t) d\tau.$$

La formule de normalisation du Ψ signifie que la probabilité totale de toutes les positions possibles du corpuscule est égale à 1.

Nous sommes maintenant conduits à une image qui peut servir à se représenter la manière dont évoluent au cours du temps les probabilités de localisation du corpuscule. En effet, au cours du temps, la répartition dans l'espace du $\Psi\Psi^*$ se modifie en général et, pour suivre ces modifications, il est commode d'imaginer un fluide fictif dont la densité à chaque instant et en chaque point serait donnée par l'expression

$$\rho(x, y, z, t) = \Psi(x, y, z, t) \Psi^*(x, y, z, t).$$

Par définition, nous attribuerons à ce fluide fictif un mouvement tel que sa vitesse au point x, y, z à l'instant t soit donnée par la formule

$$\mathbf{v} = \frac{1}{\Psi\Psi^*} \frac{h}{4\pi im} (\Psi \mathbf{grad} \Psi^* - \Psi^* \mathbf{grad} \Psi) = \frac{h}{2\pi m} \mathbf{grad} \varphi,$$

φ étant l'argument de Ψ . Nous allons démontrer que le fluide fictif ainsi défini se conserve en se déplaçant, c'est-à-dire que $\int \rho d\tau$ demeure constant au cours du temps. Nous aurons ainsi obtenu la preuve précédemment annoncée que, si la condition de normalisation (2) est satisfaite à un instant initial, elle le reste toujours ensuite.

Pour obtenir la démonstration en question, nous partirons de l'équation de propagation (1) et nous remarquerons que Ψ^* obéit à l'équation conjuguée

$$(1^*) \quad \Delta \Psi^* - \frac{8\pi^2 m}{h^2} \Psi(x, y, z, t) \Psi^* = - \frac{4\pi im}{h} \frac{\partial \Psi^*}{\partial t}.$$

Multiplions (1) par Ψ^* , (1^{*}) par Ψ et retranchons, il vient

$$\Delta \Psi^* \Psi - \Psi \Delta \Psi^* = \frac{4\pi im}{h} \frac{\partial (\Psi \Psi^*)}{\partial t} = \frac{4\pi im}{h} \frac{d\rho}{dt},$$

d'où

$$\frac{d\rho}{dt} = \frac{h}{4\pi im} (\Delta \Psi^* \Psi - \Psi \Delta \Psi^*) = \frac{h}{4\pi im} \sum_{xyz} \frac{\partial}{\partial x} \left(\Psi^* \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial \Psi^*}{\partial x} \right) = -\text{div}(\rho \mathbf{v})$$

grâce aux définitions adoptées pour ρ et pour \mathbf{v} .

Le mouvement du fluide fictif de densité ρ et de vitesse \mathbf{v} obéit donc à l'équation

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) = 0$$

appelée en hydrodynamique « équation de continuité ». Or cette équation signifie qu'il n'y a, au cours du mouvement, ni apparition, ni disparition de fluide, en d'autres termes qu'il y a conservation de la quantité totale du fluide. Nous avons ainsi une image hydrodynamique du mouvement de la « probabilité de présence » du corpuscule.

En réalité, le principe des interférences ne peut pas nous conduire à attribuer au corpuscule qui est associé à une onde Ψ connue, une position bien déterminée. Il nous permet seulement d'évaluer la probabilité pour que, lors d'une localisation par l'observation, on trouve ses coordonnées comprises dans un certain intervalle. Là où la Dynamique classique nous parlait de valeurs bien définies des coordonnées à chaque instant, la nouvelle Mécanique nous fournit seulement des valeurs possibles avec une certaine répartition de probabilité entre ces valeurs à chaque instant. Nous apercevons déjà ici le caractère probabiliste de la nouvelle Mécanique, caractère que nous verrons s'accroître au fur et à mesure que nous approfondirons davantage les conceptions de la nouvelle théorie.

3. Théorème d'Ehrenfest. — En nous servant encore du fluide fictif de probabilité que nous venons d'introduire, nous allons énoncer un théorème dû à P. Ehrenfest dont nous verrons bientôt l'importance.

Donnons d'abord une définition générale. Considérons un train d'ondes Ψ occupant un domaine limité de l'espace et soit f une grandeur définie en chaque point de cette région (f est une fonction de point). Nous appellerons « valeur moyenne de la grandeur f à l'instant t dans l'onde Ψ » la quantité

$$\bar{f} = \int f \Psi \Psi^* dz.$$

Le théorème d'Ehrenfest s'énonce comme suit :

Le centre de gravité du fluide fictif de probabilité attaché à un corpuscule de masse m se déplace dans l'espace comme le ferait, d'après la Dynamique classique, un point matériel de masse m soumis à la valeur moyenne de la force.

Naturellement, nous définissons ici le centre de gravité du fluide fictif de probabilité par ses trois coordonnées

$$x = \int x \Psi \Psi^* d\tau, \quad y = \int y \Psi \Psi^* d\tau, \quad z = \int z \Psi \Psi^* d\tau.$$

En combinant les équations (1) et (1'), on parvient aux équations

$$\begin{aligned} m \frac{d^2 x}{dt^2} &= \int \left(- \frac{\partial V}{\partial x} \right) \Psi \Psi^* d\tau = - \frac{\partial V}{\partial x}, \\ m \frac{d^2 y}{dt^2} &= \int \left(- \frac{\partial V}{\partial y} \right) \Psi \Psi^* d\tau = - \frac{\partial V}{\partial y}, \\ m \frac{d^2 z}{dt^2} &= \int \left(- \frac{\partial V}{\partial z} \right) \Psi \Psi^* d\tau = - \frac{\partial V}{\partial z}. \end{aligned}$$

ce qui exprime le théorème d'Ehrenfest.

4. Principe de décomposition spectrale ou principe de Born. — Nous avons trouvé un principe nous permettant d'évaluer la probabilité de localisation d'un corpuscule quand on connaît son onde Ψ associée. Mais dans l'ancienne Mécanique, on admettait qu'il était possible d'attribuer à chaque instant à un corpuscule non seulement des coordonnées x, y, z , mais aussi un état de mouvement défini par une énergie E et une quantité de mouvement $\mathbf{p} = m\mathbf{v}$. Or jusqu'ici, dans la nouvelle Mécanique, nous n'avons défini qu'une probabilité de position, nous n'avons pas parlé des états de mouvement. Nous allons maintenant chercher à définir une probabilité des états de mouvement pour un corpuscule associé à une onde Ψ connue.

Envisageons d'abord le cas simple de l'absence de champ. Nous avons trouvé en ce cas qu'au mouvement rectiligne et uniforme d'énergie E constante et de quantité de mouvement \mathbf{p} également constante s'opérant dans la direction de cosinus directeurs α, β, γ correspond l'onde

$$\Psi = A e^{\frac{2\pi i}{h} [Et - \sqrt{2mE}(\alpha x + \beta y + \gamma z)]}$$

de fréquence $\frac{E}{h}$ et de longueur d'onde $\frac{h}{mv}$. Cette onde représente donc un état de mouvement bien déterminé, mais elle correspond, d'après le principe des interférences, à une position entièrement indéterminée car, A étant une constante, l'intensité $|\Psi|^2 = |A|^2$ est la même en tous les points de l'espace. Nous commençons alors à apercevoir une différence essentielle entre l'ancienne et la nouvelle Mécanique : tandis que l'ancienne Mécanique permettait toujours d'attribuer au corpuscule à la

fois une position et un état de mouvement bien déterminés, nous voyons déjà qu'en Mécanique ondulatoire, si l'état de mouvement est entièrement connu, la position est complètement indéterminée.

Mais c'est là un cas limite. En général, l'onde Ψ sera formée par une superposition d'ondes planes monochromatiques constituant ce que l'on nomme en Optique classique un « train d'ondes » qui occupe une région limitée de l'espace à chaque instant, les composantes de Fourier se détruisant par interférences en dehors de cette région. Alors l'intensité $\Psi\Psi^*$ n'est différente de zéro que dans la région occupée par le train d'ondes et la position du corpuscule se trouve mieux définie que dans le cas précédent. Mais il y a une contre-partie. Nous avons, en effet, fait correspondre à toute onde plane monochromatique de fréquence ν et de longueur d'onde λ un état de mouvement défini par l'énergie $E = h\nu$ et par les composantes de quantité de mouvement

$$p_x = \alpha \frac{h}{\lambda}, \quad p_y = \beta \frac{h}{\lambda}, \quad p_z = \gamma \frac{h}{\lambda}.$$

Comme un train d'ondes de dimensions finies comporte nécessairement une superposition d'ondes monochromatiques, nous ne pouvons plus attribuer au corpuscule associé à un tel train d'ondes un état de mouvement bien défini. Ainsi, en passant du cas limite de l'onde plane monochromatique au cas du train d'ondes de dimensions finies, nous voyons apparaître une incertitude sur l'état de mouvement.

Nous pouvons encore considérer le cas limite opposé à celui de l'onde plane, celui d'un « train d'ondes ponctuel », c'est-à-dire d'un train d'ondes de dimensions infiniment petites. En ce cas, la représentation de Fourier du train d'ondes comporte, ou le démontre, la superposition d'une infinité d'ondes planes monochromatiques ayant toutes les fréquences et toutes les directions de propagation. Alors, tandis que la position du corpuscule est parfaitement déterminée puisqu'on est certain de le voir se localiser au point occupé par le train d'ondes, par contre l'état de mouvement est entièrement indéterminé, toutes les grandeurs et toutes les orientations de la quantité de mouvement étant également probables. Donc mieux la position est définie, plus grande est l'incertitude sur l'état de mouvement et inversement.

Précisons que, dans le cas d'un groupe d'ondes dont la longueur d'onde est à peu près bien définie, l'extension du groupe d'ondes dans l'espace est toujours grande par rapport à la longueur d'onde. Donc un corpuscule dont l'état de mouvement est à *peu près* bien défini est toujours non localisé dans l'espace à l'intérieur d'un domaine dont les

dimensions sont grandes par rapport à la longueur d'onde qui correspond à sa quantité de mouvement.

Les considérations qui précèdent vont nous conduire aux relations d'incertitude d'Heisenberg, mais, avant d'aborder cette question, il nous faut d'abord remarquer que, dans les raisonnements précédents, nous avons implicitement admis, à côté du principe des interférences, un deuxième principe que nous appellerons « le principe de décomposition spectrale » ou « principe de Born » et qu'il est maintenant nécessaire d'explicitier clairement. Nous bornant toujours pour l'instant au cas de l'absence de champ, nous énoncerons ce principe comme il suit :

Si l'onde Ψ associée au corpuscule est formée par une superposition d'ondes planes monochromatiques, à chacune de ces ondes monochromatiques correspond un état de mouvement possible du corpuscule, c'est-à-dire qu'une observation faite sur le corpuscule peut conduire à lui attribuer l'un de ces états de mouvement.

Mais on peut donner à ce principe un énoncé plus précis qui est dû à M. Born. Le voici :

Si l'onde Ψ associée à un corpuscule est une somme d'un nombre fini d'ondes planes monochromatiques formant un spectre discontinu, c'est-à-dire si l'on a

$$\Psi = \sum_{\mathbf{E}, \alpha, \beta} a(\mathbf{E}, \alpha, \beta) e^{\frac{2\pi i}{h} [Et - \sqrt{2mE}(\alpha x + \beta y + \sqrt{1 - \alpha^2 - \beta^2} z)]},$$

la probabilité pour qu'une mesure conduite à attribuer au corpuscule le mouvement d'énergie E dans la direction α, β, γ est

$$a(\mathbf{E}, \alpha, \beta) a^*(\mathbf{E}, \alpha, \beta) = |a(\mathbf{E}, \alpha, \beta)|^2.$$

Si l'onde Ψ est la somme d'un nombre infini d'ondes monochromatiques formant un spectre continu, c'est-à-dire si l'on a

$$\Psi = \iiint a(\mathbf{E}, \alpha, \beta) e^{\frac{2\pi i}{h} [Et - \sqrt{2mE}(\alpha x + \beta y + \sqrt{1 - \alpha^2 - \beta^2} z)]} d\mathbf{E} d\alpha d\beta,$$

la probabilité pour qu'une observation conduite à attribuer au corpuscule un mouvement d'énergie comprise entre E et $E + \Delta E$ dans

une direction correspondant aux intervalles angulaires $\alpha \rightarrow \alpha + \Delta\alpha$, $\beta \rightarrow \beta + \Delta\beta$ est égale à

$$\int_{\Delta E} \int_{\Delta\alpha} \int_{\Delta\beta} a(E, \alpha, \beta) a^*(E, \alpha, \beta) dE d\alpha d\beta \\ = \int_{\Delta E} \int_{\Delta\alpha} \int_{\Delta\beta} |a(E, \alpha, \beta)|^2 dE d\alpha d\beta.$$

Plus brièvement, on peut dire que l'intensité de chaque composante spectrale mesure la probabilité de l'état de mouvement correspondant. On peut remarquer que ce principe est la transposition de celui qu'on admet en Optique suivant lequel l'intensité de l'onde monochromatique qu'un prisme ou un réseau permet d'extraire d'une onde incidente spectralement complexe est égale au carré de l'amplitude de la composante correspondante dans le développement de Fourier de l'onde incidente.

Nous venons d'énoncer le principe de décomposition spectrale dans le cas de l'absence de champ. Pour obtenir un énoncé plus complet, il est nécessaire de se placer à un point de vue plus général : on obtient alors une théorie générale précisant dans son ensemble l'interprétation probabiliste de la Mécanique ondulatoire et permettant de trouver la probabilité des valeurs possibles d'une grandeur mécanique quelconque attachée au corpuscule quand on connaît la forme de son onde Ψ . C'est ce que nous verrons au chapitre XIII.

§. Les relations d'incertitude d'Heisenberg. — La comparaison des deux principes que nous venons d'énoncer conduit aux relations d'incertitude d'Heisenberg.

Pour y parvenir, il suffit d'examiner d'un peu plus près la représentation d'un train d'ondes limité, d'un « paquet d'ondes » comme dit Heisenberg, par une superposition d'ondes planes monochromatiques. Posons

$$\nu = \frac{E}{h}, \quad \mu_x = \frac{p_x}{h} = \frac{\alpha \sqrt{2mE}}{h}, \\ \mu_y = \frac{p_y}{h} = \frac{\beta \sqrt{2mE}}{h}, \quad \mu_z = \frac{p_z}{h} = \frac{\gamma \sqrt{2mE}}{h};$$

l'onde monochromatique caractérisée par les constantes E , α , β peut s'écrire

$$\Psi = A e^{2\pi i(\nu t - \mu_x x - \mu_y y - \mu_z z)},$$

Un « paquet d'ondes » sera donc représenté par une expression de la forme

$$\Psi = \iiint a(\mu_x, \mu_y, \mu_z) e^{2\pi i \nu (\mu_x x + \mu_y y + \mu_z z)} d\mu_x d\mu_y d\mu_z,$$

formule dans laquelle ν est considérée comme fonction de μ_x, μ_y, μ_z , car on a

$$h\nu = E = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) = \frac{h^2}{2m} (\mu_x^2 + \mu_y^2 + \mu_z^2).$$

Les coefficients $a(\mu_x, \mu_y, \mu_z)$ sont en général complexes, car les diverses ondes monochromatiques du développement de Fourier présentent, en général, entre elles des différences de phase.

Le train d'ondes limité est engendré par la superposition d'ondes monochromatiques de telle façon qu'elles se détruisent par interférences en dehors de la région occupée par le train d'ondes. En étudiant la manière dont ceci peut se produire, on arrive, par des raisonnements que je ne reproduirai pas, à la conclusion suivante : si $\Delta x, \Delta y, \Delta z$ désignent les dimensions du train d'ondes le long des trois axes rectangulaires, les ondes qui constituent le train d'ondes occupent nécessairement un domaine spectral défini par des intervalles $\Delta\mu_x, \Delta\mu_y, \Delta\mu_z$ tels que

$$\Delta x \Delta\mu_x \geq \eta, \quad \Delta y \Delta\mu_y \geq \eta, \quad \Delta z \Delta\mu_z \geq \eta,$$

η étant une quantité de l'ordre de l'unité. D'après les définitions de μ_x, μ_y, μ_z , on peut donc aussi écrire en ordre de grandeur

$$\Delta x \Delta p_x \geq h, \quad \Delta y \Delta p_y \geq h, \quad \Delta z \Delta p_z \geq h.$$

Telles sont les relations d'incertitude d'Heisenberg. Elles montrent qu'on ne peut jamais attribuer au corpuscule à un même instant une position et un état de mouvement parfaitement déterminés. Plus l'une des coordonnées est connue avec précision, plus la valeur de la composante conjuguée de quantité de mouvement est incertaine et inversement. Nous retrouvons ici sous une forme précise les considérations déjà esquissées au début du paragraphe 4.

6. Évolution de l'onde Ψ . Rôle de la mesure. — Tandis que la Mécanique classique considérait le corpuscule comme un point bien localisé décrivant une trajectoire linéaire et qu'elle attribuait à tout instant à toutes les grandeurs caractérisant ce mouvement une valeur parfaitement déterminée, nous n'avons plus à notre disposition pour décrire le

corpuscule et ses manifestations observables qu'une fonction d'onde en principe complexe et, en général, différente de zéro dans une région finie de l'espace. C'est seulement dans le cas limite des phénomènes macroscopiques, quand certaines conditions sont réalisées, que l'on peut approximativement retomber sur le schéma de la Mécanique classique (*voir* le § 9 du présent Chapitre).

Il résulte des considérations précédentes que l'onde Ψ symbolise les diverses valeurs possibles des grandeurs attachées au corpuscule avec leurs probabilités respectives. Ainsi les valeurs possibles des coordonnées du corpuscule sont celles qui correspondent aux points de la région de l'espace occupée par l'onde Ψ et les valeurs de $\Psi\Psi^*$ en chaque point mesurent les probabilités des diverses localisations possibles. De même, les valeurs de l'énergie figurant dans la décomposition spectrale de l'onde Ψ sont les valeurs qu'une observation de mesure peut amener à attribuer à l'énergie et les quantités $|\alpha(E, \alpha, \beta)|^2$ nous donnent les probabilités respectives de ces valeurs possibles. Chaque fois qu'une observation comportant mesure nous apporte à un instant t_0 un renseignement précis sur la position ou le mouvement d'un corpuscule, les probabilités se trouvent brusquement modifiées et, par suite, après la mesure, nous devons représenter l'état de nos connaissances par une nouvelle forme de l'onde Ψ : c'est ce que M. Heisenberg a appelé « la réduction du paquet de probabilité ». Ainsi que MM. Bohr et Heisenberg l'ont montré dans de fines analyses dont nous parlerons plus loin, chaque opération de mesure, chaque dispositif d'observation précise introduit, en raison même de l'existence du quantum d'action, une perturbation incontrôlable, de sorte qu'après la mesure il subsiste toujours des incertitudes qui satisfont aux relations d'Heisenberg.

Après une mesure qui a fourni tous les renseignements compatibles avec les incertitudes d'Heisenberg, l'état des probabilités est représenté par une nouvelle forme de l'onde Ψ que nous devons prendre comme donnée initiale. L'équation de propagation, qui est du premier ordre, permet alors de calculer l'évolution ultérieure de l'onde Ψ à partir de cette forme initiale et de dire quelle sera à une époque t postérieure à t_0 la probabilité pour qu'une observation fournisse telle ou telle valeur pour une certaine grandeur attachée au corpuscule. Ainsi il y aura un déterminisme de l'évolution de la probabilité, mais cette évolution sera brusquement modifiée par toute nouvelle mesure provoquant une nouvelle réduction du paquet de probabilité. Mais s'il y a encore déterminisme de l'évolution des probabilités entre les mesures, il n'y a plus déterminisme strict des phénomènes observables eux-mêmes et il faut

abandonner l'idée traditionnelle de la stricte prévisibilité de tous les phénomènes naturels.

Revenons aux relations d'incertitude. En Mécanique ondulatoire, nous admettons la possibilité de représenter l'état de nos connaissances sur un corpuscule après une mesure par une certaine onde Ψ . Si ce postulat est exact, une mesure faite avec la plus grande précision possible ne doit pouvoir nous donner sur les valeurs des coordonnées et des composantes de la quantité de mouvement que des indications compatibles avec les relations d'incertitude d'Heisenberg. En d'autres termes, une mesure de la position doit introduire nécessairement sur la valeur de la quantité de mouvement après la mesure une incertitude d'autant plus grande que la mesure de position est plus précise et, inversement, une mesure de quantité de mouvement doit introduire sur la position une incertitude d'autant plus grande que la mesure est plus précise. En effet, en analysant avec soin les expériences de mesure possibles, Bohr et Heisenberg ont montré qu'il en est bien ainsi et leurs raisonnements ont l'intérêt de prouver que ceci résulte de l'existence du quantum d'action. Des nombreux exemples qui ont été discutés, nous en résumerons seulement deux.

7. Le microscope d'Heisenberg. — Heisenberg a imaginé qu'on examine à l'aide d'un microscope un électron qui se déplace sur le porte-objet et qui est éclairé par en bas au moyen d'une radiation monochromatique de fréquence ν (*fig. 23*).

Nous supposons le microscope corrigé de toutes les aberrations, mais il subsiste nécessairement les effets de diffraction. Ces effets de diffraction limitent le pouvoir séparateur du microscope : ils ont pour conséquence que, si un photon diffusé par l'électron placé en O vient percer le plan image π en un point P, l'impression produite en P par l'arrivée de ce photon ne nous permet de déduire la position de l'électron sur l'axe des x qu'avec une incertitude au moins égale à

$$\delta x = \frac{\lambda}{2 \sin \varepsilon}.$$

C'est la formule bien connue donnant le pouvoir séparateur d'un microscope. Nous pouvons essayer de diminuer cette incertitude en employant des faisceaux éclairants de plus petite longueur d'onde λ . Mais, et c'est ici que le quantum d'action intervient d'une façon essentielle, à la longueur d'onde λ correspond la quantité de mouvement $\frac{h}{\lambda}$

pour le photon incident qui échange de la quantité de mouvement avec l'électron au moment de la diffusion. Plus nous prendrons une lumière de courte longueur d'onde pour éclairer l'électron, plus la quantité de mouvement cédée par le photon à l'électron au moment de la diffusion pourra devenir grande et plus, par suite, sera grande l'incertitude sur la valeur finale de la quantité de mouvement de l'électron.

Précisons ceci par le calcul. La diffusion s'effectue par un échange d'énergie et d'impulsion entre le photon et l'électron, par effet Compton. Si v_x est la vitesse de l'électron le long de l'axe des x avant la diffusion et v'_x sa vitesse après la diffusion et si, de plus, ν' est la fréquence du photon diffusé, nous avons en écrivant la conservation de la quantité de mouvement totale projetée sur Ox

$$mv_x = \frac{h\nu'}{c} \alpha + mv'_x,$$

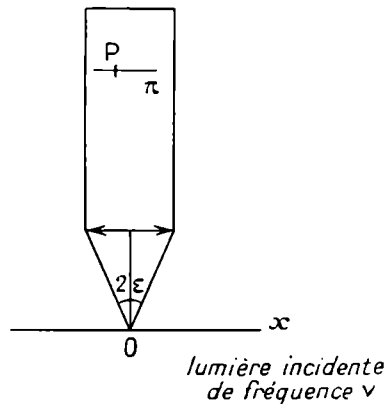


Fig. 23.

α étant le cosinus directeur de la direction du photon après diffusion par rapport à Ox . Comme il est facile de montrer que ν' diffère très peu de ν , nous avons approximativement

$$mv_x - mv'_x = \alpha \frac{h\nu}{c} = \alpha \frac{h}{\lambda}.$$

Or le photon diffusé, s'il est recueilli dans le microscope, doit être dévié de moins de ε de sa direction initiale. Les valeurs maxima et minima de α sont donc $\pm \sin \varepsilon$ et l'incertitude sur la quantité de mouvement de l'électron dans le sens Ox après la diffusion est donc au minimum (en supposant parfaitement connue la valeur initiale v_x)

$$\delta p_x = 2 \sin \varepsilon \frac{h}{\lambda}.$$

D'où

$$\delta x \delta p_{x'} = \frac{\lambda}{2 \sin \varepsilon} 2 \sin \varepsilon \frac{h}{\lambda} = h.$$

C'est là un optimum et, si la mesure n'est pas réalisée avec le maximum de précision, on aura

$$\delta x \delta p_{x'} > h.$$

Nous retrouvons la relation d'incertitude d'Heisenberg et nous voyons bien ici son origine physique : elle provient de l'impossibilité de diminuer λ sans augmenter la valeur $\frac{h}{\lambda}$ de la quantité de mouvement du photon, c'est-à-dire le quantum de quantité de mouvement mis en jeu dans l'interaction de diffusion.

8. Passage d'un corpuscule à travers une fente percée dans un écran.

— Comme deuxième exemple, nous prendrons la détermination de la position d'un corpuscule par son passage à travers une ouverture percée dans un écran plan. Pour définir avec précision les coordonnées du

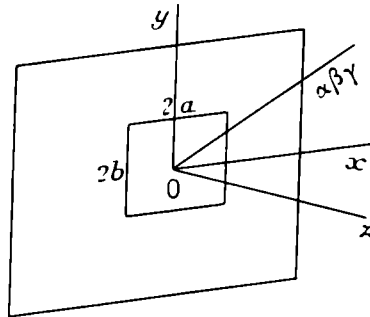


Fig. 24.

corpuscule dans le plan de l'écran quand il traverse l'ouverture, nous sommes évidemment amenés à prendre une ouverture très petite, mais plus nous prendrons l'ouverture petite, plus nous augmenterons les phénomènes de diffraction qui, d'après les idées de la Mécanique ondulatoire, doivent résulter du passage de l'onde associée à travers l'écran.

Pour développer les calculs, prenons le cas simple où le corpuscule incident tombe sur l'écran sous l'incidence normale et où l'ouverture percée dans l'écran est un rectangle. Nous prendrons le centre de cette ouverture comme origine des coordonnées, l'axe Oz étant normal au plan de l'écran et les côtés du rectangle parallèles à Ox et à Oy ayant respectivement pour longueur $2a$ et $2b$.

L'onde Ψ associée au corpuscule incident est de la forme

$$\Psi = A e^{2\pi i \left(\nu t - \frac{z}{\lambda} \right)}.$$

Le passage à travers l'ouverture rectangulaire la transforme, comme nous l'apprend la théorie de la diffraction, en un groupe d'ondes de la forme

$$\Psi = \sum_{\alpha\beta} A(\alpha, \beta) e^{2\pi i \left(\nu t - \frac{\alpha x + \beta y + \gamma z}{\lambda} \right)}, \quad \text{avec } \alpha^2 + \beta^2 + \gamma^2 = 1.$$

Les amplitudes partielles $A(\alpha, \beta)$ présentent en fonction de α et β des maxima et des minima bien connus dans la théorie de la diffraction en Optique. L'intensité des maxima successifs va rapidement en décroissant et les formules classiques de la théorie de la diffraction montrent que l'étendue de la diffraction à partir de Oz est mesurée dans le sens de variation de z et dans le sens de variation de β par

$$\delta\alpha \cong \frac{\lambda}{2a}, \quad \delta\beta \cong \frac{\lambda}{2b}.$$

Si nous posons, comme précédemment,

$$\mu_x = \frac{\alpha}{\lambda}, \quad \mu_y = \frac{\beta}{\lambda},$$

nous aurons

$$\delta\mu_x = \frac{\delta\alpha}{\lambda} \cong \frac{1}{2a}, \quad \delta\mu_y = \frac{\delta\beta}{\lambda} \cong \frac{1}{2b}.$$

Or la position du corpuscule, lors de son passage dans l'ouverture rectangulaire, est définie avec des incertitudes δx et δy qui ont évidemment pour valeurs

$$\delta x = 2a, \quad \delta y = 2b.$$

On a donc finalement

$$\delta p_x \delta x = h \delta\mu_x \delta x \cong h, \quad \delta p_y \delta y = h \delta\mu_y \delta y \cong h.$$

Cet exemple montre très clairement comment le caractère ondulatoire des corpuscules a pour conséquence d'empêcher une mesure simultanée précise des coordonnées et des quantités de mouvement.

De nombreux autres exemples ont été discutés en détail et il ne paraît pas douteux que la mesure simultanée d'une coordonnée et de la composante conjuguée de la quantité de mouvement fasse toujours apparaître les incertitudes d'Heisenberg.

9. **Le raccord avec la Mécanique classique.** — Il est essentiel de comprendre pourquoi les corpuscules et les corps plus pesants, quand on observe leur mouvement à notre échelle macroscopique, paraissent suivre les lois de la Mécanique classique. Les corpuscules matériels (et *a fortiori* les corps de plus grande masse) ont toujours dans les conditions usuelles des longueurs d'onde très petites, inférieures à 10^{-8} cm. Il en résulte deux conséquences importantes. D'abord dans les phénomènes mécaniques observables à notre échelle, les champs sont toujours sensiblement constants dans des domaines de l'ordre de la longueur d'onde; ensuite nous pouvons concevoir une région de l'espace qui contienne un grand nombre de longueurs d'onde et qui ait cependant des dimensions inférieures à ce que nous pouvons mesurer directement. Une observation du corpuscule peut donc alors nous permettre de représenter l'état du corpuscule par un groupe d'ondes presque monochromatique de dimensions négligeables à notre échelle et comprenant néanmoins un grand nombre de longueurs d'onde. Nous pouvons donc dire qu'il est possible d'attribuer au corpuscule un mouvement bien défini et, en même temps, de bien le localiser à notre échelle. Le fluide de probabilité forme en ce cas une sorte de petit globule de dimensions négligeables à l'intérieur duquel la force-grad V peut être considérée comme constante. Puisque le globule se confond sensiblement avec son centre de gravité, il résulte alors du théorème d'Ehrenfest que ce globule se meut comme un point matériel ayant la masse du corpuscule et obéissant aux lois de la Mécanique classique. Or le corpuscule ne peut manifester sa présence qu'en un des points intérieurs au globule et, comme ces points sont pratiquement indiscernables, tout se passe donc comme si le corpuscule se déplaçait suivant les lois de la Mécanique classique. A l'échelle macroscopique, à cause de l'imperfection de nos sens, nous aurons l'illusion qu'il est possible de définir à la fois la position et le mouvement du corpuscule et ceci explique pourquoi dans les expériences macroscopiques où nous observons le mouvement d'un corpuscule (par exemple l'électron dans les expériences du type Villard ou dans la chambre de Wilson), les lois de la Mécanique classique sont pratiquement valables.

Nous pouvons encore présenter ces idées en nous figurant la propagation des ondes Ψ dans le champ de force. Nous sommes ici dans le cas où l'approximation de l'Optique géométrique est valable et nous pouvons employer la notion de rayon. Nous pouvons donc envisager un pinceau de rayons extrêmement délié et un groupe d'ondes de dimensions très supérieures à la longueur d'onde et cependant sensiblement

ponctuel pour nos observations macroscopiques qui décrivent le pinceau de rayons.

D'après le théorème sur la vitesse de groupe, le petit groupe d'ondes se déplace le long du pinceau de rayons avec précisément la vitesse que la Mécanique classique attribuerait au point matériel qui décrirait un des rayons du pinceau assimilé à une trajectoire. On voit donc que tout se passe à notre échelle comme si le corpuscule décrivait avec la vitesse classique l'un des rayons-trajectoires et cette image est particulièrement intuitive.

Les considérations qui précèdent permettent d'effectuer le raccord

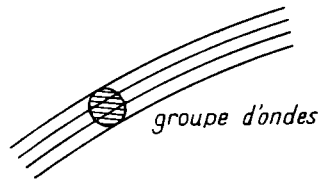


Fig. 25.

entre les conceptions de l'ancienne Mécanique et celles de la nouvelle et d'expliquer le succès de la Mécanique classique pour l'explication des phénomènes macroscopiques.

BIBLIOGRAPHIE.

- [1] LOUIS DE BROGLIE, *Introduction à l'étude de la Mécanique ondulatoire*, Hermann, Paris, 1930.
- [2] W. HEISENBERG, *Principes de la Mécanique quantique*, Gauthier-Villars, Paris, 1932.
- [3] LOUIS DE BROGLIE, *Ondes, corpuscules et Mécanique ondulatoire*, Albin Michel, Paris, 1945.
- [4] GEORGES GUINIER, *Éléments de Physique théorique moderne*, Bordas, Paris, 1949.
- [5] DAVID BOHM, *Quantum theory*, Prentice Hall, New-York, 1951.

CHAPITRE XI.

APPLICATION DE LA MÉCANIQUE ONDULATOIRE A LA QUANTIFICATION.

1. **Première interprétation de la quantification par la Mécanique ondulatoire.** — Nous avons vu précédemment Planck introduire tant bien que mal dans le cadre des théories classiques l'idée de quantification : nous avons étudié la forme des règles de quantification dans l'ancienne théorie des Quanta et constaté leurs fréquents succès. Dès son apparition [1], la Mécanique ondulatoire a permis d'apercevoir le véritable sens des règles de quantification.

Considérons un corpuscule qui décrit une trajectoire fermée dans un champ permanent : l'ancienne règle de quantification peut s'énoncer en disant que l'intégrale d'action maupertuisienne

$$\int mv ds = \int \sum_1^3 p_k dq_k$$

prise le long de la trajectoire fermée est égale à un multiple entier de h , ce qui peut s'écrire

$$(1) \quad \frac{1}{h} \oint dS_1 = n \quad (n \text{ entier}).$$

S_1 étant la fonction de Jacobi raccourcie définie précédemment. Or, d'après la Mécanique ondulatoire, lorsque l'Optique géométrique est valable pour la propagation de l'onde associée au corpuscule, celle-ci a la forme

$$\Psi = a(x, y, z) e^{\frac{2\pi i}{h} [Et - S_1(x, y, z)]}.$$

La condition (1) exprime alors que la fonction d'onde Ψ est *uniforme* le long de la trajectoire, c'est-à-dire qu'elle reprend la même valeur

quand on a décrit la trajectoire entière. La condition de quantification apparaît alors comme une condition de résonance, faisant intervenir un nombre entier comme cela est usuel pour une condition de ce genre. Ainsi se trouve expliquée l'apparition des nombres entiers dans les formules de quantification, apparition que les idées de la Mécanique classique ne permettaient aucunement de comprendre. Les conditions de quanta de Wilson-Sommerfeld pour les mouvements quasipériodiques peuvent s'interpréter d'une façon analogue.

L'interprétation des anciennes conditions de quantification que nous venons de donner explique bien la nature profonde de ces conditions et fait comprendre pourquoi, dans le cadre de la Mécanique classique encore utilisé par l'ancienne théorie des Quanta, on devrait être amené à les adopter. Mais elles sont liées à l'emploi de la Mécanique classique, c'est-à-dire, nous le savons maintenant, à la validité, pour la propagation de l'onde Ψ , de l'approximation de l'Optique géométrique. Or, quand on considère la quantification des systèmes à l'échelle atomique, les états quantifiés à petits nombres de quanta correspondent à des conditions où l'Optique géométrique n'est plus valable et ce sont justement ces états qui sont les plus intéressants à étudier parce qu'ils sont les plus fréquemment réalisés dans les systèmes quantifiés de l'échelle atomique.

Pour nous rendre compte de la non-validité de l'Optique géométrique dans le cas des états à petits nombres quantiques, rappelons d'abord les relations

$$S_1 = h z_1, \quad \sum_{xyz} \left(\frac{\partial z_1}{\partial x} \right)^2 = \frac{1}{\lambda^2}$$

qui nous donnent

$$\lambda = \frac{1}{h \sqrt{\sum_{xyz} \left(\frac{\partial S_1}{\partial x} \right)^2}} = \frac{h}{|\text{grad } S_1|} = \frac{h}{\frac{\partial S_1}{\partial s}}$$

en appelant s la variable comptée le long de la trajectoire normalement aux surfaces $S_1 = \text{const}$. La condition de quantification (1) s'écrit alors

$$n = \frac{1}{h} \oint dS_1 = \frac{1}{h} \oint \frac{\partial S_1}{\partial s} ds = \oint \frac{ds}{\lambda} = L \frac{1}{\bar{\lambda}} = \frac{L}{\bar{\lambda}},$$

L désignant la longueur totale de la trajectoire et $\bar{\lambda}$ une sorte de longueur d'onde moyenne le long de la trajectoire définie par les relations précédentes. On a donc

$$L = n \bar{\lambda}$$

et, quand n est petit, on voit que la longueur totale de la trajectoire fermée est de l'ordre de $\bar{\lambda}$. Comme le champ varie nécessairement beaucoup dans le domaine occupé par la trajectoire fermée, on voit que dans ce domaine l'indice n relatif aux ondes Ψ va varier beaucoup et què, par suite, l'approximation de l'Optique géométrique n'y sera pas valable pour la propagation de ces ondes. Dans ce cas, le plus intéressant, où n est petit, la justification donnée plus haut des anciennes formules de quantification n'est plus valable.

M. Schrödinger qui a, le premier, attiré l'attention sur ce fait, en a conclu que, bien que les anciennes conditions de quanta reçoivent de la Mécanique ondulatoire une interprétation intuitive simple, il est cependant nécessaire de reprendre en Mécanique ondulatoire toute la théorie de la quantification sur une base nouvelle.

2. La théorie de la quantification en Mécanique ondulatoire. — En Mécanique ondulatoire, on doit toujours partir de l'équation de propagation des ondes Ψ qui remplace entièrement les équations du mouvement de la Mécanique classique. Dans le cas des systèmes quantifiables qui sont toujours conservatifs, nous pouvons écrire l'équation d'ondes sous la forme de Schrödinger

$$\Delta\Psi + \frac{8\pi^2 m}{h^2} [E - V(x, y, z)]\Psi = 0$$

et nous ferons correspondre aux états stationnaires au sens de Bohr des ondes stationnaires de la forme

$$\Psi = a(x, y, z) e^{\frac{2\pi i}{h} Et},$$

la fonction $a(x, y, z)$ généralement complexe étant solution de l'équation

$$(2) \quad \Delta a + \frac{8\pi^2 m}{h^2} [E - V(x, y, z)] a = 0.$$

Conformément à la méthode employée dans la théorie générale des ondes pour la détermination des ondes stationnaires, la nouvelle méthode de quantification consistera à chercher pour quelles valeurs de la constante E , l'équation précédente admet une valeur $a(x, y, z)$ finie, uniforme, continue et satisfaisant à certaines conditions aux limites, ce qui fournira les fréquences $\frac{E}{h}$ des ondes Ψ stationnaires qui

sont possibles dans le système atomique et qui correspondront aux états quantifiés.

Cette détermination des « vibrations propres » est un problème bien connu en Physique classique. Donnons-en quelques exemples.

Soit une corde élastique homogène tendue le long de l'axe des x et fixée aux points extrêmes $x = 0$ et $x = l$. Les vibrations transversales dont la corde est susceptible doivent satisfaire à l'équation de d'Alembert

$$(3) \quad \frac{1}{V^2} \frac{\partial^2 \varphi}{\partial t^2} = \frac{\partial^2 \varphi}{\partial x^2},$$

la fonction $\varphi(x, t)$ représentant l'élongation d'un point de la corde dans le sens perpendiculaire à sa longueur et V une constante dépendant de l'élasticité et de la masse spécifique de la corde vibrante. Pour trouver les vibrations propres de la corde, on considérera une solution monochromatique stationnaire de la forme

$$\varphi(x, t) = a(x) e^{2\pi i \nu t}, \quad \frac{\partial^2 \varphi}{\partial t^2} = -4\pi^2 \nu^2 \varphi$$

et l'équation (3) prendra la forme

$$\frac{d^2 a}{dx^2} + \frac{4\pi^2 \nu^2}{V^2} a = 0.$$

Le problème de la détermination des vibrations propres nous amène à chercher pour quelle valeur de la constante ν , il existe une solution finie, uniforme et continue de l'équation précédente qui s'annule en $x = 0$ et en $x = l$. En posant $\lambda = \frac{V}{\nu}$, on trouve que les fréquences propres ν_n doivent être telles que

$$\lambda_n = \frac{V}{\nu_n} = \frac{2l}{n} \quad \text{ou} \quad l = n \frac{\lambda_n}{2}$$

la solution correspondante étant

$$a_n(x) = \sin \frac{2\pi x}{\lambda_n} = \sin \frac{n\pi x}{l}, \quad \varphi_n(x, t) = C \sin \frac{n\pi x}{l} e^{2\pi i \nu_n t}.$$

Les fonctions a_n (ou φ_n) possèdent des propriétés importantes que l'on retrouve dans tous les problèmes plus compliqués du même genre. D'abord l'on a

$$\int_0^l a_n a_m dx = 0 \quad \text{pour } n \neq m$$

comme on le vérifie aisément. Cette relation exprime que les « fonctions propres » a_n sont « orthogonales » entre elles. De plus, un état station-

naire quelconque $\varphi(x, t)$ de la corde vibrante peut se représenter par une superposition de vibrations propres sous la forme

$$\varphi(x, t) = \sum_n c_n \varphi_n(x, t).$$

les c_n étant des coefficients constants convenablement choisis. Les a_n et φ_n étant dans cet exemple simple des fonctions trigonométriques de x , les propriétés précédentes résultent immédiatement de la théorie des séries de Fourier.

Un problème un peu plus compliqué est celui des vibrations d'une membrane élastique homogène tendue sur un cadre plan de forme donnée (carré, rectangle, cercle, etc.) et fixée à ce cadre sur son pourtour. L'équation de d'Alembert prend alors la forme

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} = \frac{1}{V^2} \frac{\partial^2 \varphi}{\partial t^2},$$

$\varphi(x, y, t)$ étant le déplacement d'un point de la membrane vibrante normalement à son plan d'équilibre et V une constante qui dépend de l'élasticité et de la masse spécifique de la membrane. Pour une solution monochromatique stationnaire

$$\varphi(x, y, t) = a(x, y) e^{2\pi i \nu t},$$

on aura

$$\frac{\partial^2 a}{\partial x^2} + \frac{\partial^2 a}{\partial y^2} + \frac{4\pi^2 \nu^2}{V^2} a = 0$$

et la détermination des vibrations propres de la membrane consistera à rechercher les valeurs de ν pour lesquelles l'équation précédente admet une solution $a(x, y)$ finie, continue et uniforme dans le domaine des variations de x et de y et s'annulant aux limites de ce domaine sur le pourtour du cadre. On constate que les fonctions propres (qui dépendent naturellement de la forme du cadre) sont orthogonales entre elles et que toute vibration de la membrane peut se représenter par une superposition, avec coefficients convenablement choisis, de vibrations propres.

On peut envisager en Physique classique des problèmes plus compliqués en imposant à la fonction d'ondes d'autres conditions que de s'annuler aux limites, par exemple en imposant à sa dérivée normale d'être nulle aux limites. On peut aussi supposer que la vitesse V varie à l'intérieur du domaine. Tous ces problèmes sont classiques en Élasticité, en Optique, en Hydrodynamique et en Électromagnétisme, notamment dans la théorie des antennes de T. S. F.

Revenons maintenant à la Mécanique ondulatoire et à l'équation (2). Nous avons pour la détermination des ondes stationnaires un problème analogue à ceux que nous venons de traiter, mais rendu compliqué par le fait que le coefficient de $a(x, y, z)$ dans (2) est une fonction de x, y, z . On a ici l'analogue à trois dimensions du cas des cordes ou membranes non homogènes où la vitesse V varie d'un point à un autre du domaine. Il est cependant facile de voir comment on doit définir le problème de la détermination des fréquences propres et des vibrations propres dans le cas de l'équation (2). Tout d'abord, il est naturel d'admettre, comme dans les problèmes analogues de la théorie des ondes, que la fonction a doit être continue et uniforme. Il faut aussi qu'elle soit finie dans tout le domaine occupé par l'onde pour que la densité de probabilité de présence $\Psi\Psi^* = |a|^2$ soit finie en tout point. Enfin si l'onde ne peut occuper qu'une région finie de l'espace hors de laquelle le corpuscule ne peut se trouver, nous admettrons que a doit s'annuler aux limites de ce domaine de façon que $\Psi\Psi^*$ soit nul en dehors; si, au contraire, l'onde peut occuper l'espace entier, il faudra que a décroisse suffisamment rapidement à l'infini pour que l'intégrale de normalisation $\int |\Psi|^2 d\tau = \int |a|^2 d\tau$ soit convergente.

Tel est l'ensemble des conditions que nous imposons à l'onde Ψ stationnaire et qui doivent nous permettre de calculer les énergies des états quantifiés.

3. Définition générale et propriétés des valeurs et fonctions propres. — Pour mieux dominer la question, nous allons donner la définition générale et les propriétés d'une équation du type (2).

On peut démontrer qu'une équation aux dérivées partielles du type (2) ne peut admettre, dans le domaine D de variation des variables, de solution finie, uniforme, continue et nulle aux limites de D que pour certaines valeurs $E_1, E_2, \dots, E_i, \dots$ de la constante E . Ces valeurs particulières sont appelées les « valeurs propres » de l'équation aux dérivées partielles. En général, à chaque valeur propre E_i , correspond une seule solution finie, uniforme, continue et nulle aux limites de D : c'est la fonction propre a_i correspondante. Il peut néanmoins arriver que plusieurs fonctions propres correspondent à une même valeur propre qui est alors dite « dégénérée ».

Les valeurs propres de l'équation (2) forment en général une suite discontinue. Toutefois, il faut faire à ce sujet une remarque importante. Dans les problèmes de quantification, le domaine D est très fréquem-

ment infini : par exemple, pour un corpuscule non soumis à des liaisons, il comprend en principe tout l'espace. Or dans ce cas, on ne peut affirmer que les valeurs propres forment une suite discontinue : une partie au moins des valeurs propres *peut* former une suite continue. On appelle « spectre » d'une équation telle que (2) l'ensemble de ces valeurs propres. Si le domaine D est fini, le spectre est toujours discontinu et l'on dit que l'on a affaire à un « spectre de raies ». Si le domaine est infini, on peut avoir un « spectre continu ». Au point de vue mathématique, l'étude des spectres continus est beaucoup plus délicate que celle des spectres de raies.

Les fonctions propres de l'équation (2) peuvent être complexes. En particulier, elles ne sont déterminées qu'à une constante multiplicative complexe près puisque (2) est linéaire. On peut toujours choisir (du moins pour les spectres de raies) le module de cette constante arbitraire de façon à avoir

$$\int_0 a_i^* a_i dz = \int_0 |a_i|^2 dz = 1.$$

On dit alors que les fonctions propres sont normées. La normalisation ne fixant que le module de la constante arbitraire, les fonctions propres normées contiennent encore un facteur arbitraire de la forme $e^{i\alpha}$.

Les fonctions propres de l'équation forment une famille de fonctions propres orthogonales. Plus précisément, si E_i et E_j sont deux valeurs propres différentes, les fonctions propres a_i et a_j correspondantes satisfont à la relation

$$\int_0 a_i^* a_j dz = 0,$$

Nous démontrerons ce théorème dans le chapitre XII. La démonstration ne s'applique que si E_i et E_j sont inégales. Dans le cas d'une valeur propre dégénérée, deux fonctions propres correspondant à cette même valeur propre ne sont pas nécessairement orthogonales. Mais dans ce cas il y a une indétermination des fonctions propres car toute combinaison linéaire de fonctions propres correspondant à E_i est encore une fonction propre correspondant à E_i . En prenant comme nouvelles fonctions propres des combinaisons linéaires convenablement choisies des anciennes, on peut s'arranger pour que les nouvelles fonctions propres soient orthogonales et, si l'on opère ainsi pour toutes les valeurs propres dégénérées, l'ensemble de toutes les fonctions propres sera bien orthogonal.

En introduisant le symbole de Kronecker δ_{ij} égal à 1 si $i = j$ et à 0

si $i \neq j$, on peut écrire pour l'ensemble « orthonormal » des fonctions α_i orthogonales et normées

$$\int_{\mathbf{D}} \alpha_i^* \alpha_j \, dz = \delta_{ij}$$

et, par suite, pour les ondes stationnaires $\Psi_i = \alpha_i e^{\frac{2\pi i}{h} E_i t}$

$$\int_{\mathbf{D}} \Psi_i^* \Psi_j \, dz = \delta_{ij}.$$

Il faut remarquer que les formules d'orthogonalité et de normalisation précédentes ne sont valables que pour les spectres de raies. Leur extension aux spectres continus exige des considérations délicates que nous n'aborderons pas ici.

Les fonctions propres α_i forment un « système complet », c'est-à-dire que sous des conditions très générales, une fonction quelconque $\varphi(x, y, z)$ définie dans le domaine \mathbf{D} peut se développer sous la forme

$$\varphi(x, y, z) = \sum_i c_i \alpha_i(x, y, z),$$

les c_i étant des constantes convenablement choisies. Nous aurons à revenir sur ce résultat qui généralise celui que nous avons rencontré pour les cordes et membranes vibrantes.

Nous allons maintenant étudier la quantification, par la méthode nouvelle de la mécanique ondulatoire, d'un certain nombre de systèmes importants pour la Physique atomique.

4. Le rotateur plan. — Par définition, le rotateur plan sera constitué par un corpuscule de masse m assujéti à rester sur un cercle de rayon R . La position du corpuscule est définie par l'angle au centre θ et, comme $V = 0$, l'équation (2) se réduit ici à

$$\frac{1}{R^2} \frac{d^2 a}{d\theta^2} + \frac{8\pi^2 m}{h^2} E a = 0,$$

soit, en introduisant le moment d'inertie $I = mR^2$ du rotateur

$$\frac{d^2 a}{d\theta^2} + \frac{8\pi^2 I}{h^2} E a = 0,$$

équation dont l'intégrale générale est

$$a(\theta) = A e^{\pm i \left(\frac{2\pi}{h} \sqrt{2EI} \theta + B \right)}$$

avec deux constantes arbitraires A et B .

Pour qu'une intégrale de cette forme soit fonction propre, il faut qu'elle soit uniforme sur le cercle de rayon R qui forme ici le domaine D . Il faut donc que, quand θ augmente de 2π , l'exponentielle dans $a(\theta)$ reprenne la même valeur, ce qui donne les valeurs propres

$$(4) \quad E_n = \frac{n^2 h^2}{8\pi^2 I} \quad (n = 0, 1, 2, \dots).$$

Telles sont les énergies des états stationnaires.

L'ancienne théorie des Quanta écrivait que le moment de quantité de mouvement $I \frac{d\theta}{dt}$ était égal à $n \frac{h}{2\pi}$ et en tirait

$$E_n = \frac{1}{2} I \left(\frac{d\theta}{dt} \right)^2 = \frac{1}{2I} \left(\frac{nh}{2\pi} \right)^2 = \frac{n^2 h^2}{8\pi^2 I}$$

et obtenait donc dans ce cas le même résultat que la Mécanique ondulatoire.

Les fonctions propres normées sont ici

$$a_n(\theta) = \frac{1}{\sqrt{2\pi}} e^{i n \theta} e^{i E_n t}$$

et comme

$$\int_0^{2\pi} e^{i(n-m)\theta} d\theta = 0 \quad \text{si } n \neq m,$$

on vérifie tout de suite l'orthogonalité des a_n . L'onde stationnaire $\Psi_n(\theta, t)$ s'obtient en multipliant $a_n(\theta)$ par $e^{\frac{2\pi i}{h} E_n t}$.

On remarquera encore que toutes les valeurs propres, sauf E_0 , sont dégénérées car il y a deux fonctions propres par valeur propre. On peut d'ailleurs remplacer les fonctions propres complexes données par des fonctions propres réelles en $\sin n\theta$ et $\cos n\theta$ qui en sont des combinaisons linéaires.

Au point de vue physique, le problème du rotateur plan peut paraître soulever une difficulté, car on imagine mal comment une onde peut être assujettie à se propager le long d'une ligne. La difficulté provient du fait que le problème est un peu trop schématisé. En réalité, le corpuscule du rotateur est soumis à une liaison qui doit s'exprimer physiquement de la façon suivante : il est rappelé vers le cercle de rayon R par des forces qui croissent extrêmement rapidement s'il s'écarte sensiblement de ce cercle. La fonction potentiel est donc nulle sur le cercle et croit très rapidement quand on s'éloigne tant soit peu du cercle. Si l'on traite le problème de cette façon, l'on voit que l'onde se propage dans un tube centré sur le cercle, son amplitude décroissant très rapidement

dès que l'on s'éloigne de l'axe du tube, ce qui du point de vue de la Mécanique ondulatoire signifie que le corpuscule a une très faible chance de s'éloigner tant soit peu du cercle. Si l'on fait croître indéfiniment l'intensité de la force de rappel, on obtient à la limite le problème traité où l'onde Ψ n'est différente de zéro que sur le cercle.

5. **Le rotateur sphérique.** — Par définition, le rotateur sphérique est constitué par un corpuscule de masse m assujéti à se déplacer sur la surface d'une sphère de rayon R .

Il y a ici une première remarque à faire qui est intéressante parce qu'elle montre bien la différence entre la Mécanique ancienne et la Mécanique ondulatoire. Dans l'ancienne Mécanique, le problème du rotateur sphérique ne différerait pas de celui du rotateur plan parce qu'un point matériel assujéti à rester sur une sphère de rayon R sans subir aucune force extérieure devait se déplacer suivant une géodésique de la sphère, c'est-à-dire suivant un cercle de rayon R . En Mécanique ondulatoire, au contraire, les deux problèmes sont tout à fait distincts car, le corpuscule du rotateur sphérique pouvant se mouvoir sur toute la sphère, son onde Ψ explore toute la surface de cette sphère qui forme le domaine D ici à deux dimensions.

Écrivons l'équation (2) en coordonnées sphériques θ et φ avec $V = 0$. Il vient

$$\frac{1}{R^2 \sin^2 \theta} \frac{\partial^2 \alpha}{\partial \varphi^2} + \frac{1}{R^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \alpha}{\partial \theta} \right) + \frac{8\pi^2 m}{h^2} E \alpha = 0.$$

Posons

$$K = \frac{8\pi^2}{h^2} m R^2 E = \frac{8\pi^2 I}{h^2} E,$$

I étant le moment d'inertie du rotateur sphérique. Nous pouvons donc écrire

$$\frac{1}{\sin^2 \theta} \frac{\partial^2 \alpha}{\partial \varphi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \alpha}{\partial \theta} \right) + K \alpha = 0$$

et il faut chercher pour quelles valeurs de K il existe au moins une solution finie, continue et uniforme sur la sphère, c'est-à-dire dans le domaine D défini par les inégalités $0 \leq \varphi \leq 2\pi$, $0 \leq \theta \leq \pi$.

Nous donnerons ici le résultat de la recherche des valeurs et des fonctions propres sans indiquer comment elle s'effectue. Les valeurs propres sont données par la formule

$$K = k(k+1) \quad (k = 0, 1, 2, \dots),$$

ce qui donne pour les énergies des états stationnaires

$$(5) \quad E_k = k(k+1) \frac{h^2}{8\pi^2 I} \quad (k = 0, 1, 2, \dots).$$

Les fonctions propres correspondantes sont les fonctions sphériques de Laplace $Y_k(\theta, \varphi)$. Dans la théorie de ces fonctions, on démontre que, pour chaque valeur de k , il y a $2k+1$ fonctions sphériques distinctes linéairement indépendantes qui ont pour expression

$$Y_k^m(\theta, \varphi) = e^{\pm im\varphi} P_k^m(\cos\theta) \quad (m = 0, 1, \dots, k),$$

où les $P_k^m(\cos\theta)$ sont des polynomes en $\cos\theta$ nommés « polynomes de Legendre associés » qui s'expriment eux-mêmes en fonction des polynomes de Legendre usuels par

$$P_k^m(\cos\theta) = (\sin\theta)^{\frac{m}{2}} \frac{d^m P_k(\cos\theta)}{(d\cos\theta)^m},$$

où $P_k(\cos\theta)$ est le $k^{\text{ième}}$ polynome de Legendre.

Finalement à la valeur propre E_k , correspondent les $2k+1$ fonctions propres indépendantes

$$a_{km}(\theta, \varphi) = e^{\pm im\varphi} P_k^m(\cos\theta) \quad (k = 0, 1, 2, \dots; m = 0, 1, 2, \dots, k).$$

Toutes les valeurs propres, excepté E_0 , sont donc dégénérées et la dégénérescence augmente avec l'ordre k de la valeur propre. On peut démontrer que l'on obtient ainsi toutes les fonctions propres et que celles-ci forment un système complet et orthogonal. Si l'on veut qu'elles soient normées, il faut multiplier a_{km} par le facteur de normalisation $\frac{4\pi}{2k+1} \frac{(k+m)!}{(k-m)!}$.

Pour obtenir l'onde stationnaire $\Psi_{km}(\theta, \varphi, t)$ qui correspond à $a_{km}(\theta, \varphi)$, on multipliera $a_{km}(\theta, \varphi)$ par $e^{\frac{2\pi i}{h} E_k t}$.

Comparons maintenant les valeurs propres ainsi obtenues avec celles que fournissait l'ancienne théorie des Quanta. Celle-ci nous donnait évidemment, d'après la remarque faite au début du paragraphe, pour le rotateur sphérique comme pour le rotateur plan les énergies quantifiées (4) qui diffèrent de (5) sauf pour les valeurs très élevées de k . Mais, pour calculer les raies spectrales par la règle de Bohr, nous avons à former les différences

$$E_k - E_{k'} = \frac{h^2}{8\pi^2 I} [k(k+1) - k'(k'+1)] = \frac{h^2}{8\pi^2 I} \left[\left(k + \frac{1}{2}\right)^2 - \left(k' + \frac{1}{2}\right)^2 \right].$$

Tout se passe donc comme si l'ancienne formule (1) était valable, mais avec des nombres quantiques demi-entiers $\frac{1}{2}, \frac{3}{2}, \dots, k + \frac{1}{2}, \dots$. Or précisément, lors de l'application des anciennes règles de quanta au rotateur sphérique (par exemple dans la théorie des spectres de bandes), on avait été conduit, pour obtenir une bonne concordance avec l'expérience, à admettre des valeurs demi-entières de cette forme pour les nombres quantiques et l'on parlait alors de « demi-quanta ». Cette hypothèse, entièrement arbitraire dans l'ancienne théorie des Quanta, trouve ici son explication.

Bien entendu, au point de vue physique, on doit supposer que le corpuscule du rotateur sphérique peut en principe s'écarter de la sphère de rayon R , mais qu'il est rappelé vers la surface de cette sphère, dès qu'il s'en écarte, par une force de liaison très intense. L'onde Ψ se propage dans une mince couche sphérique centrée sur la sphère de rayon R , son amplitude décroissant rapidement quand on s'éloigne de cette sphère. L'épaisseur de la couche tend vers zéro et nous retrouvons le problème traité ci-dessus quand la force de liaison croît indéfiniment.

6. L'oscillateur harmonique. — Nous allons maintenant appliquer la nouvelle méthode de quantification à l'oscillateur linéaire harmonique qui a servi, nous l'avons vu, de point de départ aux recherches de Planck sur les Quanta.

Par définition, l'oscillateur linéaire harmonique est constitué par un corpuscule de masse m assujéti à se mouvoir sur un axe rectiligne Ox en étant rappelé vers un point de cette droite par une force proportionnelle à la distance. Le centre de force étant pris pour origine de la coordonnée x , la fonction potentiel est

$$V(x) = K \frac{x^2}{2},$$

K mesure l'intensité de la force de rappel à l'unité de distance de l'origine. En Mécanique classique, l'oscillateur a une fréquence d'oscillation mécanique indépendante de son énergie et égale à

$$\omega = \frac{1}{2\pi} \sqrt{\frac{K}{m}}.$$

L'équation (2) s'écrit ici

$$\frac{d^2 a}{dx^2} + \frac{8\pi^2 m}{h^2} \left(E - K \frac{x^2}{2} \right) a = 0$$

et il faut chercher pour quelles valeurs de E cette équation admet au moins une solution $a(x)$ finie, uniforme, continue et nulle à l'infini du domaine formé par l'axe des x tout entier.

Par des raisonnements que nous n'exposons pas, on trouve comme valeurs propres

$$(6) \quad E_k = (2k + 1) \frac{h}{4\pi} \sqrt{\frac{K}{m}} = \left(k + \frac{1}{2}\right) h\omega \quad (k = 0, 1, 2, \dots),$$

alors que l'ancienne théorie des Quanta donnait

$$(6') \quad E_k = kh\omega.$$

Cette divergence entre l'ancienne et la nouvelle théorie est très remarquable. Ici encore la comparaison de l'ancienne formule de quantification avec l'expérience avait conduit à prendre dans (6') le nombre quantique k égal à un demi-entier $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$, ce qui était en contradiction avec la base même de l'ancienne méthode de quantification. La nouvelle méthode, en fournissant la formule (6), a levé la difficulté.

La fonction propre normée correspondant à E_k s'exprime par la formule

$$a_k(x) = \frac{\sqrt{2\pi} \sqrt{\frac{m\omega}{h}}}{\sqrt{2^k \pi k!}} e^{-\frac{2\pi^2 m \omega}{h} x^2} H_k \left(2\pi \sqrt{\frac{m\omega}{h}} x \right),$$

où H_k est un polynôme de degré k , le $k^{\text{ième}}$ polynôme d'Hermite. Les polynômes d'Hermite sont définis d'une façon générale par la relation

$$H_k(q) = (-1)^k e^{q^2} \frac{d^k}{dq^k} (e^{-q^2})$$

et satisfont à l'équation différentielle

$$\frac{d^2 H_k}{dq^2} - 2q \frac{dH_k}{dq} + 2k H_k = 0$$

ainsi qu'aux relations de récurrence

$$H'_k(q) = 2k H_{k-1}(q), \quad H_{k+1}(q) - 2q H_k(q) + 2k H_{k-1}(q) = 0.$$

Leur étude est classique en Analyse.

Pour obtenir la $k^{\text{ième}}$ onde stationnaire $\Psi_k(x, t)$, il suffit de multiplier la fonction propre $a_k(x)$ par le facteur $e^{\frac{2\pi i}{h} E_k t}$.

Ici, il n'y a pas de dégénérescence. Il y en aurait, au contraire, si l'on considérait l'oscillateur harmonique *isotrope* à deux ou trois dimensions,

c'est-à-dire le cas d'un corpuscule se mouvant dans un plan ou dans l'espace en étant rappelé vers un centre fixe par une force proportionnelle à l'élongation. On peut aussi considérer le problème plus compliqué d'un oscillateur harmonique à trois dimensions *non isotrope* pour lequel la force de rappel, proportionnelle à l'élongation, dépend de la direction de l'élongation et le problème, plus compliqué encore, de l'oscillateur *non harmonique* pour lequel la force de rappel n'est plus une fonction rigoureusement linéaire de l'élongation. Tous ces cas ont été étudiés, mais nous ne pouvons nous y arrêter ici.

7. L'atome d'hydrogène. — L'atome d'hydrogène est formé d'un noyau de charge $+e$ et d'un électron-planète de charge $-e$. Le noyau étant beaucoup plus lourd que l'électron, on peut le regarder, en première approximation, comme restant immobile et étudier le mouvement de l'électron de masse m_0 et de charge $-e$ dans le champ coulombien du noyau dérivant du potentiel électrostatique $\frac{e}{r}$. On a alors $V = -\frac{e^2}{r}$ et l'équation (2) s'écrit

$$\Delta a + \frac{8\pi^2 m_0}{h^2} \left(E + \frac{e^2}{r} \right) a = 0.$$

Nous devons chercher pour quelles valeurs de E , cette équation admet au moins une solution finie, uniforme et continue dans tout l'espace et s'annulant à l'infini. Le domaine D comprend tout l'espace et si l'on prend, comme il est naturel ici, des coordonnées sphériques autour du noyau, il est défini par les inégalités

$$0 \leq r \leq \infty, \quad 0 \leq \varphi \leq 2\pi, \quad 0 \leq \theta \leq \pi.$$

On peut trouver les valeurs et fonctions propres, soit par des méthodes empruntées à la théorie des fonctions analytiques, comme l'a fait M. Schrödinger, soit par d'autres méthodes et notamment des développements en série (Sommerfeld). On trouve pour les valeurs propres

$$(7) \quad E_n = -\frac{2\pi^2 m_0 e^4}{n^2 h^2} \quad (n = 1, 2, 3, \dots).$$

Ce sont les énergies des états stationnaires de l'atome d'hydrogène : elles coïncident avec celles trouvées par Bohr à l'aide de l'ancienne théorie des Quanta. Mais ici ces valeurs propres ne sont pas les seules, car on trouve aussi que toutes les valeurs positives de E sont valeurs propres. Il y a donc un spectre discontinu formé de valeurs négatives de E qui se serrent de plus en plus vers la valeur $E = 0$, puis un spectre

continu comprenant toutes les valeurs positives de E . Les valeurs propres positives correspondent aux mouvements hyperboliques de l'électron venant passer près du noyau et s'en éloignant ensuite. Les véritables états stationnaires correspondent aux valeurs propres négatives qui ont les valeurs calculées par Bohr. Ici la nouvelle méthode ne change rien aux résultats de l'ancienne.

A la valeur propre E_n , correspondent les fonctions propres suivantes :

$$a_{nlm}(r, \theta, \varphi) = e^{\pm im\varphi} P_l^m(\cos\theta) e^{-\frac{2\pi}{h}\sqrt{-2m_0E}r} v_p(r),$$

où l'on a

$$l = 0, 1, 2, \dots, n-1, \quad m = 0, 1, \dots, l \quad \text{et} \quad p = n - (l+1).$$

La fonction $v_p(r)$ peut s'exprimer à l'aide des polynomes de Laguerre étudiés en Analyse; nous n'insisterons pas sur leurs propriétés.

On voit que a_{nlm} dépend de θ et de φ par une fonction sphérique $Y_l(\theta, \varphi)$ et nous savons que, suivant la valeur de m , il y a $2l+1$ fonctions sphériques différentes correspondant à une même valeur de l . De plus, pour n donné, l peut prendre les valeurs $0, 1, \dots, n-1$. A la

valeur propre E_n correspondent donc un nombre égal à $\sum_0^{n-1} l(2l+1) = n^2$

fonctions propres différentes. Toutes les valeurs propres, sauf E_1 , sont dégénérées et l'ordre de dégénérescence est égal au carré du rang n de la valeur propre.

L'onde stationnaire $\Psi_{nlm}(r, \theta, \varphi, t)$ s'obtient en multipliant la fonction propre $a_{nlm}(r, \theta, \varphi)$ par l'exponentielle $e^{\frac{2\pi i}{h}E_n t}$.

Les fonctions propres $a_{nlm}(r, \theta, \varphi)$ sont orthogonales : on peut les normer quand on connaît les propriétés des fonctions $v_p(r)$.

Remarquons que les fonctions $a_{nlm}(r, \theta, \varphi)$ correspondent à un certain choix arbitraire de l'axe polaire des coordonnées sphériques et du plan passant par cet axe qui définit l'origine des longitudes φ . Si nous prenions d'autres coordonnées sphériques autour du noyau, nous aurions d'autres angles polaires θ' et φ' et nous serions amenés à prendre comme fonctions propres les fonctions

$$a_{nlm}(r, \theta', \varphi') = e^{\pm im\varphi'} P_l^m(\cos\theta') e^{-\frac{2\pi}{h}\sqrt{-2m_0E}r} v_p(r).$$

Les n^2 fonctions propres correspondant à la valeur propre E_n ainsi définies sont, bien entendu, des combinaisons linéaires des fonctions propres $a_{nlm}(r, \theta, \varphi)$ qui correspondent à la même valeur propre E_n .

Nous voyons ici entrer en jeu l'indétermination des fonctions propres qui correspondent à une même valeur propre.

Dans la nouvelle méthode de quantification de l'atome d'hydrogène, le nombre entier $p = n - (l + 1)$ joue le rôle du nombre quantique radial de l'ancienne théorie des Quanta, le nombre $l + 1 = n - p$ jouant le rôle de l'ancien nombre quantique azimutal k de la théorie de Sommerfeld. Comme l ne peut prendre que les valeurs $0, 1, \dots, n - 1$, on s'explique pourquoi le nombre k de la théorie de Sommerfeld ne pouvait prendre que les valeurs $1, 2, \dots, n$ avec exclusion de la valeur zéro, circonstance qui était mal expliquée dans l'ancienne théorie.

A cette remarque se rattache la suivante qui est très importante. Dans l'ancienne théorie, les états s caractérisés par $k = 1$ possédaient un moment de quantité de mouvement non nul correspondant à l'existence supposée d'une orbite fermée plane et ses états n'étaient pas isotropes. Dans la nouvelle théorie, les états s correspondent à $l = 0$, c'est-à-dire à un moment de quantité de mouvement nul. Les fonctions d'onde correspondantes a_{n00} ne dépendent que de r , car $P_0^0 = 1$: l'état s est isotrope. Cette isotropie des états s est entièrement confirmée par l'expérience, ce qui apporte un fort appui en faveur des images fournies par la Mécanique ondulatoire et montre l'impossibilité d'admettre l'existence d'une trajectoire de l'électron à l'intérieur de l'atome, car cette trajectoire définirait un plan privilégié s'opposant à l'isotropie.

L'électron ne peut plus, en effet, être considéré comme localisé à l'intérieur de l'atome. Sa probabilité de présence donnée par $|a_{nlm}(r, \theta, \varphi)|^2$ est en général partout différente de zéro et s'étend même jusqu'à l'infini. La probabilité de trouver l'électron à une distance du noyau comprise entre r et $r + dr$ avec dr donné varie évidemment comme le produit $r^2 |a_{nlm}|^2$. Si l'on construit la courbe représentant cette grandeur en fonction de r pour les états s , on trouve qu'elle présente des maxima au voisinage des valeurs de r qui correspondent aux trajectoires circulaires de la théorie de Bohr. Ainsi l'électron n'est plus localisé et n'a plus de trajectoire, mais il a cependant une probabilité particulièrement grande de se trouver au voisinage des trajectoires prévues par l'ancienne théorie des Quanta. Remarquons enfin que, si la probabilité de présence de l'électron ne s'annule qu'à l'infini, elle décroît néanmoins très rapidement quand r croît, en raison de la présence du facteur exponentiel dans $a_{nlm}(r, \theta, \varphi)$ et de ce fait exprime que l'électron est lié à l'atome.

On peut aussi calculer les fonctions propres correspondant au spectre continu, c'est-à-dire aux valeurs propres positives. Nous laisserons cette étude de côté.

Nous savons qu'en réalité la suite des états stationnaires de l'atome H est plus compliquée que ne l'indique la formule (7) de Bohr. En fait, les raies spectrales prévues par cette formule se décomposent en plusieurs raies : il y a des structures fines. Nous avons vu comment Sommerfeld avait procédé dans le cadre de l'ancienne théorie des Quanta pour obtenir une interprétation des structures fines en introduisant la Dynamique relativiste, mais déjà avant le développement de la nouvelle Mécanique, on savait qu'en réalité les structures fines ne peuvent être entièrement expliquées qu'en introduisant le « spin » de l'électron. La Mécanique ondulatoire, sous sa forme primitive telle que nous l'exposons en ce moment, n'est pas relativiste et ne tient pas compte du spin : c'est pourquoi elle ne peut rendre compte des structures fines. Nous verrons plus tard (chap. XIV) qu'il existe une Mécanique ondulatoire de l'électron due à Dirac qui, à la fois, est relativiste et introduit le spin de l'électron : avec elle, nous pourrions rendre mieux compte des structures fines.

8. L'effet Zeeman en Mécanique ondulatoire. — Nous venons d'étudier plusieurs exemples où l'on peut opérer rigoureusement la quantification du système envisagé. Il arrive souvent qu'on ne puisse pas calculer exactement les valeurs et fonctions propres. Mais parfois un artifice conduit à la solution : c'est ce qui se produit pour l'effet Zeeman sur un atome d'hydrogène.

Considérons un atome d'hydrogène plongé dans un champ magnétique uniforme \mathbf{H} . En Mécanique classique de l'électron, nous avons le théorème de Larmor qui nous apprend que, pour les champs magnétiques pas trop grands (de l'ordre des champs usuellement réalisables), le mouvement de l'électron rapporté à un système de référence tournant autour de la direction de \mathbf{H} avec la vitesse angulaire $\omega = \frac{e\mathbf{H}}{2m_0c}$ est le même que si le champ n'existait pas et si le système ne tournait pas. Ce théorème démontré par Larmor en Mécanique classique est encore exact en Mécanique ondulatoire : dans le système tournant pour les champs \mathbf{H} pas trop intenses, l'équation de propagation de l'onde Ψ est approximativement la même que si le champ \mathbf{H} n'existait pas et si le système ne tournait pas. Nous admettrons ce théorème dont on peut trouver la démonstration dans un livre de l'auteur ([4], p. 106). Partant de là, il est facile de calculer l'effet Zeeman. Prenons l'axe des z dans la direction du champ magnétique et employons des coordonnées cylindriques autour de cet axe. Dans le système tournant de Larmor, nous aurons

des coordonnées ρ' , z' , φ' qui seront reliées aux coordonnées ρ , z , φ du système fixe de l'observateur par les relations

$$\rho' = \rho, \quad z' = z, \quad \varphi' = \varphi - \omega t \quad \left(\omega = \frac{eH}{2m_0c} \right).$$

Comme dans le système tournant, les solutions sont les mêmes que dans un système fixe sans champ magnétique, on a dans ce système les valeurs propres

$$E_n^{(0)} = -\frac{2\pi^2 m_0 e^4}{n^2 h^2} \quad (n = 1, 2, \dots)$$

et les fonctions propres

$$a_{nlm}(\rho', z', \varphi') = F(\rho', z') e^{\pm im\varphi'} \quad (m = 0, 1, 2, \dots, l),$$

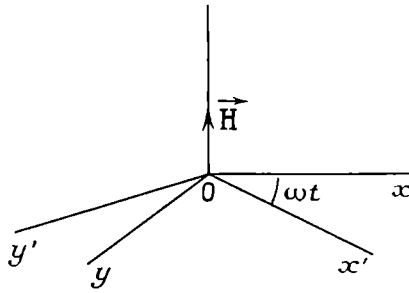


Fig. 26.

F étant une fonction dont la forme ne nous intéresse pas ici. L'onde stationnaire correspondante est

$$\Psi_{nlm}(\rho', z', \varphi', t) = F(\rho', z') e^{\pm im\varphi'} e^{\frac{2\pi i}{h} E_n t}.$$

Si nous passons au système de l'observateur par le changement de variables que nous avons indiqué plus haut, l'expression de l'onde Ψ deviendra

$$\Psi(\rho, z, \varphi, t) = F(\rho, z) e^{\pm im\varphi} e^{\frac{2\pi i}{h} \left(E_n^{(0)} - \frac{mh}{2\pi} \omega \right) t},$$

avec toujours $m = 0, 1, \dots, l$. Cette onde correspond donc à une énergie que nous pouvons écrire

$$E_{nm}^H = E_n^{(0)} + m \frac{h}{2\pi} \omega = E_n^{(0)} + m \frac{ehH}{4\pi m_0 c},$$

avec, cette fois, $m = -l, -l+1, \dots, l-1, l$.

Les valeurs propres correspondant aux diverses valeurs possibles de m sont maintenant distinctes, quoique voisines : l'action du champ magnétique a levé la dégénérescence et fait apparaître une structure fine.

A la transition ($nm \rightarrow n'm'$) en présence du champ magnétique H , correspond d'après la règle de Bohr l'émission d'une raie spectrale de fréquence

$$\nu = \frac{E_{nm}^H - E_{n'm'}^H}{h} = \frac{E_n^{(0)} - E_{n'}^{(0)}}{h} + (m' - m) \frac{eH}{4\pi m_0 c} = \nu_{nn'}^{(0)} + \delta m \frac{eH}{4\pi m_0 c},$$

$\nu_{nn'}^{(0)}$ étant la fréquence de la raie qui serait émise en l'absence de champ et δm la variation du nombre quantique « magnétique » lors de la transition.

Si l'on admet la règle de sélection $\delta m = 0, \pm 1$ que nous démontrerons plus loin (chap. XII, § 6) dans le cadre de la Mécanique ondulatoire, on retrouve exactement le triplet de l'effet Zeeman normal. Mais, pas plus que les théories antérieures, la Mécanique ondulatoire ne peut rendre compte des effets Zeeman anormaux tant qu'elle n'a pas introduit la notion de spin.

9. **Les méthodes de perturbation.** — Il arrive fréquemment qu'on ne peut trouver les valeurs propres d'un système quantifié ni par un calcul rigoureux, ni par un artifice comme celui que nous venons d'employer pour l'effet Zeeman en utilisant le système tournant de Larmor. Il faut alors tenter de faire des calculs de perturbations analogues à ceux qu'on emploie en Mécanique céleste.

Supposons, par exemple, que l'on sache calculer rigoureusement les valeurs et fonctions propres de l'équation (2) pour une certaine forme de l'énergie potentielle V . Soient E_n les valeurs propres et a_n les fonctions propres ainsi obtenues. Supposons maintenant que le système a une légère perturbation permanente, c'est-à-dire que le potentiel ne soit plus $V(x, y, z)$, mais $V(x, y, z) + v(x, y, z)$, avec $v \ll V$. Les valeurs et fonctions propres seront alors légèrement modifiées et deviendront $E_n + \delta E_n$ et $a_n + \delta a_n$, δE_n et δa_n étant de faibles variations et δa_n étant naturellement fonction de x, y, z . La méthode de perturbation à employer ici consistera à calculer approximativement les δE_n et δa_n en partant de la connaissance des E_n et des a_n . On parviendra ainsi à résoudre, au besoin par des approximations successives, les valeurs et fonctions propres du système perturbé.

Naturellement, cette méthode de perturbation permet de retrouver les résultats obtenus plus haut pour l'effet Zeeman, le champ magnétique extérieur imposé à l'atome ne perturbant que très peu l'effet du très fort champ électrostatique qui règne au voisinage du noyau. Elle s'applique aussi au cas d'un atome plongé dans un champ électrique uniforme

(effet Stark), car ici aussi les champs électriques que nous pouvons appliquer de l'extérieur à un atome sont toujours très faibles par rapport au champ coulombien intense au voisinage du noyau, de sorte qu'il n'y a encore qu'une petite perturbation de l'état quantique : on peut ainsi calculer les structures fines qui apparaissent dans l'effet Stark et le résultat est en très bonne concordance avec l'expérience.

Il existe aussi une autre méthode de perturbation due à M. Dirac et connue sous le nom de « méthode de variation des constantes ». Elle permet d'étudier les petites perturbations subies par un système sous l'action d'un champ extérieur *variable au cours du temps* et s'est montrée précieuse pour l'étude des interactions entre matière et rayonnement.

Nous n'étudierons pas en détail ces méthodes de perturbation. On pourra se reporter, pour les approfondir, aux Ouvrages cités dans la Bibliographie.

BIBLIOGRAPHIE.

- [1] COURANT et HILBERT, *Methoden der mathematischen Physik*, Springer, Berlin, 1924.
- [2] E. SCHRÖDINGER, *Mémoires sur la Mécanique ondulatoire*, Presses universitaires, Paris, 1937.
- [3] LOUIS de BROGLIE, *La quantification dans la nouvelle Mécanique*, Paris, Hermann, 1932.
- [4] LOUIS de BROGLIE, *L'électron magnétique*, Paris, Hermann, 1934.
- [5] A. SOMMERFELD, *Wellenmechanik*, Frederick Ungar, New-York.
- [6] A. MARCH, *Quantum Mechanics of particles and wave fields*, Wiley and Sons, New-York, et Chapman and Hall, Londres, 1951.

CHAPITRE XII.

MÉCANIQUE QUANTIQUE D'HEISENBERG ET PRINCIPE DE CORRESPONDANCE.

1. **Point de départ de la Mécanique quantique d'Heisenberg.** — En 1925, au moment où commençait à se développer la Mécanique ondulatoire, M. Heisenberg, et à sa suite MM. Born et Jordan, posaient les bases d'une Mécanique nouvelle qui pouvait paraître tout à fait différente de la Mécanique ondulatoire et qu'on a alors appelée la Mécanique quantique ou Mécanique des matrices.

Nous allons exposer les idées essentielles qui ont guidé M. Heisenberg. Tout d'abord, il a adopté une attitude phénoménologique, c'est-à-dire qu'il a cherché à n'introduire dans sa théorie que des éléments observables en écartant systématiquement tout élément non observable. Ainsi, dans la théorie primitive de l'atome de Bohr, on se représentait l'électron atomique comme ayant une position, une vitesse, une trajectoire; mais, en fin de compte, tous ces éléments ne servaient que d'intermédiaires de calcul, les seuls résultats utilisés pour prévoir les fréquences observables des raies spectrales étant les énergies des états stationnaires. En particulier, la fréquence de révolution de l'électron sur son orbite, fréquence qui dans la théorie classique devait déterminer la fréquence des raies spectrales émises et absorbées, n'est pas un élément directement observable. Les fréquences d'émission et d'absorption réellement observables dans les spectres, ainsi d'ailleurs que les fréquences critiques de dispersion, sont celles que l'on calcule par la règle de Bohr et, en dehors du domaine des grands nombres quantiques, elles n'ont aucune relation simple avec les fréquences de révolution. Une des idées fondamentales d'Heisenberg a été qu'une bonne théorie de l'atome doit faire abstraction de toute grandeur qui n'est pas observable et ne faire intervenir que des éléments observables tels que les fréquences, les polarisations et les intensités des raies spectrales qu'il peut émettre.

D'autre part, Heisenberg, élève de Bohr, s'est aussi laissé guider par

le principe de correspondance. Nous avons vu que dans la théorie classique le rayonnement est déterminé par le développement de Fourier des composantes du moment électrique du système émetteur. Soit, par exemple, un atome formé de particules électrisées de charges e_1, e_2, \dots, e_n . Soit q l'une quelconque des trois coordonnées x, y, z : la $k^{\text{ième}}$ particule a une coordonnée q_k par rapport à l'axe des q . La composante \mathcal{X}_q du moment électrique de l'atome est, par définition,

$$\mathcal{X}_q = \sum_1^N e_k q_k.$$

Or, chaque q_k dans la conception classique a un développement de Fourier de la forme

$$q_k = \sum_{-\infty}^{+\infty} \underline{q}_{\tau}^{(k)} e^{2\pi i(\tau\nu)t},$$

avec

$$\underline{q}_{\tau}^{(k)} = q_{\tau}^{(k)} \quad \text{et} \quad (\tau\nu) = \tau_1\nu_1 + \dots + \tau_n\nu_n$$

s'il faut n variables angulaires $\omega_1, \omega_2, \dots, \omega_n$ pour définir le système constitué par l'atome. On a donc

$$\mathcal{X}_q = \sum_1^N \sum_{-\infty}^{+\infty} e_k \underline{q}_{\tau}^{(k)} e^{2\pi i(\tau\nu)t} = \sum_{-\infty}^{+\infty} \mathcal{X}_{\tau}^{(q)} e^{2\pi i(\tau\nu)t}$$

en posant

$$\mathcal{X}_{\tau}^{(q)} = \sum_1^N e_k q_{\tau}^{(k)}.$$

D'après la théorie classique, l'atome ainsi constitué émet toutes les fréquences $(\tau\nu)$, l'énergie émise en moyenne par seconde sous forme d'une radiation de fréquence $(\tau\nu)$ dont la vibration électrique est parallèle à l'axe des q étant égale à

$$\frac{64 \pi^4 (\tau\nu)^4}{3 c^3} \left| \sum_1^N e_k q_{\tau}^{(k)} \right|^2.$$

Telle est l'image classique du rayonnement de l'atome. Passons maintenant à l'image quantique.

Dans la théorie quantique, l'atome est susceptible d'une série d'états quantifiés caractérisés chacun par un ensemble de nombres quantiques n_1, n_2, \dots . Quand l'atome passe de l'état caractérisé par l'ensemble des nombres quantiques n à l'état caractérisé par l'ensemble des

nombres quantiques n' , il y a émission de la fréquence $\nu_{nn'} = \frac{E_n - E_{n'}}{h}$.

Un ensemble d'un grand nombre d'atomes identiques de même orientation émet toutes les fréquences $\nu_{nn'}$ et Heisenberg admet comme postulat que l'énergie rayonnée en moyenne par cet ensemble d'atomes par unité de temps et par atome sous forme d'une radiation dont la vibration électrique est parallèle à l'axe des q peut toujours s'écrire sous la forme

$$\frac{64 \pi^4 \nu_{nn'}^4}{3 c^3} \left| \sum_k^N e_k q_{nn'}^{(k)} \right|^2.$$

Ce postulat définit en somme les $q_{nn'}^{(k)}$ qui déterminent ainsi l'intensité de la radiation de fréquence $\nu_{nn'}$ polarisée suivant l'axe des q et qui correspondent, par suite, à un élément observable.

Heisenberg prend donc comme point de départ que l'on peut caractériser chaque sorte d'atome par un ensemble de nombres complexes $q_{nn'}^{(k)} e^{2\pi i \nu_{nn'} t}$ où ne figurent que des quantités observables, fréquences des raies spectrales et intensités correspondant aux divers états de polarisation. Ces nombres complexes « correspondent » aux coefficients classiques du développement du moment électrique en série de Fourier bien qu'on ne puisse plus les identifier avec ces coefficients. Guidé par l'idée de correspondance, Heisenberg admet, par analogie avec la relation $q_{-\tau}^{(k)} = q_{\tau}^{(k)*}$, que l'on a

$$q_{nn'}^{(k)} = q_{n'n}^{(k)*}.$$

En effet, d'après le théorème de correspondance pour les fréquences, τ « correspond » à $n - n'$ et, par suite, $-\tau$ à $n' - n$: donc $q_{nn'}^{(k)}$ correspond à $q_{-\tau}^{(k)}$ et $q_{n'n}^{(k)}$ à $q_{-\tau}^{(k)}$.

Finalement, à chaque coordonnée q_k intervenant dans la description classique de l'atome, Heisenberg fait correspondre un ensemble de nombres complexes dont on peut dresser le tableau, les nombres n et n' servant à définir les lignes et les colonnes de ce tableau. Un tel tableau constitue une « matrice ». Une telle matrice est dite « hermitique » ou « hermitienne » parce que l'on a

$$q_{nn'}^{(k)} e^{2\pi i \nu_{nn'} t} = (q_{n'n}^{(k)} e^{2\pi i \nu_{n'n} t})^*,$$

En effet, on a $q_{nn'}^{(k)} = q_{n'n}^{(k)*}$ (postulat de Heisenberg) et $\nu_{n'n} = -\nu_{nn'}$ (loi des fréquences de Bohr). Dans ces matrices, les termes symétriques par rapport à la diagonale sont imaginaires conjugués tandis que les termes diagonaux sont réels.

Généralisant, Heisenberg a alors admis qu'à toute grandeur A caractéristique de l'atome, on pouvait faire correspondre dans sa Mécanique quantique une infinité de nombres complexes $a_{nn'}$ $e^{2\pi i\nu_{nn'}t}$ où les $\nu_{nn'}$ sont les fréquences de Bohr pour cet atome et où les $a_{nn'}$ forment une matrice hermitienne telle que $a_{n'n} = a_{nn'}^*$.

Grâce à une autre analogie avec la théorie classique, Heisenberg est arrivé à une règle importante relative aux matrices de la nouvelle Mécanique : la règle de multiplication. Considérons les développements classiques de Fourier de deux grandeurs physiques A et B

$$A = \sum_{\tau}^{+\infty} a_{\tau} e^{2\pi i(\tau\nu)t}, \quad B = \sum_{\tau}^{+\infty} b_{\tau} e^{2\pi i(\tau\nu)t}.$$

Peut-on trouver un développement de Fourier analogue pour la grandeur AB ? La réponse est affirmative, car on a

$$AB = \sum_{\tau} a_{\tau} e^{2\pi i(\nu\tau)t} \sum_{\tau'} b_{\tau'} e^{2\pi i(\tau'\nu)t} = \sum_{\sigma} \sum_{\tau'} a_{\tau} b_{\tau'} e^{2\pi i((\tau+\tau')\nu)t}$$

et, en posant $\sigma = \tau + \tau'$, on peut écrire

$$AB = \sum_{\sigma} (ab)_{\sigma} e^{2\pi i(\sigma\nu)t} \quad \left((\sigma\nu) = \sum_i \sigma_i \nu_i \right),$$

avec la définition

$$(1) \quad (ab)_{\sigma} = \sum_{\tau}^{+\infty} a_{\tau} b_{\sigma-\tau}.$$

Cette formule nous fournit la règle permettant d'obtenir le coefficient d'indice σ du développement de Fourier du produit AB en fonction des coefficients des développements de Fourier de A et de B .

Comment peut-on transposer ces considérations classiques dans la théorie quantique d'Heisenberg? Ici, aux deux grandeurs A et B , nous faisons correspondre deux tableaux de nombres complexes suivant le schéma

$$A \sim a_{nn'} e^{2\pi i\nu_{nn'}t}, \quad B \sim b_{nn'} e^{2\pi i\nu_{nn'}t}.$$

Si nous faisons « correspondre » dans la formule (1) σ à $n - n'$, c'est-à-dire à la paire d'indices nn' , τ à $n - l$ avec l quelconque, c'est-à-dire à la paire d'indices nl , enfin $\sigma - \tau$ à $n - n' - (n - l) = l - n'$, c'est-à-dire à la paire d'indices ln' , la formule (1) se transformera par correspondance en la suivante :

$$(1') \quad (ab)_{nn'} = \sum_l a_{nl} b_{ln'}.$$

L'élément d'indices nn' de la matrice AB est donc égal à $\sum_l a_{nl} b_{ln'}$: c'est la formule de multiplication des matrices bien connue en algèbre. Il en résulte qu'en général la matrice AB ne coïncide pas avec la matrice BA . Dans l'algèbre des matrices, il n'y a pas commutativité de la multiplication.

Comme on a évidemment

$$A + B \sim (a_{nn'} + b_{nn'}) e^{2\pi i \nu_{nn'} t},$$

c'est-à-dire

$$(1'') \quad (a + b)_{nn'} = a_{nn'} + b_{nn'};$$

l'ensemble des formules (1') et (1'') montre que les matrices d'Heisenberg ont toutes les propriétés des matrices algébriques.

Remarquons encore que l'on peut écrire à la place de (1')

$$(ab)_{nn'} e^{2\pi i \nu_{nn'} t} = \sum_l a_{nl} e^{2\pi i \nu_{nl} t} b_{ln'} e^{2\pi i \nu_{ln'} t}.$$

En effet, on a, quel que soit l , $\nu_{nn'} = \nu_{nl} + \nu_{ln'}$ en vertu de la loi des fréquences de Bohr, car

$$\frac{1}{h}(E_n - E_{n'}) = \frac{1}{h}(E_n - E_l) + \frac{1}{h}(E_l - E_{n'}).$$

Cette remarque montre la cohérence des définitions d'Heisenberg avec la loi de Bohr,

2. Équations canoniques et relations de commutation. — Pour développer sa Mécanique quantique, Heisenberg a admis que la forme des équations de cette Mécanique devait être la même que celle des équations de la Mécanique classique, mais avec substitution aux grandeurs qui interviennent dans celle-ci des matrices qui désormais les remplacent. Il écrit donc les équations de Hamilton

$$(2) \quad \frac{dP_k}{dt} = - \frac{\partial H}{\partial Q_k}, \quad \frac{dQ_k}{dt} = \frac{\partial H}{\partial P_k},$$

où les Q_k et les P_k sont les matrices qui remplacent les variables de Lagrange et où H est la matrice obtenue en substituant dans l'expression hamiltonienne classique de l'énergie pour le problème considéré les matrices Q_k et P_k aux variables q_k et p_k . Pour donner un sens précis à ces nouvelles équations canoniques, il faut définir les dérivées qui y

figurent. Pas de difficulté pour les dérivées $\frac{d}{dt}$, car les éléments de la matrice A étant $a_{nm}e^{2\pi i\nu_{nm}t}$, on admettra tout naturellement que $\frac{dA}{dt}$ est la matrice dont les éléments sont $2\pi i\nu_{nm}a_{nm}e^{2\pi i\nu_{nm}t}$. Plus délicates à donner sont les définitions des dérivées $\frac{\partial H}{\partial Q_k}$ et $\frac{\partial H}{\partial P_k}$ d'une matrice par rapport à une autre matrice. On peut cependant en donner une définition parfaitement logique et cohérente, comme l'a montré notamment M. Born, et les nouvelles équations de Hamilton prennent ainsi un sens parfaitement défini.

A ces équations canoniques, Heisenberg a vu que, pour introduire la constante de Planck et obtenir une Mécanique quantique, il fallait adjoindre des relations contenant h . Toujours guidé par des considérations de correspondance que nous ne reproduirons pas, il a admis que, pour deux variables canoniquement conjuguées, on doit poser

$$(3) \quad P_k Q_k - Q_k P_k = -\frac{h}{2\pi i} 1$$

le 1 du second membre représentant la « matrice unité »

$$\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{array}$$

dont les termes diagonaux, seuls différents de zéro, sont tous égaux à 1 , de sorte que $1_{nm} = \delta_{nm}$. La matrice $P_k Q_k - Q_k P_k$ est donc une matrice diagonale dont les termes diagonaux, seuls différents de zéro, sont égaux à $\frac{h}{2\pi i}$. Par contre, les matrices correspondant à des variables non canoniquement conjuguées commutent et l'on a

$$(3') \quad \begin{cases} Q_i Q_k - Q_k Q_i = 0, & P_i P_k - P_k P_i = 0 & \text{pour tout } i \text{ et tout } k; \\ P_i Q_k - Q_k P_i = 0 & & \text{pour } i \neq k, \end{cases}$$

(3) et (3') sont les relations de commutation (Vertauschungsrelationen) de Heisenberg.

S'appuyant sur l'ensemble des équations (2), (3) et (3'), Heisenberg est parvenu à déterminer les énergies des états stationnaires des systèmes quantifiés tels que : oscillateur linéaire, atome d'hydrogène, etc. Le principe qui lui a servi pour effectuer cette détermination est le suivant : quand on a trouvé la matrice $H(Q, P)$ qui correspond à l'énergie dans

le problème considéré, on ramène cette matrice à la forme diagonale, ce qui est toujours possible, et l'on a alors

$$H_{nm} = \delta_{nm} E_n.$$

Les termes diagonaux E_n sont alors, par définition, les énergies des états stationnaires. Lorsqu'en 1925, Heisenberg est parvenu à obtenir les énergies quantifiées des systèmes atomiques, la Mécanique ondulatoire n'était pas encore arrivée à faire ce calcul. C'est en 1926 que Schrödinger exécuta le calcul des états quantifiés par les méthodes de la Mécanique ondulatoire étudiées au chapitre XI, et les résultats qu'il a ainsi obtenus se sont trouvés en concordance parfaite avec ceux de la Mécanique quantique. Au premier abord, cette concordance a paru étrange, car les deux sortes de Mécanique nouvelle semblaient tout à fait différentes et sans lien entre elles. Schrödinger a dissipé ce mystère en montrant que la Mécanique quantique des matrices n'est qu'une transposition, sous une forme mathématique différente, de la Mécanique ondulatoire. Les matrices d'Heisenberg peuvent se définir à l'aide des fonctions d'onde de la Mécanique ondulatoire : alors les équations canoniques (2) et les relations de commutation (3) et (3') reçoivent une interprétation simple. Le calcul d'Heisenberg constitue donc finalement une méthode qui permet d'obtenir les mêmes résultats essentiels que la Mécanique ondulatoire en ce qui concerne la valeur des énergies quantifiées, sans faire le calcul explicite des fonctions d'onde. Suivant les cas, cette méthode peut être plus ou moins commode que celle de la Mécanique ondulatoire, mais elle lui est équivalente. L'accord des résultats obtenus apparaît alors comme tout naturel.

3. Retour sur certains points de la Mécanique ondulatoire. — Pour pouvoir démontrer l'identité de la Mécanique de Heisenberg avec la Mécanique ondulatoire, il nous faut revenir sur certains points de celle-ci.

Nous allons indiquer une manière très simple et en quelque sorte automatique de retrouver la forme de l'équation des ondes Ψ . Soit $H(q_k, p_k, t)$ la fonction hamiltonienne d'un corpuscule en coordonnées cartésiennes, c'est-à-dire la fonction qui, en Mécanique classique, représente pour le problème considéré l'énergie du corpuscule en fonction de l'ensemble des coordonnées q_k et des moments de Lagrange p_k et éventuellement du temps t . Dans l'expression de cette fonction, remplaçons chaque moment p_k par l'opérateur (symbole d'opération) $-\frac{h}{2\pi i} \frac{d}{dq_k}$,

où h est la constante de Planck. Nous obtenons ainsi un opérateur hamiltonien

$$\mathfrak{H}\left(q_k, -\frac{h}{2\pi i} \frac{\partial}{\partial q_k}, t\right)$$

et pour trouver l'équation des ondes Ψ en Mécanique ondulatoire, il suffit d'écrire

$$\frac{h}{2\pi i} \frac{\partial \Psi}{\partial t} = \mathfrak{H}\left(q_k, -\frac{h}{2\pi i} \frac{\partial}{\partial q_k}, t\right) \Psi.$$

En effet, pour le corpuscule de masse m dans un champ dérivant de la fonction potentiel $U(x, y, z, t)$, on a en Mécanique classique

$$H(q_k, p_k, t) = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + U(x, y, z, t)$$

et l'équation des ondes doit être, d'après la règle précédente,

$$\frac{h}{2\pi i} \frac{\partial \Psi}{\partial t} = -\frac{h^2}{8\pi^2 m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \Psi + U \Psi,$$

ce qui est équivalent à

$$\Delta \Psi - \frac{8\pi^2 m}{h^2} U \Psi = \frac{4\pi i m}{h^2} \frac{\partial \Psi}{\partial t},$$

c'est-à-dire à l'équation des ondes trouvée par d'autres considérations au chapitre IX.

Dans certains cas, il pourrait y avoir ambiguïté dans la formation de l'opérateur hamiltonien à partir de la fonction hamiltonienne classique. Si, par exemple, on avait dans celle-ci un terme en qp , il ne reviendrait pas au même d'écrire $q\left(-\frac{h}{2\pi i} \frac{\partial}{\partial q}\right)$ ou $-\frac{h}{2\pi i} \frac{\partial}{\partial q} q$. Mais en coordonnées cartésiennes rectangulaires, ces difficultés ne se présentent jamais et nous les laisserons ici de côté.

Dans le cas des systèmes quantifiables, le potentiel est indépendant du temps et l'on peut chercher des solutions monochromatiques

$$\Psi = a(x, y, z) e^{\frac{2\pi i}{h} Et},$$

où $a(x, y, z)$ obéit à l'équation de Schrödinger

$$\Delta a + \frac{8\pi^2 m}{h^2} (E - U) a = 0.$$

Nous savons qu'il existe une série de valeurs de la constante E , soit E_1, E_2, \dots , dites « valeurs propres » pour lesquelles cette équation

admet au moins une solution finie, continue, uniforme et nulle aux limites (fonction propre). Les fonctions propres $a_n(x, y, z)$ ou, si l'on veut, les fonctions d'onde correspondantes $\Psi_n(x, y, z, t)$ qui n'en diffèrent que par le facteur de module unité $e^{\frac{2\pi i}{h} E_n t}$, forment un système complet de fonctions orthogonales que l'on peut supposer normées par la condition

$$\int_{\mathfrak{D}} |a_n|^2 d\tau = \int_{\mathfrak{D}} |\Psi_n|^2 d\tau = 1$$

et qui sont alors bien déterminées à un facteur de la forme $e^{i\alpha}$ près.

L'orthogonalité des fonctions propres résulte du fait que l'opérateur hamiltonien \mathfrak{H} est un opérateur « hermitique » ou « self-adjoint ». D'une façon générale, un opérateur \mathfrak{A} est dit hermitique ou self-adjoint quand, $f(x, y, z)$ et $g(x, y, z)$ étant deux fonctions finies, continues, uniformes et nulles à l'infini, on a

$$\int_{\mathfrak{D}} f^*(xyz) \mathfrak{A} g(xyz) d\tau = \int_{\mathfrak{D}} g \mathfrak{A}^* f^* d\tau.$$

On vérifie facilement que c'est bien le cas de l'opérateur hamiltonien \mathfrak{H} et, comme de plus, \mathfrak{H} est réel, on a alors

$$\int_{\mathfrak{D}} f^* \mathfrak{H} g d\tau = \int_{\mathfrak{D}} g \mathfrak{H} f^* d\tau.$$

Voici alors comment l'orthogonalité des Ψ_n entre eux résulte de l'hermiticité de \mathfrak{H} . On a d'abord

$$\mathfrak{H} \Psi_m = E_m \Psi_m, \quad \mathfrak{H} \Psi_n^* = E_n^* \Psi_n^*,$$

d'où l'on tire aisément

$$\int_{\mathfrak{D}} [\Psi_n^* \mathfrak{H} \Psi_m - \Psi_m \mathfrak{H} \Psi_n^*] d\tau = (E_m - E_n^*) \int_{\mathfrak{D}} \Psi_n^* \Psi_m d\tau.$$

Le premier membre est nul en vertu de l'hermiticité de \mathfrak{H} et il reste

$$(E_m - E_n^*) \int_{\mathfrak{D}} \Psi_n^* \Psi_m d\tau = 0,$$

équation qui doit être exacte pour toutes les valeurs de m et de n .

Faisons d'abord $n = m$: nous trouvons $E_n = E_n^*$, ce qui nous montre que les valeurs propres E_n sont réelles, comme l'exige la signification physique que nous leur donnons.

Faisons ensuite $m \neq n$: nous trouvons, si $E_m \neq E_n$,

$$\int_{\mathfrak{D}} \Psi_n^* \Psi_m d\tau = 0.$$

Donc les fonctions propres correspondant à des valeurs propres différentes sont orthogonales.

Mais il peut arriver qu'à une valeur propre « dégénérée » E_i correspondent plusieurs fonctions propres linéairement indépendantes $\Psi_{i,1}, \Psi_{i,2}, \dots, \Psi_{i,p}$ telles que

$$\mathfrak{H} \Psi_{i,l} = E_i \Psi_{i,l} \quad (l = 1, 2, \dots, p).$$

On voit alors aisément que toute combinaison linéaire des p fonctions propres $\Psi_{i,l}$ est encore une fonction propre correspondant à la valeur propre E_i . Le système des fonctions propres n'est alors défini qu'à une substitution linéaire près. Si l'on a trouvé par un procédé quelconque les p fonctions propres $\Psi_{i,1}, \dots, \Psi_{i,p}$, elles ne sont pas en général orthogonales, mais on peut les remplacer par p de leurs combinaisons linéaires qui soient linéairement indépendantes et orthogonales entre elles : les p nouvelles fonctions propres, dont le choix reste d'ailleurs partiellement arbitraire, constituent donc avec les fonctions propres correspondant aux autres valeurs propres un système orthogonal.

Le système des fonctions propres orthonormales de l'opérateur \mathfrak{H} est un système complet, c'est-à-dire que si $f(q_1, \dots, q_k, \dots)$ est une fonction des coordonnées q satisfaisant à des conditions très générales, on peut toujours développer cette fonction sous la forme

$$f(q_k) = \sum_i c_i a_i(q_k) = \sum_i d_i(t) \Psi_i(q_k, t),$$

avec

$$d_i(t) = c_i e^{-\frac{2\pi i}{h} E_i t}.$$

La valeur des coefficients c_i et d_i est facile à calculer en tenant compte de l'orthonormalité des fonctions propres et l'on trouve

$$c_i = \int_{\mathfrak{D}} a_i^*(q_k) f(q_k) d\tau, \quad d_i(t) = \int_{\mathfrak{D}} \Psi_i^*(q_k, t) f(q_k) d\tau.$$

Les formules donnant les coefficients d'une série de Fourier ordinaire sont des cas particuliers des formules précédentes. Aussi les coefficients c_i et d_i sont-ils souvent nommés les « coefficients de Fourier » du développement de $f(q_k)$ suivant les fonctions propres a_i ou Ψ_i .

Si la fonction $f(q_k)$ est elle-même normée, c'est-à-dire si

$$\int_{\mathfrak{D}} |f|^2 d\tau = 1,$$

on voit aisément que

$$\sum_i |c_i|^2 = \sum_i |d_i|^2 = 1.$$

4. Interprétation des matrices d'Heisenberg en Mécanique ondulatoire. — Quand nous avons formé plus haut l'opérateur hamiltonien, nous avons pris l'expression classique de l'énergie en fonction des coordonnées et des moments dans le problème considéré et nous l'avons transformée en opérateur. On a été amené en Mécanique ondulatoire, comme nous le verrons plus en détail dans le chapitre XIII, à généraliser l'idée qui nous a servi à faire cette transformation et l'on a admis qu'à toute grandeur physique mesurable (*observable*, au sens de Dirac) définie par la théorie classique, on doit en Mécanique nouvelle faire correspondre un opérateur \mathfrak{A} linéaire et hermitique, c'est-à-dire tel que

$$\mathfrak{A}(f + g) = \mathfrak{A}(f) + \mathfrak{A}(g), \quad \int_{\mathfrak{D}} f^* \mathfrak{A} g d\tau = \int_{\mathfrak{D}} g \mathfrak{A}^* f^* d\tau.$$

Comment doit-on procéder pour obtenir l'opérateur correspondant à une grandeur observable de la mécanique classique ? Tout d'abord à la coordonnée q_k d'un corpuscule, on fera correspondre l'opérateur « multiplication par q_k », soit q_k , puis à un moment de Lagrange p_k , on fera correspondre l'opérateur $-\frac{h}{2\pi i} \frac{\partial}{\partial q_k}$. Toutes les grandeurs mécaniques classiques peuvent s'exprimer à l'aide des q_k et des p_k et, pour obtenir l'opérateur quantique correspondant à l'une de ces grandeurs classiques, on n'a qu'à remplacer dans l'expression classique de cette grandeur chaque q_k par q_k , et chaque p_k par $-\frac{h}{2\pi i} \frac{\partial}{\partial q_k}$. En particulier, c'est bien cette méthode qui nous a fourni l'opérateur hamiltonien correspondant à l'énergie. En employant des coordonnées cartésiennes rectangulaires, il ne se présente jamais de difficultés quant à l'ordre des facteurs.

La Mécanique ondulatoire a donc à envisager des opérateurs linéaires et hermitiques correspondant aux grandeurs classiques mesurables. Soit \mathfrak{A} l'un de ces opérateurs. Nous pouvons l'appliquer à l'une des fonctions propres Ψ_l de l'opérateur \mathfrak{H} du système envisagé. Nous obten-

nous ainsi une fonction $\mathfrak{A}\Psi_l$ que nous pouvons développer ensuite suivant le système complet des fonctions propres Ψ_k sous la forme

$$\mathfrak{A}\Psi_l = \sum_k A_{kl}\Psi_k,$$

les coefficients A_{kl} que l'on désigne souvent par le symbole $(k|\mathfrak{A}|l)$ ayant pour valeur

$$(4) \quad A_{kl} \equiv (k|\mathfrak{A}|l) = \int_{\mathfrak{D}} \Psi_k^* \mathfrak{A}\Psi_l d\tau.$$

L'ensemble des A_{kl} pour toutes les valeurs de k et de l définit une matrice A , qui est engendrée par l'opérateur \mathfrak{A} dans le système complet des fonctions propres orthonormales de l'opérateur hamiltonien \mathfrak{H} .

Il est facile de vérifier que, les opérateurs \mathfrak{A} considérés en Mécanique ondulatoire étant toujours hermitiques, les matrices A formées à partir de ces opérateurs sont elles aussi hermitiques, car l'hermiticité de A entraîne que $A_{kl} = A_{lk}^*$.

Les matrices A que nous venons de définir ont toutes les propriétés des matrices de la Mécanique quantique d'Heisenberg et peuvent être identifiées avec elles. Vérifions que leur algèbre est la même, c'est-à-dire qu'elles admettent leurs propriétés d'addition et de multiplication.

Il est d'abord évident d'après la définition (4) que, si la matrice C , somme de deux matrices A et B , correspond par définition à l'opérateur \mathfrak{C} , somme des opérateurs \mathfrak{A} et \mathfrak{B} , elle a pour élément $C_{kl} = A_{kl} + B_{kl}$. C'est la règle d'addition des matrices algébriques. Pour retrouver la règle de multiplication, partons de l'idée que la matrice C , produit de la matrice A par la matrice B , correspond à l'opérateur $\mathfrak{C} = \mathfrak{A}\mathfrak{B}$ et écrivons les développements

$$\mathfrak{A}\Psi_l = \sum_n A_{nl}\Psi_n, \quad \mathfrak{B}\Psi_k = \sum_l B_{lk}\Psi_l,$$

d'où nous tirons

$$\mathfrak{A}\mathfrak{B}\Psi_k = \mathfrak{A} \sum_l B_{lk}\Psi_l = \sum_l B_{lk} \mathfrak{A}\Psi_l = \sum_l B_{lk} \sum_n A_{nl}\Psi_n.$$

On peut encore écrire

$$\mathfrak{C}\Psi_k = \sum_n C_{nk}\Psi_n$$

en posant

$$C_{nk} = \sum_l A_{nl}B_{lk},$$

ce qui est bien la formule de multiplication des matrices (1). On n'a donc pas, en général, la relation matricielle $AB = BA$, parce que les opérateurs \mathfrak{A} et \mathfrak{B} ne sont pas, en général, équivalents. Lorsque l'on a par exception $AB = BA$, on dit que les matrices A et B commutent et alors il en est de même des opérateurs \mathfrak{A} et \mathfrak{B} , et inversement.

Ayant montré que les matrices de la Mécanique ondulatoire satisfont aux mêmes règles d'addition et de multiplication que celles d'Heisenberg montrons qu'elles satisfont aussi aux règles de commutation (3) et (3'). En effet, si q_k et p_k sont deux variables canoniquement conjuguées, on a entre les opérateurs quantiques \mathfrak{q}_k et \mathfrak{p}_k correspondant la relation

$$\mathfrak{p}_k \mathfrak{q}_k - \mathfrak{q}_k \mathfrak{p}_k = \left(-\frac{h}{2\pi i} \frac{\partial}{\partial q_k} \right) q_k - q_k \left(-\frac{h}{2\pi i} \frac{\partial}{\partial q_k} \right) = -\frac{h}{2\pi i}.$$

car

$$\frac{\partial}{\partial q_k} q_k f(q_k) - q_k \frac{\partial}{\partial q_k} f(q_k) = f(q_k).$$

On a donc pour les éléments de la matrice $P_k Q_k - Q_k P_k$

$$(P_k Q_k - Q_k P_k)_{nm} = \int_{\mathfrak{D}} \Psi_n^* (\mathfrak{p}_k \mathfrak{q}_k - \mathfrak{q}_k \mathfrak{p}_k) \Psi_m d\tau = -\frac{h}{2\pi i} \delta_{nm},$$

en raison de l'orthonormalité des Ψ_n . D'où la relation matricielle (3). Les relations de commutation (3') sont évidentes.

On peut démontrer que la signification des équations canoniques (2) de Heisenberg est la suivante : les matrices employées par Heisenberg sont celles qui sont engendrées par l'opérateur hamiltonien \mathfrak{H} (et non pas dans un autre système complet de fonctions orthonormales).

Il nous reste à justifier la manière employée par Heisenberg pour déterminer les énergies des états stationnaires. Comme on a

$$\mathfrak{H} \Psi_n = \frac{h}{2\pi i} \frac{\partial \Psi_n}{\partial t} = E_n \Psi_n,$$

il est évident que, si H est ramenée à sa forme diagonale, on a

$$H_{nn} = \int_{\mathfrak{D}} \Psi_n^* \mathfrak{H} \Psi_n d\tau = E_n.$$

Les termes diagonaux de la matrice H donnent donc bien les énergies des états quantifiés correspondant aux fonctions propres.

(1) M. Nicolas Cabrera a attiré l'attention sur le fait que, pour des raisons d'uniformité de convergence, le raisonnement précédent n'est pas rigoureux et que, dans certains cas, la règle de multiplication des matrices peut être en défaut.

Nous avons ainsi montré la complète équivalence de la Mécanique quantique d'Heisenberg avec la Mécanique ondulatoire.

§. **Le principe de correspondance dans la nouvelle Mécanique.** — Rappelons d'abord que, d'après le postulat initial d'Heisenberg, si l'on a un grand nombre d'atomes dans le même état quantique symbolisé par n , l'énergie rayonnée en moyenne par unité de temps et par atome sous la forme d'une radiation de fréquence $\nu_{nn'}$ dont le champ électrique vibre parallèlement à l'axe des q est donnée par

$$\frac{64 \pi^4 \nu_{nn'}^4}{3 c^3} \left| \sum_k e_k q_{nn'}^{(k)} \right|^2,$$

la sommation portant sur les corpuscules chargés qui forment le système; $q_{nn'}^{(k)}$ est l'élément de matrice correspondant à la coordonnée q du $k^{\text{ième}}$ corpuscule électrisé de charge e_k .

Cet énoncé doit être conservé en Mécanique ondulatoire en y donnant aux matrices d'Heisenberg l'interprétation que nous venons de préciser. En nous bornant au cas où le système est formé d'un seul électron, l'énergie rayonnée dans les conditions spécifiées plus haut sera

$$\frac{64 \pi^4 \nu_{nn'}^4}{3 c^3} e^2 |q_{nn'}|^2 \quad (q = x, y, z),$$

avec

$$q_{nn'} = \int_D \Psi_n^* q \Psi_{n'} d\tau,$$

puisque la matrice dont les éléments sont $q_{nn'}$ correspond à l'opérateur q .

On peut chercher à donner une sorte d'interprétation physique à la règle d'Heisenberg à l'aide d'une image du type classique développée par Schrödinger. Il est utile de connaître cette interprétation bien qu'en définitive elle ne puisse être considérée comme satisfaisante. Pour l'établir, remarquons d'abord que pour un corpuscule dans un état Ψ la probabilité de localisation dans l'élément $d\tau$ est $\Psi\Psi^* d\tau$, de sorte que la valeur moyenne de la coordonnée q du corpuscule est

$$\bar{q} = \int q \Psi^* \Psi d\tau.$$

Dès lors, pour un atome à un électron dont la fonction d'onde est

$$\Psi = \sum_i c_i \Psi_i,$$

la valeur moyenne de la composante \mathfrak{X}_q du moment électrique doit naturellement s'écrire

$$\bar{\mathfrak{X}}_q = -e \int_{\mathfrak{D}} \Psi^* q \Psi d\tau = \sum_{ik} c_i^* c_k (-e) \int_{\mathfrak{D}} \Psi_i^* q \Psi_k d\tau = \sum_{ik} c_i^* c_k (\mathfrak{X}_q)_{ik}$$

en posant

$$(\mathfrak{X}_q)_{ik} = (-e) \int_{\mathfrak{D}} \Psi_i^* q \Psi_k d\tau = -e \int_{\mathfrak{D}} a_i^* q a_k d\tau e^{\frac{2\pi i}{h}(\mathbf{E}_k - \mathbf{E}_i)t}.$$

Le moment électrique moyen $\bar{\mathfrak{X}}$ ainsi défini contient donc des termes présentant toutes les fréquences de combinaison de la forme $\frac{\mathbf{E}_k - \mathbf{E}_i}{h}$ prévues par la loi de Bohr. On serait donc tenté de dire que l'atome dans l'état défini par la fonction Ψ rayonne comme il le ferait d'après la théorie classique s'il possédait un moment électrique égal au moment électrique moyen défini par la Mécanique ondulatoire : on aurait ainsi une sorte d'interprétation physique de la loi des fréquences de Bohr. Malheureusement, cette interprétation séduisante est trompeuse, car elle serre de trop près les conceptions classiques et méconnaît la véritable nature des discontinuités quantiques. Si, en effet, on prenait cette image à la lettre, on serait amené à regarder l'atome comme émettant simultanément et continûment toutes les fréquences de Bohr, alors qu'il est certain que l'émission de ces fréquences est liée aux processus discontinus de transitions quantiques. Disons aussi, sans pouvoir insister ici sur ce point, qu'on serait également conduit à des difficultés en ce qui concerne le rapport des intensités des émissions avec le nombre des atomes émetteurs qui se trouvent initialement dans les divers états quantiques.

Bref, il vaut mieux renoncer à l'image de Schrödinger et se borner à énoncer, à la façon d'Heisenberg, la règle relative à la prévision des intensités d'émission de la manière suivante :

Si l'on a initialement N atomes dans l'état quantique d'énergie \mathbf{E}_i , l'intensité de la radiation de fréquence $\nu_{ik} = \frac{\mathbf{E}_i - \mathbf{E}_k}{h}$ émise en moyenne par unité de temps par les N atomes avec champ électrique vibrant parallèlement à l'axe des q est $N \frac{64 \pi^4 \nu_{ik}^4}{3 c^3} |(\mathfrak{X}_q)_{ik}|^2$.

On peut vérifier que cette règle appliquée au domaine des grands nombres quantiques donne les mêmes résultats que l'application des lois classiques, comme l'exige le principe de correspondance. On peut donc

considérer la règle d'Heisenberg comme constituant la forme précise que prend le principe de correspondance dans la nouvelle Mécanique, et ce principe ne présente plus ainsi les ambiguïtés qu'il présentait dans le cadre de l'ancienne théorie des Quanta (*voir* chap. VIII, § 4).

6. Démonstration des règles de sélection. — Ayant ainsi obtenu une forme nouvelle, plus précise et plus satisfaisante du principe de correspondance pour les intensités, nous allons pouvoir reprendre sur cette base la démonstration des « règles de sélection » indiquée au chapitre VIII, paragraphe 5.

Reprenons la démonstration de la règle de sélection pour le nombre quantique magnétique m dans l'effet Zeeman pour l'atome d'hydrogène. Prenons le système de référence suivant où l'axe Oz coïncide avec la direction du champ magnétique \mathbf{H} .

Les fonctions propres qui représentent les états stationnaires de l'atome d'hydrogène sont toujours de la forme

$$a_{nlm}(r, \theta, \varphi) = F_{nlm}(r, \theta) e^{im\varphi} \quad (m \text{ entier});$$

m est ici le nombre quantique magnétique. Soit alors un certain état stationnaire initial correspondant à $a_1 = F_1 e^{im_1\varphi}$ et un état stationnaire final correspondant à $a_2 = F_2 e^{im_2\varphi}$. Envisageons la transition ($1 \rightarrow 2$) et les éléments de matrice du moment électrique correspondant. Comme on a

$$x = r \sin \theta \cos \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \theta,$$

il vient (au facteur $-e$ près)

$$\begin{aligned} (\mathcal{X}_x + i \mathcal{X}_y)_{12} &= \int_{\mathbf{D}} \Psi_1^*(x + iy) \Psi_2 d\tau = \int_{\mathbf{D}} \Psi_1^* r \sin \theta e^{i\varphi} \Psi_2 r^2 \sin \theta dr d\theta d\varphi \\ &= \int \Phi(r, \theta) dr d\theta \int_0^{2\pi} e^{i(m_2 - m_1 + 1)\varphi} d\varphi, \\ (\mathcal{X}_x - i \mathcal{X}_y)_{12} &= \int_{\mathbf{D}} \Psi_1^*(x - iy) \Psi_2 d\tau = \int_{\mathbf{D}} \Psi_1^* r \sin \theta e^{-i\varphi} \Psi_2 r^2 \sin \theta dr d\theta d\varphi \\ &= \int \Phi(r, \theta) dr d\theta \int_0^{2\pi} e^{i(m_2 - m_1 - 1)\varphi} d\varphi, \\ (\mathcal{X}_z)_{12} &= \int_{\mathbf{D}} \Psi_1^* z \Psi_2 d\tau = \int_{\mathbf{D}} \Psi_1^* r \cos \theta \Psi_2 r^2 \sin \theta dr d\theta d\varphi \\ &= \int \Phi'(r, \theta) dr d\theta \int_0^{2\pi} e^{i(m_2 - m_1)\varphi} d\varphi. \end{aligned}$$

Si $\delta m = m_2 - m_1$ n'est égal ni à 0, ni ± 1 , les trois composantes de \mathcal{X} sont nulles et la transition est interdite.

Si $\delta m = 0$, on a

$$(\mathcal{X}_x)_{12} = (\mathcal{X}_y)_{12} = 0, \quad (\mathcal{X}_z)_{12} \neq 0.$$

Il y a émission d'une raie avec vibration parallèle à Oz.

Si $\delta m = +1$, on a

$$(\mathcal{X}_z)_{12} = (\mathcal{X}_x + i\mathcal{X}_y)_{12} = 0, \quad (\mathcal{X}_x - i\mathcal{X}_y)_{12} \neq 0.$$

Il y a émission d'une raie avec polarisation circulaire droite dans le plan xOy .

Si $\delta m = -1$, on a

$$(\mathcal{X}_z)_{12} = (\mathcal{X}_x - i\mathcal{X}_y)_{12} = 0, \quad (\mathcal{X}_x + i\mathcal{X}_y)_{12} \neq 0.$$

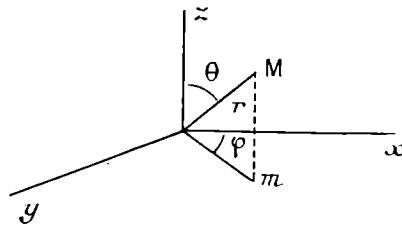


Fig. 27.

Il y a émission d'une raie avec polarisation circulaire gauche dans le plan xOy .

On a donc ainsi retrouvé la règle de sélection $\delta m = 0, \pm 1$ pour le nombre quantique magnétique, ce qui ramène au triplet Zeeman normal.

7. La formule de dispersion de Kramers-Heisenberg. — La Mécanique nouvelle a pu retrouver une théorie de la dispersion analogue à la théorie classique. Pour cela, elle calcule la perturbation créée dans l'état d'un atome, dont on connaît l'état stationnaire initial, par l'action d'une onde lumineuse, puis elle en déduit le moment électrique moyen acquis par l'atome sous l'influence de cette perturbation. En supposant l'atome isotrope, on trouve une relation de la forme $\vec{\mathcal{P}} = \alpha' \vec{\mathbf{h}}$, \mathbf{h} étant le champ électrique de l'onde lumineuse incidente. En employant alors la formule classique $n^2 - 1 = 4\pi N \alpha'$ (voir chap. I), on trouve une formule de dispersion, dite formule de Kramers-Heisenberg, très analogue à la formule classique. Elle rend compte comme celle-ci de l'allure générale des phénomènes de dispersion, mais elle contient en plus certaines caractéristiques nouvelles qui sont en accord avec l'expérience.

En outre, la formule de Kramers-Heisenberg permet de lever une difficulté de la théorie classique. Celle-ci donne une formule où figurent les nombres des électrons constituant dans l'atome des oscillateurs de diverses fréquences. Or, en comparant la formule classique avec les résultats expérimentaux, on peut calculer ces nombres et l'on trouve qu'ils sont fractionnaires, ce qui est en contradiction avec leur définition. Dans la théorie quantique, au contraire, chaque électron dans un état stationnaire initial donné peut subir diverses transitions, chaque transition correspondant à l'un des oscillateurs de la conception classique : tout se passe donc comme si chaque électron intraatomique contribuait à plusieurs oscillateurs, se partageait en quelque sorte statistiquement entre ces oscillateurs. L'intervention des nombres fractionnaires s'explique alors immédiatement et la formule de Kramers-Heisenberg, en introduisant des nombres dits « intensités d'oscillateur », en rend compte quantitativement.

BIBLIOGRAPHIE.

- [1] W. HEISENBERG, *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen* (*Zeits. Physik*, Bd 33, 1925, p. 879).
- [2] E. SCHRÖDINGER, *Mémoires sur la Mécanique ondulatoire*, Presses universitaires, Paris, 1937.
- [3] BORN et JORDAN, *Elementare quantenmechanik*, Springer, Berlin, 1930.
- [4] LOUIS de BROGLIE, *Le principe de correspondance*, Hermann, Paris, 1938.

CHAPITRE XIII.

L'INTERPRÉTATION PROBABILISTE DE LA MÉCANIQUE ONDULATOIRE.

1. **Idées générales.** — Nous avons déjà vu qu'il existe une différence profonde entre les conceptions de l'ancienne Mécanique et celles de la nouvelle, plus généralement entre les conceptions de la Physique classique et celles de la Physique quantique. L'image des phénomènes fournie par la Physique classique suppose essentiellement que nous pouvons définir le monde physique par la valeur de certaines grandeurs sans nous préoccuper d'une façon essentielle de la manière dont nous parvenons à connaître ces valeurs, c'est-à-dire à « *mesurer* » ces grandeurs. On suppose donc implicitement que l'on peut, par une technique expérimentale assez fine, diminuer indéfiniment la perturbation qu'une opération de mesure exerce sur l'état de choses existant avant la mesure, de sorte que chaque mesure traduit exactement la situation telle qu'elle existait immédiatement avant la mesure. La Physique quantique a introduit la notion de quantum d'action et elle s'est aperçu que l'existence de ce quantum d'action ne permettait pas de diminuer indéfiniment la perturbation qu'une mesure exerce sur une situation antérieure : la perturbation résiduelle, minima, est insignifiante à grande échelle, mais elle devient prépondérante à l'échelle des phénomènes élémentaires : c'est ce que montrent des exemples comme celui du microscope d'Heisenberg.

La mesure d'une grandeur ne révèle donc pas un état de choses existant avant la mesure, mais un état de choses créé par la mesure elle-même, et il est indispensable, quand on discute une situation physique, de faire intervenir le processus de mesure qui nous a permis de la connaître. La mesure « *extrait* » de l'état de choses antérieur une valeur bien déterminée de la grandeur mesurée. Nous ne pouvons plus

dire, en général, que, dans l'état de choses antérieur à la mesure représenté en Mécanique ondulatoire par une certaine fonction Ψ , une grandeur physique ait une valeur bien déterminée, mais seulement qu'elle a des valeurs possibles, c'est-à-dire différentes valeurs qu'une opération de mesure peut extraire de l'état de choses. La Physique quantique représentera donc une situation physique par l'ensemble des valeurs possibles des diverses grandeurs physiques, mais elle peut aller plus loin et attribuer à ces diverses grandeurs des probabilités. La Physique classique décrit le monde matériel par la connaissance précise des grandeurs physiques et leur évolution au cours du temps. Au contraire, la Physique quantique décrit les situations physiques, au moyen de la fonction Ψ , par la connaissance des valeurs possibles des grandeurs physiques et de leurs probabilités respectives et elle cherche à suivre l'évolution au cours du temps de ces probabilités. Naturellement, en Physique quantique, chaque opération de mesure, en fournissant des valeurs précises pour certaines grandeurs, bouleverse l'état des probabilités d'une façon qui n'est pas prévisible *a priori*; mais aucune opération de mesure ne peut conduire à un état de choses tout à fait déterminé au sens de l'ancienne Physique puisque, d'après les analyses de Bohr et de Heisenberg résumées par les relations d'incertitude, aucune opération de mesure ne peut fournir en même temps d'une façon précise les valeurs de deux grandeurs canoniquement conjuguées. Une opération de mesure, faite avec toute la précision que permet l'existence du quantum d'action, laisse toujours incertaines les valeurs d'au moins la moitié des grandeurs physiques, de telle sorte que, dans l'état de choses créé par la mesure, on ne peut attribuer à ces grandeurs que diverses valeurs possibles affectées de probabilités.

Nous avons déjà vu comment la dualité d'aspect onde-corpuscule conduit aux relations d'incertitude d'Heisenberg. Nous voulons maintenant montrer que, si l'on cherche à associer des photons aux ondes lumineuses, on rencontre tout de suite des questions analogues à celles que l'on rencontre pour l'électron et les autres corpuscules en Mécanique ondulatoire, et cet examen jette un jour très vif sur la nature de la représentation par ondes.

Considérons une onde lumineuse quelconque. On peut décomposer cette onde de plusieurs manières en se plaçant à des points de vue différents. D'abord, au point de vue des fréquences, on peut faire la décomposition « spectrale » en une suite finie ou infinie de composantes monochromatiques, ce qui s'effectue mathématiquement en développant la fonction d'onde en série ou en intégrale de Fourier. Si l'on envoie

Fondu sur un appareil qui sépare les fréquences, un réseau par exemple, on obtient chaque composante de Fourier séparément avec son intensité propre. Si maintenant on associe, par la pensée, des photons aux ondes lumineuses, on voit que l'action du réseau doit avoir pour effet de répartir les photons entre les composantes spectrales proportionnellement à l'intensité de celles-ci. Avant le passage à travers le réseau, chaque photon ne peut pas être considéré en général comme ayant une fréquence déterminée, puisqu'il est associé à une onde incidente où généralement plusieurs fréquences sont présentes, mais après le passage dans le réseau, le photon a une fréquence déterminée et la probabilité pour que ce soit telle fréquence de l'onde incidente est proportionnelle à l'intensité de la composante de Fourier correspondante dans l'onde incidente. Le réseau est donc un appareil qui permet d'attribuer au photon une fréquence et, par suite, de mesurer son énergie $h\nu$. On peut dire que la composition spectrale de l'onde incidente représente les valeurs possibles de la fréquence (ou de l'énergie) du photon associé et leurs probabilités respectives. Ces valeurs possibles sont les valeurs des fréquences qui figurent dans la décomposition spectrale et les probabilités correspondantes sont les intensités des composantes spectrales. Appliquée à l'électron, la même correspondance entre les valeurs de l'énergie et la composition spectrale de l'onde associée conduit au principe de Born que nous avons déjà énoncé au chapitre X.

Mais l'on peut aussi décomposer une onde lumineuse en se plaçant, par exemple, au point de vue des polarisations. On peut notamment décomposer une onde lumineuse en deux composantes rectilignement polarisées dans deux directions A et B à angle droit. Un nicol convenablement orienté laissera passer la vibration parallèle à A et arrêtera la vibration parallèle à B. La vibration transmise à travers le nicol sera désormais « simple » par rapport aux directions A, B : elle sera totalement transmise par un nicol parallèle au premier, totalement arrêtée par un nicol tourné de 90° . C'est ⁽¹⁾ un « cas pur » pour la décomposition en vibrations rectilignement polarisées suivant A et B. Mais ce ne sera pas un cas pur pour la décomposition suivant une autre paire de directions rectangulaires A', B', puisque l'onde sortant du premier nicol ne sera que partiellement transmise par un second nicol qui laisse passer les vibrations parallèles à la direction A'. Introduisons maintenant l'idée de photon. La présence du nicol a pour effet de répartir les pho-

(1) Le mot « cas pur » n'est pas pris ici exactement dans le même sens que dans la théorie des mélanges de J. von Neumann.

tons en deux catégories : ceux qui sont transmis et dont l'onde associée est rectilignement polarisée, et ceux qui sont arrêtés et à qui l'on doit évidemment attribuer la polarisation perpendiculaire. On ne peut pas attribuer à un photon avant l'action du nicol un état de polarisation rectiligne, mais seulement des probabilités de polarisation rectiligne suivant les axes A et B, mesurées par l'intensité des composantes correspondantes. Après le passage à travers le nicol, le photon constitue « un cas pur » pour la polarisation rectiligne suivant A, B, c'est-à-dire qu'il sera sûrement transmis par un nicol parallèle au premier et sûrement arrêté par un nicol croisé avec le premier. Mais le photon ainsi « préparé » par le passage à travers le premier nicol ne sera pas un cas pur quant à la polarisation pour un nicol dont les axes A', B' ne seraient pas parallèles à ceux du premier, car l'onde de ce photon polarisé rectilignement suivant A aura des composantes non nulles suivant A' et B'. De même, il ne sera pas un cas pur pour un analyseur circulaire et, inversement, un photon préparé par un analyseur circulaire ne sera pas un cas pur pour un analyseur rectiligne.

En résumé, tout appareil analyseur placé sur le trajet d'une onde lumineuse oblige les photons associés à cette onde à se révéler comme ayant tel ou tel état de polarisation et la probabilité pour qu'un photon se révèle ainsi avec tel ou tel état de polarisation s'obtient en décomposant l'onde d'une façon qui correspond à l'analyseur employé (décomposition en composantes rectilignement polarisées à angle droit dans le cas du nicol, décomposition en composantes circulaires de sens inverse dans le cas des analyseurs circulaires, etc.). Et toujours, c'est l'intensité de chaque composante qui donne la probabilité correspondante.

Ces idées se transposent exactement dans l'interprétation probabiliste générale de la Mécanique ondulatoire. Tout dispositif permettant de mesurer exactement l'une des grandeurs qui caractérisent un corpuscule de l'échelle atomique oblige ce corpuscule à se révéler comme étant dans un état où cette grandeur a une valeur bien déterminée, mais antérieurement à l'action de l'appareil de mesure, il est généralement impossible d'attribuer à la grandeur en question une valeur bien déterminée, on peut seulement lui attribuer des valeurs possibles affectées de probabilités. Pour trouver ces valeurs possibles et ces probabilités, on doit effectuer une certaine décomposition de l'onde Ψ associée initialement au corpuscule, décomposition déterminée par la nature de la grandeur à mesurer ou, si l'on préfère, par le processus de mesure à employer. Dans cette décomposition, chaque composante correspond à l'une des valeurs possibles de la grandeur et son intensité mesure la

probabilité de cette valeur possible. Après la mesure, l'état du corpuscule est un cas pur, au sens adopté ici de cette expression, pour la grandeur en question, mais il ne l'est pas, en général, pour une autre grandeur dont la mesure exige un autre dispositif expérimental.

Les conceptions que nous venons d'esquisser sont très différentes des conceptions classiques sur la mesure des grandeurs corpusculaires. Il faut bien réfléchir sur ces conceptions si l'on veut bien comprendre le sens des prévisions de la nouvelle Physique. Nous verrons dans ce chapitre qu'on ne peut pas sans cela saisir le véritable sens de la « quantification dans l'espace » du moment de la quantité de mouvement et qu'il en est de même pour le spin.

Passons maintenant à l'énoncé précis des postulats qui servent à donner à la Mécanique ondulatoire son interprétation probabiliste générale.

2. Grandeurs et opérateurs en Mécanique ondulatoire. — Nous avons déjà vu qu'en Mécanique ondulatoire, on fait correspondre à toute grandeur mesurable un opérateur linéaire et hermitique. Nous avons appris à former ces opérateurs pour les grandeurs mécaniques à l'aide de la correspondance

$$q_k \Rightarrow q_k, \quad p_k \Rightarrow -\frac{h}{2\pi i} \frac{\partial}{\partial q_k}.$$

Rappelons les définitions suivantes :

1° Un opérateur \mathfrak{A} est linéaire si l'on a

$$\mathfrak{A}(\varphi_1 + \varphi_2) = \mathfrak{A}\varphi_1 + \mathfrak{A}\varphi_2, \quad \mathfrak{A}c\varphi = c\mathfrak{A}\varphi,$$

c étant une constante complexe quelconque.

2° Un opérateur est hermitique si l'on a

$$\int_{\mathfrak{D}} f^* \mathfrak{A}g \, d\tau = \int_{\mathfrak{D}} g \mathfrak{A}^* f^* \, d\tau,$$

les intégrations étant faites dans le domaine \mathfrak{D} de variation des variables du problème envisagé et les fonctions f et g de ces variables étant finies, continues et uniformes dans \mathfrak{D} et tendant vers zéro aux limites de \mathfrak{D} assez vite pour que les intégrales de surface introduites par des intégrations par partie soient nulles. \mathfrak{A}^* est l'opérateur obtenu en changeant i en $-i$ dans la définition de \mathfrak{A} .

Il faut d'ailleurs faire une distinction importante entre deux catégories d'opérateurs qui s'introduisent en Mécanique ondulatoire. Les uns intéressent l'ensemble des variables du domaine \mathfrak{D} et sont appelés

pour cette raison « *opérateurs complets* ». Les autres n'intéressent, au contraire, qu'une partie des dites variables et sont nommés « *opérateurs incomplets* ». L'opérateur hamiltonien \mathfrak{H} est un exemple d'opérateur complet, tandis que l'opérateur $\mathfrak{p}_x = -\frac{h}{2\pi i} \frac{\partial}{\partial x}$ est, dans le cas d'un corpuscule pouvant se déplacer dans l'espace à trois dimensions, l'exemple d'un opérateur incomplet.

Comme il résulte des considérations générales du paragraphe 1, le but de la Mécanique ondulatoire doit être de nous dire quels sont les nombres, nécessairement *réels*, qu'une mesure précise peut nous fournir comme valeurs d'une grandeur physique. De l'opérateur linéaire et hermitique que la nouvelle Mécanique fait correspondre à toute grandeur physique mesurable, nous devons donc pouvoir déduire une liste de nombres réels représentant les résultats possibles d'une mesure précise de cette grandeur. Or ceci est possible parce que tous les opérateurs linéaires et hermitiques de la Mécanique ondulatoire possèdent une suite de valeurs propres qui sont des nombres réels.

Soit une grandeur physique mesurable et \mathfrak{A} l'opérateur linéaire et hermitique qui lui correspond en Mécanique ondulatoire. L'équation (valable dans le domaine D)

$$\mathfrak{A}\varphi = \alpha\varphi,$$

où α est une constante et φ une fonction des variables sur lesquelles opère \mathfrak{A} est « l'équation aux valeurs propres » de la grandeur considérée. Nous admettrons qu'une telle équation possède toujours une suite continue ou discontinue de valeurs propres α_i auxquelles correspondent une ou plusieurs fonctions propres φ_i assujetties à satisfaire des conditions que nous avons antérieurement précisées. S'il y a plusieurs fonctions propres $\varphi_{i,1}, \dots, \varphi_{i,p}$ linéairement indépendantes correspondant à une même valeur propre α_i , cette valeur propre est dite « *dégénérée* » ou « *multiple* ».

En nous basant sur le caractère hermitique de \mathfrak{A} , nous pourrions ici encore démontrer, exactement comme nous l'avons fait pour l'opérateur \mathfrak{H} au paragraphe 3 du chapitre XII, les deux résultats suivants :

1° Les valeurs propres α_i (qui peuvent éventuellement ici dépendre du temps si \mathfrak{A} en dépend) sont *réelles*;

2° Deux fonctions propres φ_i et φ_j correspondant à deux valeurs propres *distinctes* α_i et α_j sont orthogonales, c'est-à-dire que l'on a

$$\int_{\mathfrak{D}} \varphi_i^* \varphi_j d\tau = 0.$$

Nous pourrions reproduire ici ce que nous avons dit sur l'indétermination des fonctions propres correspondant à une valeur propre dégénérée et la manière d'utiliser cette indétermination pour obtenir, même dans ce cas, des fonctions propres orthogonales, sur la possibilité de « normer » les fonctions propres φ_i , sur les difficultés particulières qui se présentent dans le cas des spectres continus.

Ici encore le système des fonctions propres φ_i constituera un système « complet » permettant d'écrire des développements de la forme

$$f = \sum_i d_i \varphi_i,$$

le signe Σ devant être remplacé par une intégration dans le cas des spectres continus. Compte tenu de la relation d'orthonormalité

$$\int_D \varphi_i^* \varphi_j d\tau = \delta_{ij},$$

on trouve

$$d_j = \int_D \varphi_j^* f d\tau$$

et les d_j peuvent être appelés les coefficients de Fourier du développement de la fonction f suivant les fonctions propres normées de l'opérateur \mathfrak{A} .

3. Principes généraux de l'interprétation probabiliste de la Mécanique ondulatoire. — En Mécanique ondulatoire nous devons représenter tout état d'un corpuscule par une fonction d'onde $\Psi(x, y, z, t)$ solution de l'équation de propagation

$$(1) \quad \frac{h}{2\pi i} \frac{\partial \Psi}{\partial t} = \mathfrak{H} \Psi,$$

où \mathfrak{H} est l'opérateur hamiltonien du problème considéré, et nous devons toujours supposer que cette fonction d'ondes est normée par la condition

$$(2) \quad \int_D |\Psi|^2 d\tau = 1.$$

Nous sommes maintenant en mesure d'énoncer les principes généraux de l'interprétation probabiliste de la Mécanique ondulatoire.

Premier principe (principe de quantification). — Les valeurs possibles à un instant t d'une grandeur physique mesurable attachée à un

corpuscule, c'est-à-dire les résultats possibles d'une mesure de cette grandeur faite à l'instant t , sont les valeurs propres (calculées à l'instant t) de l'opérateur linéaire et hermitique correspondant à la grandeur.

Deuxième principe (principe de décomposition spectrale généralisée). — Si nous savons que l'état d'un corpuscule est représenté par une fonction d'onde $\Psi(x, y, z, t)$, solution de l'équation (1), la probabilité pour qu'une mesure précise d'une grandeur physique mesurable correspondant à un opérateur \mathfrak{A} complet et à valeurs propres non dégénérées fournisse à l'instant t une certaine valeur propre de cet opérateur est égale au carré du module du coefficient de la fonction propre correspondante dans le développement de la fonction d'onde Ψ suivant les fonctions propres orthonormales de l'opérateur \mathfrak{A} .

En d'autres termes, si le développement du Ψ suivant les fonctions propres φ_i de l'opérateur complet à valeurs propres non dégénérées \mathfrak{A} est de la forme

$$\Psi = \sum_i c_i(t) \varphi_i,$$

c'est $|c_i(t)|^2$ qui donne la probabilité de la valeur α_i à l'instant t .

On vérifie aisément que, le Ψ étant normé, la probabilité totale de toutes les hypothèses possibles, c'est-à-dire $\sum_i |c_i|^2$ est toujours égale à 1.

Nous avons introduit dans le précédent énoncé les conditions restrictives que l'opérateur \mathfrak{A} est complet et à valeurs propres non dégénérées. Il est facile de lever ces restrictions de la façon suivante :

Si l'opérateur \mathfrak{A} admet des valeurs propres multiples, l'énoncé du second principe doit être modifié comme il suit. Soit α_i une valeur propre multiple à laquelle correspondent p fonctions propres $\varphi_{i,1}, \varphi_{i,2}, \dots, \varphi_{i,p}$ normées et orthogonales, linéairement indépendantes. La probabilité de trouver la valeur α_i par la mesure de la grandeur en question doit alors être prise égale à

$$|c_{i,1}|^2 + |c_{i,2}|^2 + \dots + |c_{i,p}|^2 = \sum_k^p |c_{i,k}|^2,$$

c'est-à-dire que l'on doit bloquer ensemble les carrés des modules des coefficients de $\varphi_{i,1}, \dots, \varphi_{i,p}$ dans le développement du Ψ suivant les fonctions propres de \mathfrak{A} . La probabilité ainsi obtenue se trouve être indépendante de la façon, en partie arbitraire, dont on peut choisir les

fonctions propres orthonormales $\varphi_{i,k}$ correspondant à α_i , ce qui évidemment devait être.

Quand l'opérateur \mathbf{A} est incomplet, l'énoncé du second principe doit subir une autre modification. Dans ce cas, en effet, les fonctions propres de \mathbf{A} ne contiennent pas toutes les variables du domaine D et, par suite, dans le développement $\Psi = \sum_i c_i \varphi_i$, les c_i sont en général des fonctions des variables qui ne figurent pas dans les φ_i . La probabilité d'une valeur propre α_i ne peut donc pas être égale à $|c_i|^2$, quantité qui dépendrait encore de certaines variables. Pour obtenir la probabilité de α_i , il faudra intégrer $|c_i|^2$ pour toutes les valeurs des variables du domaine D qui ne figurent pas dans φ_i . Ainsi, si \mathbf{A} n'opère pas sur x , mais seulement sur y et z , les φ_i seront fonctions seulement de y et z et les c_i dépendront de x : la probabilité de la valeur propre α_i sera alors, non pas $|c_i|^2$, mais $\int_D |c_i(x)|^2 dx$. On vérifie qu'avec cette définition des probabilités, la somme des probabilités de toutes les hypothèses possibles est bien égale à l'unité.

Les deux principes que nous venons d'énoncer permettent, pour tout état du système représenté par une fonction Ψ connue, d'attribuer aux valeurs de toute grandeur physique mesurable une certaine répartition de probabilité traduisant le fait que la probabilité est nulle pour toute valeur différente des α_i et a une valeur, en général, différente de zéro et calculable à partir du Ψ pour chaque α_i . On peut donc appliquer à ces répartitions de probabilité les résultats généraux qui sont classiques en Calcul des Probabilités. C'est ainsi qu'on peut les définir à l'aide d'une « fonction caractéristique ». Dans sa Thèse de Doctorat, M. Arnous a donné la forme de la fonction caractéristique qui traduit à la fois les deux principes énoncés plus haut, et il a ainsi obtenu une très élégante présentation des résultats essentiels de la Mécanique ondulatoire.

Connaissant la loi de répartition de la probabilité pour une grandeur A dans un état représenté par une fonction Ψ , on peut naturellement définir une « dispersion » σ_A au sens du Calcul des Probabilités qui est, par définition, la racine carrée de la valeur moyenne du carré de l'écart par rapport à la valeur probable

$$\sigma_A = \sqrt{(\overline{A - A})^2}.$$

4. Applications des principes généraux. — Les principes généraux que nous venons d'énoncer sont, en somme, la traduction mathé-

matique des idées que nous avons développées au paragraphe 1. Nous allons montrer que ces principes généraux nous permettent de retrouver les deux principes particuliers que nous avons rencontrés au chapitre X en cherchant à interpréter physiquement la Mécanique ondulatoire.

Considérons d'abord l'application des principes généraux à l'opérateur hamiltonien \mathfrak{H} d'un système qui est, nous le savons, un opérateur complet. Pour définir les valeurs propres de \mathfrak{H} , nous devons écrire l'équation

$$\mathfrak{H} \varphi = E \varphi,$$

en désignant ici par E la constante appelée α dans le cas général. On a des valeurs propres E_i et des fonctions propres Ψ_i . Nos principes généraux indiquent alors qu'une mesure précise de l'énergie ne peut conduire qu'à lui attribuer l'une des valeurs E_i (ce qui justifie la façon dont nous avons opéré la quantification en Mécanique ondulatoire) et, de plus, que la probabilité d'obtenir la valeur E_k est égale au carré du module du coefficient de la fonction Ψ_k dans le développement

$$\Psi = \sum_i c_i \Psi_i,$$

suivant les fonctions propres de \mathfrak{H} . Or c'est là précisément le principe de décomposition spectrale ou principe de Born énoncé au chapitre X, paragraphe 4.

Plus délicate, au point de vue de la rigueur mathématique, est la déduction du principe fondamental des interférences donné au paragraphe 2 du chapitre X. Cherchons à appliquer nos principes généraux à la coordonnée x du corpuscule, à laquelle correspond l'opérateur x . « multiplication par x ». L'équation aux valeurs propres est ici

$$x \cdot \varphi = \alpha \varphi.$$

Avec un peu d'indulgence au point de vue de la rigueur mathématique, on peut considérer cette équation comme satisfaite par la « fonction singulière de Dirac » $\delta(x - \alpha)$ qui jouit des propriétés suivantes :

1° C'est une fonction paire de l'argument $x - \alpha$;

2° L'intégrale $\int_a^b f(x) \delta(x - \alpha) dx$ est nulle si l'intervalle d'intégration ne contient pas α et est égale à 1 si cet intervalle contient α , et ceci quelles que soient les valeurs de a et de b .

Nous avons donc ici comme valeurs propres toutes les valeurs réelles de $-\infty$ à $+\infty$. D'après le premier principe, une mesure de la coordonnée x peut *a priori* nous donner toutes les valeurs de $-\infty$ à $+\infty$. Pour trouver les probabilités respectives de ces valeurs, on remarque qu'en vertu des propriétés de la fonction δ de Dirac, on a

$$\Psi(x, y, z, t) = \int_{-\infty}^{+\infty} \Psi(\alpha, y, z, t) \delta(x - \alpha) d\alpha.$$

Le second principe, appliqué à l'opérateur incomplet x , nous indique donc que la probabilité pour que la coordonnée x soit trouvée avoir la valeur α à l'instant t est

$$\int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dz |\Psi(\alpha, y, z, t)|^2.$$

En raisonnant de même pour les coordonnées y et z , on voit aisément que $|\Psi(\alpha, \beta, \gamma, t)|^2$ est la probabilité pour que la mesure de x, y, z donne respectivement les valeurs α, β, γ . Donc la probabilité pour qu'une expérience permette de localiser le corpuscule à l'instant t dans un élément de volume $d\tau$ est $|\Psi(x, y, z, t)|^2 d\tau$. C'est bien le principe de localisation que nous avons appelé « principe des interférences ».

§. **Mesure simultanée de deux grandeurs.** — Des principes généraux adoptés, nous pouvons déduire une conséquence très importante : deux grandeurs A et B ne peuvent être mesurées simultanément avec précision que si les opérateurs correspondants commutent, c'est-à-dire si $\mathbf{AB} \equiv \mathbf{BA}$.

En effet, si φ_i et χ_j sont les fonctions propres des opérateurs \mathbf{A} et \mathbf{B} , α_i et β_j leurs valeurs propres, pour qu'une mesure simultanée des grandeurs A et B puisse se faire avec précision, il faut qu'après la mesure on puisse attribuer simultanément avec certitude à la grandeur A une certaine valeur α_i et à la grandeur B une certaine valeur β_j . D'après le second principe, il faut donc qu'après la mesure, on puisse mettre simultanément la fonction d'onde sous les deux formes

$$\Psi = c_i \varphi_i, \quad \Psi = d_j \chi_j, \quad .$$

avec

$$|c_i| = |d_j| = 1.$$

De ces relations, on déduit

$$\begin{aligned} \mathbf{AB}\Psi &= \mathbf{AB} d_j \chi_j = \mathbf{A} d_j \beta_j \chi_j = \beta_j \mathbf{A} c_i \varphi_i = \beta_j \alpha_i \Psi, \\ \mathbf{BA}\Psi &= \mathbf{BA} c_i \varphi_i = \mathbf{B} c_i \alpha_i \varphi_i = \alpha_i \mathbf{B} d_j \chi_j = \alpha_i \beta_j \Psi. \end{aligned}$$

On a donc

$$\mathfrak{A}\mathfrak{B}\Psi = \mathfrak{B}\mathfrak{A}\Psi,$$

et cela doit être vrai pour toutes les valeurs de i et de j , ce qui entraîne la relation opératorielle

$$\mathfrak{A}\mathfrak{B} \equiv \mathfrak{B}\mathfrak{A}.$$

Donc, si \mathfrak{A} et \mathfrak{B} ne commutent pas, les grandeurs ne sont pas simultanément mesurables avec précision : l'une au moins est affectée d'incertitudes. L'exemple le plus simple et le plus important de deux grandeurs non simultanément mesurables est celui d'une coordonnée q et de la composante conjuguée de la quantité de mouvement p . Ici l'on a

$$\mathfrak{A} = q, \quad \mathfrak{B} = -\frac{h}{2\pi i} \frac{\partial}{\partial q}, \quad \mathfrak{A}\mathfrak{B} - \mathfrak{B}\mathfrak{A} = \frac{h}{2\pi i} \left(\frac{\partial}{\partial q} q - q \frac{\partial}{\partial q} \right) = \frac{h}{2\pi i} \neq 0.$$

Donc, les opérateurs ne commutant pas, les grandeurs correspondantes ne peuvent être simultanément mesurées avec précision. Une coordonnée et le moment conjugué de Lagrange ne peuvent donc jamais être connus qu'avec des incertitudes Δx et Δp_x qui ne peuvent être nulles toutes les deux à la fois. En appliquant les principes généraux, on peut se rendre compte qu'on a toujours en ordre de grandeur

$$\Delta x \Delta p_x \geq h.$$

On retrouve les incertitudes d'Heisenberg.

Nous avons dit que l'on peut définir, à partir des répartitions de probabilité, des dispersions au sens du Calcul des Probabilités. On peut démontrer que le produit des dispersions σ_x et σ_{p_x} sur x et sur p_x satisfait à l'inégalité

$$\sigma_x \sigma_{p_x} \geq \frac{h}{4\pi}.$$

Cette formule est, en un sens, plus précise que les relations d'incertitude d'Heisenberg, mais elle ne leur est pas tout à fait équivalente.

6. Valeur moyenne d'une grandeur en Mécanique ondulatoire. — Nous allons maintenant pouvoir définir la valeur moyenne d'une grandeur mesurable attachée à un corpuscule dont on connaît l'état Ψ . Comme il y a en général pour la grandeur plusieurs valeurs possibles dont les probabilités ne sont pas nulles, on ne peut plus ici parler sans équivoque de la valeur de la grandeur à chaque instant, mais on peut toujours définir sa « valeur moyenne » en entendant par là, comme

d'habitude, la somme des produits de chaque valeur possible par la probabilité correspondante (espérance mathématique).

Supposons que nous ayons un très grand nombre N de corpuscules dans des états identiques, c'est-à-dire ayant même fonction Ψ et qu'à l'instant t nous mesurons simultanément la valeur de la grandeur envisagée pour les N corpuscules. Ces mesures ne fourniront pas, en général, la même valeur puisqu'il y a plusieurs valeurs possibles pour la grandeur. Si \mathfrak{A} est l'opérateur correspondant à la grandeur, α_i et φ_i ses valeurs et fonctions propres à l'instant t , chacune des mesures pourra, d'après le premier principe général, fournir l'un des α_i : n_1 des N mesures fourniront α_1 , n_2 fourniront α_2 , etc. D'après le second principe général, si le développement du Ψ commun aux N corpuscules suivant les φ_i est de la forme $\Psi = \sum_i c_i \varphi_i$, les nombres n_i sont respectivement proportionnels aux $|c_i|^2$. La moyenne des valeurs fournies par les N expériences de mesure simultanées est donc

$$\bar{A} = \sum_i \alpha_i |c_i|^2;$$

c'est là la valeur moyenne de la grandeur A pour un corpuscule dans l'état Ψ . On peut l'écrire aussi sous la forme

$$\bar{A} = \int_{\mathfrak{D}} \Psi^* \mathfrak{A} \Psi d\tau,$$

comme on le vérifie facilement en remplaçant Ψ et Ψ^* par leurs développements suivant les φ_i et les φ_i^* et en tenant compte de l'orthonormalité des fonctions propres.

La définition de \bar{A} fournit une intéressante interprétation statistique des éléments diagonaux des matrices d'Heisenberg. Supposons, en effet, que le développement du Ψ se réduise à un seul terme, c'est-à-dire que l'état envisagé soit un cas pur pour la grandeur A . On a alors

$$\Psi = c_i \varphi_i \quad \text{avec} \quad |c_i| = 1.$$

Nous sommes alors sûrs que la mesure de A fournira la valeur α_i . Soit maintenant une autre grandeur B attachée au corpuscule et correspondant à un opérateur \mathfrak{B} . En général, l'état Ψ n'est pas un cas pur pour B et cette grandeur n'a pas de valeur certaine, mais on peut toujours calculer sa valeur moyenne qui est

$$B = \int_{\mathfrak{D}} \Psi^* \mathfrak{B} \Psi d\tau = \int_{\mathfrak{D}} \varphi_i^* \mathfrak{B} \varphi_i d\tau.$$

Or, la seconde intégrale n'est pas autre chose que l'élément diagonal B_{ii} de la matrice engendrée par l'opérateur \mathfrak{B} dans le système des fonctions propres φ_i .

D'où le théorème :

L'élément diagonal d'indices ii de la matrice engendrée par l'opérateur \mathfrak{B} dans le système des fonctions propres de l'opérateur \mathfrak{A} est égal à la valeur moyenne de la grandeur B quand on sait que la grandeur A a la valeur α_i .

7. Les intégrales premières en Mécanique ondulatoire. — En Mécanique classique, on nomme *intégrale première* une grandeur mécanique s'exprimant à l'aide des coordonnées, des moments et parfois du temps, qui reste constante en vertu même des équations du mouvement. En Mécanique ondulatoire, on ne peut pas conserver la même définition parce que les grandeurs n'y ont plus, en général, une valeur bien déterminée. On est alors amené à dire qu'en Mécanique ondulatoire, une grandeur mesurable est « intégrale première » ou « constante du mouvement » quand ses valeurs possibles et leurs probabilités respectives sont constantes au cours du temps. On démontre que la condition pour que la grandeur correspondant à un opérateur \mathfrak{A} soit intégrale première, dans un problème où l'opérateur hamiltonien est \mathfrak{H} , est

$$\frac{\partial \mathfrak{A}}{\partial t} + \frac{2\pi i}{h} (\mathfrak{A}\mathfrak{H} - \mathfrak{H}\mathfrak{A}) \equiv 0.$$

Ici $\frac{\partial \mathfrak{A}}{\partial t}$ est l'opérateur obtenu en dérivant formellement l'expression de l'opérateur \mathfrak{A} par rapport au paramètre t . Dans le cas très fréquent où \mathfrak{A} ne dépend pas de t , la condition précédente exprime simplement que \mathfrak{A} commute avec \mathfrak{H} .

Si le champ est permanent, \mathfrak{H} est indépendant du temps et, comme il commute avec lui-même, l'énergie est intégrale première. On retrouve l'analogie du théorème classique sur les systèmes conservatifs.

Si la composante x du champ est nulle, \mathfrak{H} ne dépend pas de x (car $\frac{\partial V}{\partial x} = 0$) et, par suite, commute avec $p_x \sim \frac{\partial}{\partial x}$: donc la composante x de la quantité de mouvement est alors une intégrale première, résultat encore analogue à un théorème classique.

L'exemple le plus intéressant est celui du moment de la quantité de mouvement (ou moment de rotation). Quand le champ est partout dans un même plan avec un axe Oz , \mathfrak{H} ne dépend pas de l'azimut φ

autour de cet axe (car $\frac{\partial V}{\partial \varphi} = 0$). En prenant des axes directs et en convenant qu'à une rotation de sens positif dans le plan xOy correspond un moment de rotation positif, l'opérateur qui correspond au moment de rotation autour de Oz est

$$\mathfrak{M}_z = x \cdot p_y - y \cdot p_x = \frac{h}{2\pi i} \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right).$$

En prenant des coordonnées sphériques avec Oz comme axe polaire, on trouve

$$(3) \quad \mathfrak{M}_z = - \frac{h}{2\pi i} \frac{\partial}{\partial \varphi},$$

formule qui traduit le fait que M_z est le moment de Lagrange conjugué de φ .

Par suite, dans le cas envisagé, l'opérateur \mathfrak{M}_z indépendant du temps commute avec \mathfrak{H} : la grandeur M_z est donc intégrale première. Si le champ est radial par rapport à l'origine O , M_x , M_y et M_z sont toutes trois intégrales premières.

On peut définir un opérateur correspondant au carré du moment de rotation total en posant

$$\mathfrak{M}^2 = \mathfrak{M}_x^2 + \mathfrak{M}_y^2 + \mathfrak{M}_z^2.$$

On trouve aisément en coordonnées sphériques

$$(4) \quad \mathfrak{M}^2 = - \frac{h^2}{4\pi^2} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right]$$

et l'on démontre que, si le champ est radial, M^2 est intégrale première en appelant M^2 la grandeur $M^2 = M_x^2 + M_y^2 + M_z^2$ à laquelle correspond l'opération \mathfrak{M}^2 .

8. La quantification dans l'espace. — Dans l'ancienne théorie des quanta, on avait été amené à l'énoncé suivant :

Dans un système quantifié, comme par exemple l'atome d'hydrogène, la projection du moment de quantité de mouvement sur la direction d'un champ extérieur (champ magnétique dans l'effet Zeeman, par exemple), est toujours égale à $m \frac{h}{2\pi}$, m étant un nombre entier positif ou négatif.

Comme, d'autre part, la longueur du vecteur \vec{M} devait être égale à $k \frac{h}{2\pi}$, k étant le nombre quantique azimutal, on en concluait que

l'angle θ entre \vec{M} et la direction du champ devait être tel que $\cos\theta = \frac{m}{k}$, m variant par valeurs entières de $-k$ à $+k$. C'est ce que montrait clairement le schéma vectoriel suivant (*fig. 28*).

Le vecteur \vec{M} et, par suite, le plan de la trajectoire électronique qui lui est normal dans le cas de l'atome d'hydrogène, ne pouvait donc prendre que certaines orientations quantifiées par rapport à Oz : on désignait cette circonstance sous le nom de « quantification dans l'espace ».

Or, à y regarder de près, l'idée de quantification dans l'espace soulevait de grandes difficultés. La direction Oz est bien définie s'il existe un champ (magnétique ou électrique) dans lequel l'atome est plongé.

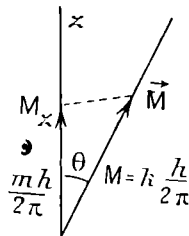


Fig. 28.

Mais la quantification dans l'espace doit exister si faible que soit le champ extérieur. Qu'arrive-t-il quand le champ tend vers zéro ? A la limite, l'axe Oz devient arbitraire et l'on ne peut évidemment pas admettre qu'il y ait quantification dans l'espace par rapport à n'importe quel axe Oz issu de O . De plus, s'il y a plusieurs champs agissant sur le système quantifié comme c'est le cas pour un atome plongé dans un champ magnétique et dans un champ électrique qui peuvent faire entre eux un angle quelconque, il y a plusieurs directions Oz privilégiées et l'on ne sait par rapport à laquelle on doit quantifier. Les conceptions probabilistes de la nouvelle Mécanique ont levé toutes ces difficultés en montrant que la représentation du moment de quantité de mouvement par un vecteur a quelque chose de trompeur.

En effet, si l'on forme l'expression des opérateurs correspondant à M_x , M_y , M_z ,

$$\mathfrak{M}_x = \frac{h}{2\pi i} \left(z \frac{\partial}{\partial y} - y \frac{\partial}{\partial z} \right), \quad \mathfrak{M}_y = \frac{h}{2\pi i} \left(x \frac{\partial}{\partial z} - z \frac{\partial}{\partial x} \right),$$

$$\mathfrak{M}_z = \frac{h}{2\pi i} \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right),$$

on constate que ces opérateurs ne commutent pas, car l'on a

$$\mathfrak{M}_x \mathfrak{M}_z - \mathfrak{M}_z \mathfrak{M}_x = \frac{h}{2\pi i} \mathfrak{M}_y \neq 0$$

et deux relations analogues obtenues par permutation circulaire sur x , y , z . Donc, sauf dans le cas où M_x , M_y , M_z sont tous trois nuls, on ne peut simultanément mesurer les trois composantes de \vec{M} et, par suite, connaître ce vecteur. La représentation du moment de rotation par un vecteur utilisé dans la figure faite plus haut a donc quelque chose de fallacieux et il ne faut plus s'étonner qu'elle conduise à des contradictions.

Mais rien n'empêche de mesurer *l'une* des composantes du moment de rotation, M_z par exemple. D'après l'expression (3) de \mathfrak{M}_z obtenue plus haut, les valeurs propres de la grandeur M_z sont déterminées par l'équation

$$-\frac{h}{2\pi i} \frac{\partial f}{\partial \varphi} = \alpha f,$$

où f est la fonction propre et φ l'azimut autour de Oz . Nous obtenons pour les valeurs et fonctions propres

$$\alpha = m \frac{h}{2\pi}, \quad f = \frac{e^{-im\varphi}}{\sqrt{2\pi}},$$

où m est un entier positif ou négatif. Donc, d'après le premier principe général, si nous parvenons par une mesure à attribuer une valeur à la composante z du moment de rotation, nous trouverons un multiple entier, positif ou négatif, de $\frac{h}{2\pi}$, ce qui est en accord avec l'ancienne théorie des Quanta.

Comme nous l'avons vu dans le cas de l'atome d'hydrogène, les fonctions propres de l'opérateur \mathfrak{H} sont les suivantes :

$$\Psi_{nlm} = F_{nlm}(r, \theta) e^{-im\varphi} \quad (-l \leq m \leq l).$$

Supposons que la fonction Ψ de l'atome se réduise à Ψ_{nlm} . Nous aurons alors un cas pur pour le moment de rotation autour de Oz , c'est-à-dire qu'une mesure de cette grandeur nous donnera nécessairement la valeur $m \frac{h}{2\pi}$. Mais qu'arriverait-il si, au lieu de mesurer le moment de rotation autour de Oz , nous mesurons le moment de rotation autour d'un autre axe Oz' ? En répétant le raisonnement fait ci-dessus, on voit

facilement que l'on doit aussi lui trouver une valeur égale à un nombre entier de fois $\frac{h}{2\pi}$, ce qui est encore le résultat de l'ancienne théorie des Quanta. Mais dans cette ancienne théorie qui admettait les idées de la Physique classique, on arrivait ainsi à la conclusion inadmissible que M_z et $M_{z'}$ devaient tous deux être égaux à un multiple entier de $\frac{h}{2\pi}$, quel que fut l'angle des axes Oz et Oz' . Avec notre nouveau point de vue, le paradoxe est écarté, parce que les mesures de M_z et de $M_{z'}$ sont incompatibles et que, si l'on mesure par exemple M_z , il y aura nécessairement une incertitude après la mesure sur la valeur de $M_{z'}$.

Précisons ce dernier point. Dans le système de coordonnées polaires r, θ', φ' dont Oz' est l'axe polaire, les fonctions propres de l'opérateur \mathfrak{H} sont de la forme

$$\Psi'_{n'l'm'} = F_{n'l'm'}(r, \theta') e^{-im'\varphi'}$$

et dans le cas envisagé plus haut où l'on a $\Psi = F_{nlm}(r, \theta) e^{-im\varphi}$, Ψ pourra se développer suivant le système complet des $\Psi'_{n'l'm'}$, puisque ces dernières fonctions forment un système complet. On trouve un développement de la forme

$$\Psi = F_{nlm}(r, \theta) e^{-im\varphi} = \sum_{m'}^{+l} d_{m'} F_{nlm'}(r, \theta') e^{-im'\varphi'}.$$

L'état défini par la fonction d'onde Ψ , qui était un cas pur pour le moment de rotation autour de Oz , n'est pas un cas pur pour le moment de rotation autour de Oz' . La mesure de $M_{z'}$ pourra donner diverses valeurs de la forme $m' \frac{h}{2\pi}$ avec des probabilités respectives proportionnelles aux $|d_{m'}|^2$. La mesure initiale de M_z qui avait permis de donner au Ψ la forme de cas pur pour M_z que nous avons adoptée, comportait donc, conformément aux conclusions ci-dessus exposées, une incertitude sur la valeur de $M_{z'}$.

Terminons en disant quelques mots de l'opérateur \mathfrak{M}^2 défini au paragraphe 7 par la formule (4). Il lui correspond l'équation aux valeurs propres

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 f}{\partial \varphi^2} + \frac{4\pi^2}{h^2} \alpha f = 0.$$

Nous avons déjà rencontré une équation de cette forme dans la

théorie du rotateur sphérique (Chap. XI, § 3) et nous avons vu alors que ses valeurs propres sont

$$K = \frac{4\pi^2}{h^2} \alpha = l(l+1) \quad (l = 0, 1, 2, \dots),$$

ce qui nous donne

$$\alpha_l = l(l+1) \frac{h^2}{4\pi^2}.$$

Dans l'ancienne théorie des Quanta où $|\vec{M}| = k \frac{h}{2\pi}$, on trouvait pour les valeurs possibles de $M^2 : \frac{k^2 h^2}{4\pi^2}$, avec $k = 1, 2, \dots$. On voit que la divergence entre les deux théories est importante, notamment en ceci que, dans la nouvelle théorie, la plus petite valeur de M^2 est zéro, alors que dans l'ancienne théorie elle était $\frac{h^2}{4\pi^2}$. Ceci correspond au fait qu'en Mécanique ondulatoire, les états s sont isotropes.

Notons enfin que, l'opérateur \mathbf{M}^2 commutant avec \mathbf{M}_x , \mathbf{M}_y et \mathbf{M}_z , on peut mesurer simultanément M^2 et l'une des composantes M_x , M_y , M_z .

BIBLIOGRAPHIE.

- [1] P. A. M. DIRAC, *Les principes de la Mécanique quantique*, Presses universitaires, Paris, 1931.
- [2] LOUIS DE BROGLIE, *La quantification dans la nouvelle Mécanique*, Hermann, Paris, 1932.
- [3] LOUIS DE BROGLIE, *Théorie générale des particules à spin*, Gauthier-Villars, Paris, 1943.
- [4] E. ARNOUS, *Les lois de probabilité en Mécanique ondulatoire* (Thèse de Doctorat), Centre de documentation universitaire Tournier et Constant, Paris, 1946.

CHAPITRE XIV.

LE SPIN DE L'ÉLECTRON. LA THÉORIE DE DIRAC.

1. Nécessité d'introduire un élément supplémentaire dans la définition de l'électron. — Nous allons pour un instant revenir un peu en arrière et nous replacer dans l'état d'esprit de l'ancienne théorie des Quanta, antérieur à l'éclosion de la nouvelle Mécanique, et nous allons voir que, même en restant dans ce cadre ancien, on peut apercevoir la nécessité d'introduire un élément nouveau dans la définition de l'électron. C'est d'ailleurs ce dont certains physiciens s'étaient bien aperçus dès avant l'apparition de la nouvelle Mécanique.

Deux catégories de faits inexplicables suggéraient principalement cette nécessité : des faits spectroscopiques et des faits magnétiques. Nous allons les résumer :

a. Anomalies spectroscopiques. — L'ancienne théorie des Quanta avait conduit, après le travail de Sommerfeld, à classer les états stationnaires d'énergies distinctes des atomes à l'aide de deux nombres quantiques n et k , prenant les valeurs $n = 1, 2, 3, \dots$ et $k = 1, 2, \dots, n$. On prévoyait ainsi un niveau d'énergie minimum $n = 1, k = 1$: c'est le niveau **K** des spectres de rayons X. Ensuite doivent venir deux niveaux $n = 2, k = 1, 2$, niveaux **L** des spectres X, puis trois niveaux $n = 3, k = 1, 2, 3$ ou niveaux **M**, etc. D'une façon générale, k pouvant varier de 1 à n pour un nombre n donné, on trouvait n niveaux correspondant au nombre quantique total et formant la « couche » définie par ce nombre quantique. Tel était le résultat essentiel de la théorie des Sommerfeld.

Or, l'expérience prouve bien que le nombre des niveaux par couche augmente avec le nombre n , mais en réalité il y a plus de niveaux, la structure fine est plus riche que ne le prévoit la théorie précédente. Ainsi, dans le spectre de rayons X, on reconnaît bien un seul niveau **K**, mais on identifie, très nettement séparés, trois niveaux **L**, cinq niveaux **M**,

etc. D'une façon générale, l'expérience indique ainsi qu'il existe, pour un nombre n donné, $2n - 1$ niveaux distincts et non pas n . On peut interpréter empiriquement ce résultat en disant que les niveaux prévus par la théorie des Quanta sont tous dédoublés, sauf ceux pour lesquels $k = 1$. Ainsi le niveau K, ainsi que le niveau L le plus profond (L_I), pour lesquels $k = 1$, restent simples; mais le niveau L le moins profond prévu par la théorie de Sommerfeld et pour lequel $k = 2$ est en réalité dédoublé en deux niveaux distincts L_{II} et L_{III} . Pour une couche définie par le nombre quantique total n , l'un des niveaux ($k = 1$) reste simple et tous les autres ($k = 2, \dots, n$) sont ainsi dédoublés, de sorte qu'au total la couche comprend bien $1 + 2(n - 1) = 2n - 1$ niveaux comme l'indique l'expérience.

En présence de ces faits, Sommerfeld a cherché à spécifier chaque niveau pour un ensemble de nombres quantiques en introduisant assez arbitrairement, à côté des nombres n et k de sa théorie, un troisième nombre quantique j qui pour $k = 1$ ne pourrait prendre qu'une seule valeur, mais qui, pour $k > 1$, pourrait en prendre deux. Il suffit, par exemple, pour cela d'admettre que le nombre quantique peut prendre les valeurs $k - \frac{1}{2}$ et $k - \frac{3}{2}$, mais qu'il est nécessairement positif, ce qui exclut la deuxième valeur dans le cas $k = 1$. Guidé par des idées classiques qui n'ont pas été confirmées par la suite, Sommerfeld avait donné à ce nombre j , dont l'intervention restait assez mystérieuse, le nom de « nombre quantique interne ».

Des circonstances analogues se rencontrent dans l'étude du spectre optique des atomes alcalins. Les états $k = 1$ de ces spectres appelés « états s » sont simples, tandis que les autres états (tels que les états $k = 2$ ou états p , les états $k = 3$ ou états d , etc.) sont tous doubles. De là vient que le spectre des alcalins est un spectre de doublets. Pour les éléments autres que les alcalins, on rencontre des apparences plus compliquées encore avec spectre de triplets, de quadruplets, etc. Nous n'y insisterons pas ici.

Nous avons déjà vu au chapitre VII qu'une autre anomalie s'est présentée en ce qui concerne la formule de structure fine de Sommerfeld. C'est, en effet, dans les spectres de rayons X, entre les niveaux L_{II} et L_{III} que l'on observe le doublet prévu par Sommerfeld dont la largeur augmente comme Z^4 . Ceci paraissait au début bien cadrer avec la théorie de Sommerfeld, car on attribuait alors aux niveaux L_{II} et L_{III} les nombres quantiques $n = 2, k = 1$ et $n = 2, k = 2$ respectivement et la théorie de Sommerfeld prévoit l'existence des doublets en Z^4 entre les niveaux

dont les nombres k diffèrent d'une unité. Or, les considérations théoriques sur le nombre j et l'analogie des spectres de rayons X avec les spectres de doublets des alcalins n'ont pas tardé à indiquer clairement que l'on devait faire l'attribution suivante des nombres quantiques aux niveaux L

$$\begin{aligned} L_I : \quad n = 2, \quad k = 1, \quad j = \frac{1}{2}; \\ L_{II} : \quad n = 2, \quad k = 2, \quad j = \frac{1}{2}; \\ L_{III} : \quad n = 2, \quad k = 2, \quad j = \frac{3}{2} \end{aligned}$$

et ceci ne va plus du tout avec la formule de Sommerfeld qui prévoit le doublet de relativité entre le niveau L_I et l'un des niveaux L_{II} et L_{III} , ce qui est contraire à l'expérience. Ajoutons qu'en examinant avec soin le spectre à structure fine beaucoup serrée de l'atome d'hydrogène, on a pu reconnaître que le même genre de difficultés se présentait même dans l'interprétation des doublets de la série de Balmer. On a alors pu croire que le succès de la théorie de Sommerfeld était fortuit : il n'en est rien, mais cette théorie qui a le mérite d'introduire les considérations relativistes reste incomplète parce qu'elle n'introduit pas le *spin*.

b. Anomalies magnétiques. — A côté des anomalies spectroscopiques, on a vu se manifester des anomalies magnétiques que nous allons étudier. Commençons par les *anomalies gyromagnétiques*.

Pour bien comprendre la nature de ces anomalies, il faut d'abord se souvenir de la relation qui existe entre le moment magnétique et le moment de rotation correspondant à une orbite électronique. Considérons un électron décrivant une orbite fermée : il est équivalent à un courant qui, lui-même, est équivalent à un feuillet dont il est facile d'évaluer le moment magnétique, Soient $\vec{\mathfrak{M}}$ ce moment magnétique et \mathbf{M} le moment de rotation de l'électron sur sa trajectoire, moment supposé constant. Si $-e$ est la charge de l'électron en u. e. s., m_0 sa masse, c la vitesse de la lumière dans le vide, il est aisé de montrer que

$$(1) \quad \vec{\mathfrak{M}} = - \frac{e}{2m_0c} \mathbf{M}.$$

Einstein a démontré que cette formule est encore valable pour un ensemble d'électrons entièrement contenu dans un petit volume. Elle doit donc être applicable à tout atome du type Rutherford-Bohr.

Or, d'après la théorie des Quanta, le moment de rotation $\vec{\mathfrak{M}}$ doit être égal à un multiple entier de $\frac{h}{2\pi}$. On en conclut que l'on a

$$\mathfrak{M} = k \frac{eh}{4\pi m_0 c} \quad (k \text{ entier}).$$

Donc le moment magnétique d'un atome doit toujours être un multiple entier du « magnéton de Bohr » égal à $\frac{eh}{4\pi m_0 c}$, à peu près cinq fois plus petit que le magnéton de Weiss dont la signification théorique est douteuse.

L'existence réelle du magnéton de Bohr a été prouvée par l'expérience de Stern et Gerlach, mais néanmoins la relation prévue entre $\vec{\mathfrak{M}}$ et \mathbf{M} ne s'est pas vérifiée d'une façon générale. La théorie indique, en effet, qu'un barreau magnétique doit entrer en rotation si on le soumet à un champ magnétique et que le rapport entre son moment magnétique et son moment de rotation doit être donné par (1); mais si le phénomène a bien lieu (expériences d'Einstein et de De Haas), le rapport des deux moments est plus voisin de $\frac{e}{m_0 c}$ que $\frac{e}{2m_0 c}$. Barnett a retrouvé le même rapport anormal en étudiant le phénomène inverse : création d'un moment magnétique par rotation d'un barreau dans un champ magnétique. L'idée, classique depuis Ampère, suivant laquelle le magnétisme a pour origine des mouvements de particules (électrons) dans la matière, se trouve donc ne pas représenter l'ensemble des faits.

Les mêmes difficultés se présentent sous une autre forme dans la théorie de l'effet Zeeman. Nous savons que la théorie classique de l'effet Zeeman conduit à la conclusion suivante : si un atome émet, en dehors de tout champ, une radiation de fréquence ν , placé dans un champ magnétique uniforme, il peut émettre les trois raies

$$\nu - \frac{eH}{4\pi m_0 c}, \quad \nu, \quad \nu + \frac{eH}{4\pi m_0 c},$$

e et m_0 étant les constantes classiques de l'électron. Cette prévision de la théorie classique, ainsi que les prévisions relatives à l'état de polarisation des raies, avaient été vérifiées par l'expérience et constituaient l'un des grands succès des conceptions de Lorentz. La théorie des Quanta n'avait rien introduit d'essentiellement nouveau quant à la prévision de l'effet Zeeman. Soit un électron ayant un état stationnaire dans un atome en l'absence de champ magnétique extérieur. Désignons par W_0 l'énergie

de cet électron et par $\vec{\mathcal{M}}$ le moment magnétique de son orbite. En présence du champ extérieur, l'énergie de l'électron sera, d'après (1),

$$W_{\text{II}} = W_0 - (\vec{\mathcal{M}} \cdot \mathbf{H}) = W_0 + \frac{e}{2m_0c} (\mathbf{M} \cdot \mathbf{H}).$$

Comme, d'après la théorie des Quanta, la composante de \mathbf{M} dans la direction de \mathbf{H} est égale à $m \frac{h}{2\pi}$ avec m entier, on a

$$W_{\text{II}} = W_0 + m \frac{eh}{4\pi m_0c} \text{II},$$

m est appelé le nombre quantique magnétique : il correspond à la quantification dans l'espace du moment de rotation \mathbf{M} par rapport à la direction de \mathbf{H} .

Si $\nu_0 = \frac{W_0 - W'_0}{h}$ désigne la fréquence de la raie émise en l'absence de champ lors de la transition ($W_0 \rightarrow W'_0$), cette raie aura pour fréquence en présence du champ II

$$\nu_{\text{II}} = \frac{W_{\text{II}} - W'_{\text{II}}}{h} = \nu_0 + (m - m') \frac{e\text{II}}{4\pi m_0c}.$$

Si l'on admet la règle de sélection $\delta m = 0, \pm 1$, on retrouve l'effet Zeeman normal, la constante h s'étant éliminée dans le calcul.

Malheureusement, l'effet Zeeman normal est très exceptionnel : c'est par un hasard heureux que Zeeman a d'abord étudié un cas où il est réalisé. En général, la décomposition magnétique des raies est très compliquée (effet Zeeman anomal) et Lorentz avait fait inutilement de grands efforts pour parvenir à l'interprétation des effets anomaux. La théorie des Quanta n'a pas non plus apporté la solution du problème puisqu'elle conduit, elle aussi, nous venons de le voir, à l'effet normal. Les physiciens ont fini par se persuader que, pour expliquer l'existence et la complication des effets Zeeman anomaux, il fallait introduire une idée entièrement nouvelle. Cette idée a été celle du « spin » de l'électron; mais, avant d'en parler, nous devons dire encore quelques mots de l'effet Zeeman anomal.

L'effet Zeeman anomal est un phénomène d'apparence très compliquée. Nous nous bornerons à le décrire dans le cas le plus simple, celui des atomes alcalins pour lesquels on peut considérer le spectre visible comme émis par l'électron le plus extérieur, l'électron de valence. Les résultats expérimentaux relatifs à l'effet Zeeman anomal des alcalins peuvent se résumer de la façon suivante : En l'absence de champ,

les énergies des états stationnaires de l'électron de valence ont pour valeurs $E_0(n, l, j)$, chaque état étant défini par les trois nombres quantiques n, l, j (n nombre quantique total, $l = k - 1$, j nombre quantique interne de Sommerfeld égal à $l \pm \frac{1}{2}$). En présence du champ magnétique H , chacun des niveaux d'énergie se décompose en $2j + 1$ niveaux d'énergie

$$W_{II}(n, l, m) = W_0(n, l) + mg \frac{ehH}{4\pi m_0 c},$$

avec

$$(2) \quad g = \frac{2j + 1}{2l + 1},$$

et où m est le nombre quantique magnétique qui peut prendre les $2j + 1$ valeurs demi-entières $-j, -j + 1, \dots, j - 1, j$. Telles sont les formules empiriques qui représentent le dédoublement réellement observé des niveaux par effet Zeeman ; g est le « facteur de Landé » et c'est la présence de ce facteur dans l'expression de W_{II} qui correspond au caractère anormal de l'effet Zeeman. On remarquera que pour les niveaux s pour lesquels $l = 0$, on a $j = \frac{1}{2}$, $g = 2$ et $m = \pm \frac{1}{2}$, de sorte qu'en ce cas l'effet Zeeman est normal.

La formule donnant W_{II} est valable seulement dans le cas des champs faibles, c'est-à-dire des champs tels que le déplacement des termes spectraux par effet Zeeman soit faible par rapport aux écarts des composantes des doublets qui constituent le spectre de l'alcalin en l'absence de champ. Pour des champs plus intenses n'obéissant plus à cette condition, on observe un phénomène plus compliqué représenté par une formule due à Voigt. Enfin, pour des champs très forts, on observe un phénomène simple dans lequel il y a décomposition Zeeman normale par rapport au centre de gravité de chaque doublet. Ni la théorie classique, ni l'ancienne théorie des Quanta ne pouvaient interpréter ces faits.

Tout ce que nous venons de dire concerne le cas relativement simple des alcalins. Dans les autres cas, on peut toujours utiliser la formule de Landé pour W_{II} , mais il faut remplacer l'expression (2) du facteur de Landé par une expression plus générale dont nous parlerons plus loin et dont (2) est un cas particulier.

2. L'hypothèse d'Uhlenbeck et Goudsmit (1925). Le spin de l'électron. — L'ensemble des anomalies que nous venons d'étudier conduisait

naturellement à penser que le magnétisme des systèmes atomiques ne provient pas uniquement des mouvements orbitaux des électrons. En réfléchissant à cette idée, deux physiciens hollandais, Uhlenbeck et Goudsmit, sont parvenus en 1925 à la conception que l'électron devait posséder un magnétisme propre, c'est-à-dire qu'il n'était pas seulement assimilable à une petite charge électrique, mais aussi à un petit aimant. Cherchant alors à se faire une nouvelle représentation de l'électron dans le cadre des images classiques, ils se le sont représenté comme une petite sphère d'électricité négative tournant autour de l'un de ses diamètres et possédant, par suite, à la fois un moment magnétique et un moment de rotation. On appelle cette caractéristique nouvelle de l'électron le Spin (d'un mot anglais signifiant rotation).

D'après cette hypothèse nouvelle, l'électron doit donc posséder un moment de rotation propre M_0 et un moment magnétique \mathfrak{M}_0 étroitement liés l'un à l'autre. Quel rapport doit-on attribuer au rapport $\frac{\mathfrak{M}_0}{M_0}$ des deux moments propres ? Le premier mouvement pourrait être de lui attribuer la valeur classique $\frac{e}{2m_0c}$ correspondant à la formule (1). Mais on retomberait alors sur les conséquences inexactes des théories antérieures et, en particulier, on ne retrouverait toujours que l'effet Zeeman normal. Uhlenbeck et Goudsmit ont alors remarqué que les expériences sur les effets gyromagnétiques avaient donné pour le rapport $\frac{\mathfrak{M}}{M}$ une valeur anormale $\frac{e}{m_0c}$, double de la valeur classique donnée par (1) et ils ont admis que l'électron possédait un « double magnétisme propre », c'est-à-dire que le rapport $\frac{\mathfrak{M}_0}{M_0}$ des moments propres de l'électron avait la valeur $\frac{e}{m_0c}$. Or, d'après l'ancienne théorie des Quanta, le rapport normal $\frac{e}{2m_0c}$ était réalisé pour les moments orbitaux parce que, d'une part, le moment de rotation devait être égal à $k \frac{h}{2\pi}$ et que, d'autre part, le moment magnétique devait être égal à k fois le magnéton de Bohr, soit $k \frac{eh}{4\pi m_0c}$. Pour expliquer la valeur anormale du rapport des moments propres, on pouvait alors envisager deux hypothèses simples : ou bien le moment de rotation propre serait égal à $\frac{h}{2\pi}$, le moment magnétique propre s'élevant à deux magnétons de Bohr ; ou bien, le moment magnétique propre étant égal à un magnéton de Bohr, le moment de rotation propre aurait seulement pour valeur $\frac{h}{4\pi}$. Diverses

considérations ont montré que la seconde hypothèse était la meilleure et Uhlenbeck et Goudsmit ont posé

$$M_0 = \frac{1}{2} \frac{h}{2\pi}, \quad \mathfrak{M}_0 = \frac{eh}{4\pi m_0 c},$$

En se tenant dans le cadre de l'ancienne théorie des Quanta encore presque seule en usage en 1925, ils ont montré, ce qui a été ensuite confirmé par les travaux plus approfondis de Thomas et de Frenkel, qu'avec ces hypothèses il est possible d'interpréter de façon satisfaisante les anomalies spectroscopiques et magnétiques signalées plus haut.

Au point de vue spectroscopique, en effet, on sera amené, pour caractériser les états stationnaires, à introduire le spin de l'électron. Le moment de rotation orbital de l'électron étant égal à $k \frac{h}{2\pi}$, son moment propre $\frac{h}{4\pi}$ pourra s'ajouter ou se retrancher au moment orbital donnant un moment total $\left(k \pm \frac{1}{2}\right) \frac{h}{2\pi}$, ce qui introduit un nouveau nombre quantique j égal à $k \pm \frac{1}{2}$. Maintenant nous savons qu'en Mécanique ondulatoire le moment de rotation orbital n'est pas égal à $k \frac{h}{2\pi}$, mais à $l \frac{h}{2\pi}$ avec $l = k - 1$, ce qui nous conduit pour j aux valeurs $j = l \pm \frac{1}{2}$ ou $k - \frac{1}{2}$ et $k - \frac{3}{2}$ que nous avons précédemment adoptées. L'introduction du nombre quantique « interne » j de Sommerfeld se trouve ainsi recevoir une justification : il correspond à la quantification du moment de rotation total formé par l'addition algébrique du moment orbital et du moment propre. Des calculs plus complets avaient d'ailleurs montré que l'on peut ainsi retrouver la structure fine de Sommerfeld et prévoir les doublets de Relativité là où ils sont réellement, c'est-à-dire entre les niveaux dont les nombres quantiques j diffèrent d'une unité. Mais ces calculs développés dans le cadre de l'ancienne théorie des Quanta n'étaient cependant pas entièrement satisfaisants.

Passons maintenant à l'interprétation des effets Zeeman anomaux. Nous allons employer, pour retrouver la formule de Landé, un raisonnement mi-classique et mi-quantique qui n'est pas très satisfaisant, mais qui montre comment on pouvait dès 1925 parvenir à interpréter l'effet Zeeman anomal. Soit un atome pouvant contenir plusieurs électrons qui contribuent à l'émission des raies spectrales considérées. La somme de leurs moments de rotation orbitaux est quantifiée et a pour valeur $l \frac{h}{2\pi}$,

celle de leurs spins a pour valeur $s \frac{h}{2\pi}$, s se réduisant à $\frac{1}{2}$ dans le cas d'un seul électron et pouvant avoir des valeurs semi-entières plus élevées dans le cas de plusieurs électrons. Nous admettons que le moment de rotation total est quantifié et égal à $j \frac{h}{2\pi}$ et qu'il est formé par l'addition géométrique du moment orbital et du moment de spin suivant le schéma ci-dessous (*fig. 29*).

Sous l'influence du champ magnétique \mathbf{H} , le niveau d'énergie correspondant aux nombres quantiques l, s, j passe de la valeur W_0 qu'il possède en l'absence de champ à la valeur $W_H = W_0 + \Delta W$ et l'on a

$$\Delta W = - \mathfrak{M}_H H.$$

Or \mathfrak{M}_H , projection du moment magnétique sur la direction de \mathbf{H} a,

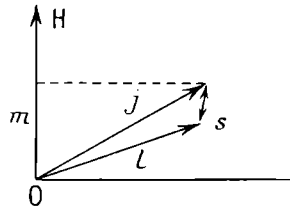


Fig. 29.

en vertu de l'hypothèse du double magnétisme propre de l'électron, la valeur

$$\mathfrak{M}_H = - \frac{eh}{4\pi m_0 c} [l \cos \widehat{lH} + 2s \cos \widehat{sH}].$$

Le vecteur \mathbf{l} a pour composantes parallèle et perpendiculaire au vecteur \mathbf{j} , d'une part $l \cos \widehat{l\mathbf{j}}$ et, d'autre part, $l \sin \widehat{l\mathbf{j}}$. On peut, par des considérations mécaniques, montrer que les vecteurs \mathbf{l} et \mathbf{s} ont un mouvement de précession autour du vecteur \mathbf{j} et que celui-ci à son tour a un mouvement de précession autour de \mathbf{H} , de sorte qu'en moyenne $\overline{l \sin \widehat{l\mathbf{j}}} = 0$. On a alors

$$\overline{l \cos \widehat{lH}} = l \cos \widehat{l\mathbf{j}} \cos \widehat{\mathbf{j}H}$$

et, de même,

$$\overline{s \cos \widehat{sH}} = s \cos \widehat{s\mathbf{j}} \cos \widehat{\mathbf{j}H}.$$

Or, on voit sur la figure 29 que l'on a les expressions

$$\cos \widehat{l\mathbf{j}} = \frac{j^2 + l^2 - s^2}{2jl}, \quad \cos \widehat{s\mathbf{j}} = \frac{j^2 + s^2 - l^2}{2js}, \quad \cos \widehat{\mathbf{j}H} = \frac{m}{j},$$

car la projection sur la direction de \mathbf{H} du moment de rotation total est égale à $m \frac{h}{2\pi}$, m étant le nombre quantique magnétique. On obtient donc

$$\begin{aligned}\Delta W &= \frac{ehH}{4\pi m_0 c} \left[\overline{l \cos \widehat{\mathbf{lH}} + 2s \cos \widehat{\mathbf{sH}}} \right] \\ &= \frac{ehH}{4\pi m_0 c} \left[l \cos \widehat{\mathbf{lj}} + 2s \cos \widehat{\mathbf{s j}} \right] \frac{m}{j} \\ &= m \frac{ehH}{4\pi m_0 c} \left[1 + \frac{j^2 + s^2 - l^2}{2j^2} \right].\end{aligned}$$

Mais nous savons maintenant qu'en Mécanique ondulatoire un carré tel de j^2 est remplacé en règle générale par le produit $j(j+1)$. Nous sommes ainsi amené à penser que la véritable expression de ΔW doit être

$$\Delta W = m \frac{ehH}{4\pi m_0 c} \left[1 + \frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)} \right].$$

Nous retrouverons ainsi la formule de Landé avec

$$g = 1 + \frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)},$$

ce qui est bien la forme générale du facteur de Landé conforme à l'expérience. Dans le cas des alcalins où l'émission des raies spectrales est due au seul électron de valence, on doit poser $s = \frac{1}{2}$ et que l'on ait

$$j = l - \frac{1}{2} \quad \text{ou} \quad j = l + \frac{1}{2},$$

on trouve toujours la valeur (2)

$$g = \frac{2j+1}{2l+1}$$

que nous avons donnée précédemment.

On voit donc que l'on parvenait, à l'aide, il est vrai, de raisonnements imparfaits, à imaginer comment l'hypothèse du spin devait permettre de comprendre l'origine des anomalies spectroscopiques et magnétiques.

3. **Le spin de l'électron en Mécanique ondulatoire.** — Comme nous l'avons vu, Uhlenbeck et Goudsmit avaient introduit leur hypothèse dans le cadre de l'ancienne théorie des Quanta. Mais, à peu près à la même époque, on a compris la nécessité de transformer l'ancienne

théorie des Quanta dans le sens de la Mécanique ondulatoire. On pouvait donc pressentir que l'explication des anomalies spectroscopiques et magnétiques par le spin de l'électron ne pouvait devenir satisfaisante que par l'incorporation du spin à la Mécanique ondulatoire.

L'image de l'électron tournant autour d'un de ses diamètres proposée par Uhlenbeck et Goudsmit était admissible dans le cadre des anciennes théories pour donner une idée intuitive du spin, mais elle cessait de l'être dans la nouvelle Mécanique où l'on ne peut connaître à la fois la position et le mouvement d'un corpuscule et où, à plus forte raison, on ne peut pas se représenter les détails de structure d'une particule élémentaire. De plus, nous le savons, dans la nouvelle Mécanique, on ne peut plus représenter un moment de rotation par un vecteur : on doit le définir par un opérateur et par des énoncés où intervient la notion de probabilité. Nous avons vu qu'on peut ainsi comprendre ce qu'est réellement la « quantification dans l'espace », conception qui est incompréhensible avec la représentation vectorielle classique du moment de rotation. Il doit évidemment en être de même pour le spin ou moment de rotation propre de l'électron.

Nous devons donc définir le spin en termes de probabilité en disant : Étant donné un certain axe Oz , quelles sont les valeurs possibles de la composante du spin suivant Oz et les probabilités respectives de ces valeurs possibles ? A cette question, on doit évidemment répondre d'abord que les valeurs possibles en question sont $\pm \frac{1}{2} \frac{h}{2\pi}$ et M. Pauli a vu le premier que, pour attribuer des probabilités à ces deux valeurs possibles, il était nécessaire de supposer que la fonction Ψ a deux composantes Ψ_1 et Ψ_2 telles que $|\Psi_1|^2 d\tau$ donne la probabilité de trouver l'électron localisé dans l'élément de volume $d\tau$ avec le spin $+\frac{h}{4\pi}$ dans la direction Oz et que $|\Psi_2|^2 d\tau$ donne la probabilité de trouver l'électron localisé dans l'élément de volume $d\tau$ avec le spin $-\frac{h}{4\pi}$ dans la direction Oz .

La question du spin de l'électron se pose ainsi d'une façon assez analogue à celle de la polarisation de la lumière. Pour la lumière, nous devons décomposer l'onde en deux composantes correspondant à deux directions rectangulaires Ox et Oy pour obtenir la probabilité de passage d'un photon à travers un nicol dont la diagonale coïncide avec l'un ou l'autre de ces axes : par exemple $|\Psi_x|^2$ mesure alors la probabilité pour que le photon traverse un nicol qui laisse passer les vibrations parallèles à Ox . Néanmoins, il y a une différence importante entre le

spin et la polarisation : c'est que la polarisation rectiligne est définie par une direction (sans sens), tandis que le spin est défini par une direction et un sens. Aussi pour la lumière, nous devons décomposer l'onde en deux composantes correspondant à deux directions rectangulaires, tandis que pour l'électron nous devons la décomposer en deux composantes correspondant aux deux sens d'une même direction Oz .

Pour la théorie de Pauli, on définit donc les composantes Ψ_1 et Ψ_2 de la fonction d'ordre Ψ en relation avec la composante du spin dans une direction Oz choisie arbitrairement. Si l'on considère ensuite une autre direction Oz' , on devra faire une autre décomposition du Ψ en deux composantes Ψ'_1 et Ψ'_2 et ce sont les carrés des modules de ces composantes qui donneront les probabilités relatives aux deux valeurs possibles $\pm \frac{h}{4\pi}$ que peut donner la détermination de la composante du spin suivant Oz' . Pauli a montré que l'on passe des composantes Ψ_1 et Ψ_2 aux composantes Ψ'_1 et Ψ'_2 par des transformations linéaires qui dépendent de l'orientation de Oz par rapport à Oz' . Ceci est encore tout à fait analogue à la façon dont se transforment les composantes rectangulaires d'une onde lumineuse quand on fait tourner le système des axes rectangulaires. Néanmoins, ici encore, il y a analogie et non identité, car les composantes rectangulaires d'une fonction d'onde lumineuse se transforment comme des composantes de vecteur, alors qu'il résulte des calculs de Pauli que les composantes Ψ_1 et Ψ_2 de la fonction Ψ de l'électron se transforment, mais suivant une loi qui n'est pas celle de la transformation des composantes d'un vecteur.

La tentative de Pauli a eu le grand mérite de montrer clairement comment on devait introduire le spin en Mécanique ondulatoire et a indiqué la nécessité de considérer une fonction d'onde à plus d'une composante. Mais cette tentative était encore incomplète et c'est la théorie de l'électron de Dirac qui a apporté la solution du problème.

4. La théorie de Dirac. — M. Dirac a trouvé l'équation d'ondes fondamentale de sa théorie en cherchant à obtenir une Mécanique ondulatoire *relativiste* qui conserverait pour la densité de probabilité de présence une forme définie positive (c'est-à-dire positive ou nulle, mais jamais négative) analogue à la forme $\Psi\Psi^*$ de la théorie antérieure.

En cherchant à donner un caractère relativiste à la Mécanique ondulatoire, on était parvenu déjà, en 1926, à proposer une théorie dans laquelle la probabilité de présence n'avait plus la forme $\Psi\Psi^*$, mais une forme plus compliquée. Nous avons vu (chap. III, § 6) que la forme rela-

tiviste de la fonction hamiltonienne pour un point matériel libre de masse propre m_0 est

$$H(p_x, p_y, p_z) = c \sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2}.$$

Pour éviter d'avoir un opérateur \mathfrak{H} défini par une expression irrationnelle dont la signification n'est pas claire, on avait proposé (De Donder, Louis de Broglie, Klein, Gordon, Fock, etc.) de poser en l'absence de champ

$$(3) \quad \left(\frac{h}{2\pi i} \right)^2 \frac{\partial^2 \Psi}{\partial t^2} = \mathfrak{H}^2 \Psi = [c^2(\mathfrak{p}_x^2 + \mathfrak{p}_y^2 + \mathfrak{p}_z^2) + m_0^2 c^4] \Psi,$$

où $\mathfrak{p}_x, \mathfrak{p}_y, \mathfrak{p}_z$ sont les opérateurs

$$- \frac{h}{2\pi i} \frac{\partial}{\partial x}, \quad - \frac{h}{2\pi i} \frac{\partial}{\partial y}, \quad - \frac{h}{2\pi i} \frac{\partial}{\partial z}.$$

Cette équation du second ordre conduisait à prendre comme probabilité de présence l'expression (réelle)

$$\varrho = - \frac{h}{4\pi i m_0 c} \left(\Psi \frac{\partial \Psi^*}{\partial t} - \Psi^* \frac{\partial \Psi}{\partial t} \right).$$

Dirac a considéré cette forme de la probabilité de présence comme inadmissible parce qu'elle n'a pas nécessairement toujours une valeur positive et il a admis que la probabilité de présence devait garder la forme définie positive $\Psi\Psi^*$, ou du moins la forme légèrement plus générale $\sum_{\kappa} \Psi_{\kappa} \Psi_{\kappa}^*$ si l'on admet que la fonction d'onde peut avoir plu-

sieurs composantes, comme Pauli l'avait suggéré. Dès que l'on admet cette hypothèse, on peut voir que l'équation en Ψ , ou plus exactement le système des équations en Ψ_{κ} , doit être du premier ordre par rapport au temps sans quoi la condition, évidemment nécessaire, que la probabilité totale $\int_{\mathcal{D}} \sum_{\kappa} \Psi_{\kappa} \Psi_{\kappa}^* d\tau$ de toutes les positions possibles dans l'espace

soit égale à l'unité ne pourrait se trouver automatiquement réalisée à tout instant, si elle l'est à l'instant initial. La symétrie relativiste entre les coordonnées d'espace et de temps indique alors que l'on doit chercher un système d'équations aux dérivées partielles simultanées entre les Ψ_{κ} qui soit du premier ordre par rapport aux quatre variables d'espace-temps.

Dirac a montré qu'on devait, pour y parvenir, prendre au moins quatre fonctions Ψ_{κ} (au lieu de deux seulement comme dans la théorie de Pauli) et il a admis qu'on pouvait se borner à quatre. Il a donc cherché

un système de quatre équations aux dérivées partielles simultanées du premier ordre qui, en l'absence de champ, ait pour conséquence que chacune des fonctions Ψ_k obéisse à l'équation (3). Pour cela, il a introduit quatre matrices à quatre lignes et quatre colonnes que nous nommerons $\alpha_1 \alpha_2 \alpha_3 \alpha_4$ jouissant des propriétés suivantes :

$$\alpha_i^2 = \mathbf{1}, \quad \alpha_i \alpha_j + \alpha_j \alpha_i = \mathbf{0} \quad \text{pour } i \neq j;$$

les matrices doivent aussi être hermitiques, c'est-à-dire que

$$(\alpha_i)_{kl} = (\alpha_i)_{lk}^*.$$

Puis il a défini l'opération linéaire symbolisée par $\alpha_i \Psi_k$ en posant

$$\alpha_i \Psi_k = \sum_l^4 (\alpha_i)_{kl} \Psi_l.$$

Avec les définitions précédentes, Dirac a écrit l'ensemble des quatre équations aux dérivées partielles simultanées en Ψ_k sous la forme condensée

$$(4) \quad \frac{h}{2\pi i} \frac{1}{c} \frac{\partial \Psi_k}{\partial t} + (\mathfrak{p}_x \alpha_1 + \mathfrak{p}_y \alpha_2 + \mathfrak{p}_z \alpha_3 + m_0 c \alpha_4) \Psi_k = 0, \quad \text{avec } k = 1, 2, 3, 4;$$

$\mathfrak{p}_x, \mathfrak{p}_y, \mathfrak{p}_z$ sont toujours les opérateurs

$$-\frac{h}{2\pi i} \frac{\partial}{\partial x}, \quad -\frac{h}{2\pi i} \frac{\partial}{\partial y}, \quad -\frac{h}{2\pi i} \frac{\partial}{\partial z}.$$

Si l'on applique à l'équation (4) l'opération

$$\left(\frac{h}{2\pi i} \frac{1}{c} \frac{\partial}{\partial t} - \mathfrak{p}_x \alpha_1 - \mathfrak{p}_y \alpha_2 - \mathfrak{p}_z \alpha_3 - m_0 c \alpha_4 \right),$$

on trouve aisément, compte tenu des propriétés des α_i , les quatre relations

$$\left[\left(\frac{h}{2\pi i} \right)^2 \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \mathfrak{p}_x^2 - \mathfrak{p}_y^2 - \mathfrak{p}_z^2 - m_0^2 c^2 \right] \Psi_k = 0 \quad (k = 1, 2, 3, 4)$$

qui expriment bien que chaque Ψ_k obéit séparément à l'équation (3). Mais ici les Ψ_k doivent, de plus, obéir aux relations plus restrictives (4) qui établissent des liens entre les valeurs de ces quatre composantes du Ψ .

Les conditions imposées aux α_i ne les déterminent pas d'une façon univoque. Leur choix, qui détermine la forme des équations (4) reste donc en partie arbitraire, ce qui peut paraître étrange, car le change-

ment de l'équation d'ondes entraîne une modification des Ψ_k . Mais on peut démontrer que, pour tout choix de α_i compatible avec les conditions qui leur sont imposées, les valeurs obtenues pour les Ψ_k conduisent aux mêmes prévisions pour les phénomènes observables, car ce ne sont pas les Ψ_k eux-mêmes mais certaines de leurs combinaisons bilinéaires qui représentent des prévisions vérifiables et le changement des α_i qui modifie les Ψ_k laisse précisément invariante ces combinaisons bilinéaires.

Il est souvent commode (mais non obligatoire d'après ce que nous venons de dire) de faire pour les α_i le choix suivant qui satisfait aux conditions requises :

$$\alpha_1 = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix}, \quad \alpha_2 = \begin{vmatrix} 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{vmatrix},$$

$$\alpha_3 = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{vmatrix}, \quad \alpha_4 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}$$

et dans ce qui suivra nous supposons que l'on utilise ces matrices.

S'inspirant de l'ancienne Mécanique relativiste de l'électron, Dirac a ensuite généralisé les équations (4) pour le cas où l'électron se déplace dans un champ électromagnétique défini par un potentiel scalaire V et un potentiel vecteur \mathbf{A} . Il écrit alors à la place de (4)

$$(5) \quad \left(\frac{\hbar}{2\pi i} \frac{1}{c} \frac{\partial}{\partial t} + \frac{e}{c} V \right) \Psi_k + \left[\left(p_x + \frac{e}{c} A_x \right) \alpha_1 + \left(p_y + \frac{e}{c} A_y \right) \alpha_2 + \left(p_z + \frac{e}{c} A_z \right) \alpha_3 + m_0 c \alpha_4 \right] \Psi_k = 0$$

pour $k = 1, 2, 3, 4$.

5. Caractéristiques diverses de la théorie de Dirac. — Dirac a démontré le fait suivant : si l'on passe d'un système de référence galiléen à un autre système de référence galiléen en soumettant les coordonnées d'espace et de temps à une transformation de Lorentz, on peut conserver les équations (4) ou (5) avec les mêmes matrices α_i , mais à condition de remplacer les quatre Ψ_k primitifs par quatre nouvelles composantes Ψ'_k qui sont des combinaisons linéaires des anciennes dont la forme dépend naturellement de la transformation de Lorentz envisagée. Si l'on examine cette transformation linéaire des composantes du Ψ , on s'aperçoit que les Ψ_k ne se comportent pas comme les quatre composantes d'un vecteur d'espace-temps. Les Ψ_k sont des grandeurs d'un

type inconnu en Relativité ancienne, ce sont des composantes de « spineur » ou de « demi-vecteur ». Pour approfondir ce point, on se reportera aux Ouvrages sur la théorie des spineurs.

Quand on écrit explicitement les équations (4) ou (5) avec emploi des matrices α_i choisies plus haut, on constate que ces équations sont très dissymétriques et qu'elles font jouer un rôle particulier à l'axe des z . Ainsi les équations (4) s'écrivent

$$\begin{aligned}\frac{h}{2\pi i} \frac{1}{c} \frac{\partial \Psi_1}{\partial t} &= \frac{h}{2\pi i} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \Psi_4 + \frac{h}{2\pi i} \frac{\partial \Psi_3}{\partial z} - m_0 c \Psi_1, \\ \frac{h}{2\pi i} \frac{1}{c} \frac{\partial \Psi_2}{\partial t} &= \frac{h}{2\pi i} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \Psi_3 - \frac{h}{2\pi i} \frac{\partial \Psi_4}{\partial z} - m_0 c \Psi_2, \\ \frac{h}{2\pi i} \frac{1}{c} \frac{\partial \Psi_3}{\partial t} &= \frac{h}{2\pi i} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \Psi_2 + \frac{h}{2\pi i} \frac{\partial \Psi_1}{\partial z} + m_0 c \Psi_3, \\ \frac{h}{2\pi i} \frac{1}{c} \frac{\partial \Psi_4}{\partial t} &= \frac{h}{2\pi i} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \Psi_1 - \frac{h}{2\pi i} \frac{\partial \Psi_2}{\partial z} + m_0 c \Psi_4.\end{aligned}$$

Pour expliquer cette dissymétrie, il faut se souvenir que le rôle des fonctions d'onde en Mécanique ondulatoire est toujours de permettre l'évaluation de certaines probabilités. Or, comme Pauli l'a remarqué, pour l'électron, à spin, les questions de probabilités relatives au spin doivent être posées par rapport à un axe de référence. On peut, par exemple, se demander quelle est la composante du moment de rotation propre (spin) de l'électron dans une direction de référence D. Les équations de Dirac telles que nous venons de les écrire correspondent au cas où l'on a pris la direction de référence D comme axe des z , ce qui explique le rôle particulier de cet axe dans les équations. Comme le montre une étude plus approfondie de la théorie de Dirac, la probabilité pour que la composante z du spin soit $+\frac{h}{4\pi}$ est donnée par $|\Psi_1|^2 + |\Psi_3|^2$ et la probabilité pour que la composante z du spin soit $-\frac{h}{4\pi}$ par $|\Psi_2|^2 + |\Psi_4|^2$. Si donc on veut répondre à des questions de probabilité relatives au spin pour lesquelles la direction D de référence ne coïncide pas avec l'axe Oz initialement choisi, on devra d'abord faire un changement de coordonnées qui fasse coïncider le nouvel axe des z avec la direction D et ce sont les nouvelles composantes Ψ'_k , combinaisons linéaires des anciennes, qui fourniront, par les expressions que nous venons de donner, les probabilités cherchées.

Ce qu'il y a de plus remarquable dans la théorie de Dirac, c'est qu'étant partie d'un effort pour obtenir une Mécanique ondulatoire relativiste, sans faire aucune hypothèse pour introduire un moment magné-

tique propre et un moment de rotation propre de l'électron, elle s'est trouvée avoir introduit automatiquement ces deux éléments.

Dirac a, en effet, montré d'abord que, si l'on considère le mouvement d'un électron dans un champ électromagnétique conformément à l'équation (5), tout se passe comme si l'électron possédait un moment magnétique propre et un moment électrique propre, chacune de leurs composantes étant définie par un opérateur, comme cela doit être dans une théorie quantique. Le moment magnétique a comme valeur possible (valeur propre) $\frac{eh}{4\pi m_0 c}$, c'est-à-dire un magnéton de Bohr en conformité avec l'hypothèse d'Uhlenbeck et Goudsmit. Le moment électrique apparaît, conformément aux lois de la Relativité, par suite du mouvement de l'électron, doué de magnétisme.

En ce qui concerne le moment de rotation, on constate en théorie de Dirac que le moment orbital de l'électron ne satisfait pas au même théorème de conservation que dans les Mécaniques antérieures; il ne reste pas constant dans un champ de force central. Pour obtenir une grandeur qui jouisse de cette propriété classique de conservation, il faut ajouter au moment de rotation orbital de l'électron un « moment de rotation propre » dont les composantes sont naturellement encore définies par des opérateurs. Ces opérateurs sont

$$\mathfrak{M}_x^{(0)} = \frac{h}{4\pi} i \alpha_2 \alpha_3, \quad \mathfrak{M}_y^{(0)} = \frac{h}{4\pi} i \alpha_3 \alpha_1, \quad \mathfrak{M}_z^{(0)} = \frac{h}{4\pi} i \alpha_1 \alpha_2$$

qui ont, tous trois, pour valeurs propres $\pm \frac{h}{4\pi}$, mais ne sont pas simultanément mesurables. On retrouve bien l'hypothèse d'Uhlenbeck et Goudsmit, mais complétée par l'introduction des idées nouvelles sur la nature des moments de rotation.

Donc, dans la théorie de Dirac, on retrouve bien toutes les caractéristiques essentielles de l'électron magnétique et tournant d'Uhlenbeck et Goudsmit, mais indépendantes de l'image trop classique d'une sphère d'électricité en rotation, et incorporées dans la nouvelle Mécanique. Nous allons maintenant résumer les autres succès de cette théorie et dire aussi un mot de ses insuffisances.

6. Succès et insuffisances de la théorie de Dirac. — La théorie de Dirac est parvenue à interpréter les anomalies spectroscopiques et magnétiques, du moins dans le cas où il est possible de ne considérer qu'un électron dans l'atome (atome d'hydrogène ou atome hydrogénoïde tel que ceux des alcalins).

Ainsi pour les spectres hydrogénoïdes, et également pour les spectres X, si l'on admet les approximations déjà introduites dans les théories antérieures pour la description de ces spectres, la théorie de Dirac permet de déterminer les niveaux d'énergie, c'est-à-dire les valeurs de E pour lesquelles il existe un ensemble de quatre fonctions

$$\Psi_k = a_k(q_i) e^{\frac{2\pi i}{h} E t}$$

finies, uniformes, continues, nulles aux limites du domaine D et satisfaisant aux équations de Dirac. On trouve que chaque niveau d'énergie est caractérisé par trois nombres quantiques n, l, j tels que $0 \leq l < n$ et $j = l \pm \frac{1}{2}$. Ainsi se trouvent dédoublés tous les niveaux d'énergie prévus par la théorie de Sommerfeld, sauf les niveaux pour lesquels $l = 0$ et nous savons que ceci est en accord avec l'expérience. De plus, la théorie de Dirac trouve comme énergie E du niveau n, l, j dans un atome de nombre atomique Z

$$E(n, l, j) = - \frac{R h (Z - z(n, l))^2}{n^2} \left[1 + \frac{\alpha^2 (Z - z(n, l))^2}{n^2} \left(\frac{n}{j + \frac{1}{2}} - \frac{3}{4} \right) \right],$$

où R est la constante de Rydberg; α la constante de structure fine $\frac{2\pi e^2}{hc}$ et $z(n, l)$ le nombre d'écran. Cette formule est semblable à celle donnée par la théorie de Sommerfeld [form. (5) du chap. VII], mais avec cette différence essentielle que le nombre quantique azimutal k de Sommerfeld est remplacé par $j + \frac{1}{2}$ dans la dernière parenthèse. En raison de cette différence, la formule de Dirac prévoit l'existence des « doublets de relativité » entre les niveaux ayant mêmes nombres quantiques n et l , mais des nombres j différents d'une unité : c'est bien là aussi ce que l'expérience exige. La théorie de Dirac remet ainsi entièrement dans la bonne forme les résultats de la théorie de la structure fine de Sommerfeld montrant ainsi que le succès de la théorie de Sommerfeld n'était pas fortuit et qu'en définitive les doublets réguliers sont bien dus à l'influence de la Relativité; mais pour avoir une théorie tout à fait correcte, il fallait introduire non seulement la Relativité, mais le spin.

En ce qui concerne les anomalies magnétiques, la théorie de Dirac parvient à retrouver pour les alcalins (pour lesquels on peut approxi-

mativement considérer isolément l'électron de valence) la formule de Landé

$$g = \frac{j + \frac{1}{2}}{l + \frac{1}{2}}.$$

Une des insuffisances de la théorie de Dirac est qu'elle ne peut être appliquée qu'à un seul électron dans un champ donné et non à un ensemble d'électrons en interaction. Le problème de constituer une Mécanique ondulatoire relativiste des ensembles d'électrons est un problème difficile qui ne peut être abordé qu'en introduisant la théorie quantique du rayonnement. Néanmoins, on a pu introduire approximativement l'existence du spin pour interpréter les spectres lumineux des atomes plus compliqués que ceux des alcalins et pour rendre compte des effets Zeeman anomaux des corps autres que les alcalins. Nous n'entrerons pas dans le détail de ces théories, nous bornant à dire que l'introduction du spin a partout permis d'interpréter correctement les faits.

Il est donc aujourd'hui certain que l'électron n'est pas caractérisé uniquement par sa charge et par sa masse, mais aussi par un moment magnétique et un moment de rotation propres. Le spin est une propriété intrinsèque de l'électron et sans doute aussi de tous les corpuscules élémentaires de matière. En partant de la théorie de Dirac considérée comme la théorie des corpuscules élémentaires de spin $\frac{1}{2} \frac{h}{2\pi}$, on a pu, par généralisation, construire une théorie générale des particules douées de spin. Dans le cadre de cette théorie générale, le photon apparaît comme une particule de spin $\frac{h}{2\pi}$ dont l'onde associée est l'onde électromagnétique de Maxwell. Nous renvoyons pour ces questions aux Ouvrages spéciaux [4].

Les équations de Dirac admettent des solutions à énergie négative qui devraient correspondre à des électrons dans des états à propriétés paradoxales qui n'ont jamais été observés. Il semblait y avoir là une difficulté pour la théorie de Dirac. Mais Dirac a pu interpréter les états à énergie négative comme indiquant l'existence possible d'antiélectrons ou « électrons positifs » ayant même masse que l'électron avec une charge égale, mais de signe contraire. De tels électrons positifs ont été ensuite effectivement décelés dans le rayonnement cosmique et parmi les produits de la désintégration des corps radioactifs artificiels, de sorte que ce qui semblait un inconvénient de la théorie de Dirac a finalement

turné à son avantage. Nous ne pouvons insister ici sur la question, d'ailleurs très importante, des électrons positifs.

Enfin signalons une dernière insuffisance de la théorie de Dirac. Récemment, en examinant avec beaucoup de soin par des procédés radioélectriques la structure fine de la série de Balmer, Lamb et Retherford ont montré que les formules de Dirac, malgré le progrès qu'elles réalisent sur celles de Sommerfeld ne suffisent pas encore à prévoir quantitativement d'une manière tout à fait exacte cette structure fine. Cette divergence a été ensuite interprétée théoriquement par des considérations dont le principe est dû à H. Bethe. Ces considérations assez compliquées ne peuvent être exposées ici.

BIBLIOGRAPHIE.

- [1] P. A. M. DIRAC, *The quantum theory of the Electron* (*Proc. Roy. Soc.*, A. 117, 1928, p. 610).
- [2] P. A. M. DIRAC, *Les principes de la Mécanique quantique*, Presses universitaires, Paris, 1931.
- [3] LOUIS DE BROGLIE, *L'électron magnétique (théorie de Dirac)*, Hermann, Paris, 1934.
- [4] LOUIS DE BROGLIE, *Théorie générale des particules à spin*, Gauthier-Villars, Paris, 1943.

CHAPITRE XV.

LE PRINCIPE DE PAULI ET LA MÉCANIQUE ONDULATOIRE DES SYSTÈMES DE CORPUSCULES.

1. La répartition des électrons dans l'atome. — La meilleure manière de comprendre l'origine expérimentale du principe d'exclusion de Pauli est d'étudier la répartition des électrons dans l'atome.

Dans un atome, les électrons se répartissent en un certain nombre de groupes que l'on a nommés par ordre d'énergie de liaison décroissante : électrons K, électrons L, électrons M, etc. D'après la théorie de Bohr, l'arrachement d'un de ces électrons crée une « place libre » sur le niveau d'énergie où il était placé et met l'atome en état d'émettre une des raies de la série Röntgen correspondante. Mais, nous l'avons vu, ces séries, à l'exception de la série K, sont complexes et se subdivisent en sous-séries. On est donc amené à dire que les groupes d'électrons L, M, ... se subdivisent en sous-groupes qui correspondent à des niveaux d'énergie voisins, mais un peu différents. On dit souvent que les électrons d'un même groupe forment une « couche » : couche K, couche L, etc. La couche K est simple, mais la couche L se subdivise en trois niveaux : L_I , L_{II} , L_{III} ; la couche M en cinq niveaux M_I , ..., M_V , etc.

Pour comprendre la structure électronique de l'atome, il a été nécessaire d'admettre une « saturation » des niveaux d'énergie, c'est-à-dire qu'il y a un nombre maximum d'électrons pouvant se trouver sur chaque niveau. Si cela n'était pas, tous les électrons devraient, en vertu de la loi de répartition de Boltzmann-Gibbs, se trouver normalement sur le niveau K de moindre énergie. Toute l'étude systématique de la spectroscopie X et des propriétés physico-chimiques des éléments montre qu'il n'en est pas ainsi. Quand on passe d'un élément au suivant dans le système périodique des éléments, un nouvel

électron vient s'ajouter sur le niveau de faible énergie *sur lequel il y a encore une place libre*. Quand on suit ainsi pas à pas l'édification du système électronique à travers les quatre-vingt-douze éléments de la série de Mendéléeff, on doit parvenir à interpréter la variation, dans l'ensemble de ces éléments des propriétés chimiques, optiques, Röntgen et même magnétiques. C'est ce que Bohr a très bien vu. Mais, comme l'étude théorique *a priori* de la structure des éléments était pratiquement impossible, Bohr et ses continuateurs ont suivi la voie inverse : partant des données expérimentales sur les propriétés chimiques, optiques, Röntgen et magnétiques des éléments, ils ont cherché comment les couches et les niveaux devaient se remplir successivement, au fur et à mesure que le système atomique se complique par l'adjonction de nouveaux électrons, de façon à rendre compte de la variation de ces propriétés [1].

Les efforts faits en ce sens, notamment par Bohr, Stoner et Main Smith, on abouti à préciser la répartition des électrons entre couches et niveaux et à fixer le maximum des électrons que chaque niveau peut comporter. Comme nous l'avons vu, on peut considérer chaque niveau interne comme caractérisé par trois nombres quantiques n, l, j . Les niveaux ayant même nombre n forment une « couche ». D'autre part, pour n donné, l peut prendre les valeurs $0, 1, 2, \dots, n-1$ et j peut prendre pour une valeur donnée de l les valeurs $l \pm \frac{1}{2}$ (saut pour $l=0$ où la valeur négative $l - \frac{1}{2}$ est exclue). Dans une couche définie par le nombre n , il y aura donc $1 + 2(n-1) = 2n-1$ niveaux différents. Dans la couche K, ($n=1$), il y a un seul niveau $n=1$, $l=0, j=\frac{1}{2}$. Dans la couche L ($n=2$), il y a trois niveaux

$$\begin{aligned} L_I : n = 2, \quad l = 0, \quad j = \frac{1}{2}; \quad L_{II} : n = 2, \quad l = 1, \quad j = \frac{1}{2}; \\ L_{III} : n = 2, \quad l = 1, \quad j = \frac{3}{2}; \quad \dots \end{aligned}$$

Les recherches de Bohr, Stoner et Main Smith ont permis d'obtenir avec certitude la règle suivante :

RÈGLE DE STONER. — *Le nombre maximum des électrons qui peuvent trouver place sur le niveau (n, l, j) est égal à $2j + 1$.*

La règle précédente permet de calculer le nombre maximum d'électrons par couche, nombre qui était bien connu empiriquement avant l'établissement de la règle de Stoner. L'application de cette règle

donne, en effet, pour le nombre maximum d'électrons pouvant trouver place sur la couche caractérisée par le nombre n

$$2 + \sum_1^{n-1} \left[2 \binom{l+1}{2} + 1 + 2 \binom{l-1}{2} + 1 \right] = 2 + \sum_1^{n-1} (4l+2)$$

$$= 2 + 2(n-1) + 4 \frac{n(n-1)}{2} = 2n^2,$$

ce qui concorde bien avec l'expérience.

Appliquée aux différents niveaux, la règle de Stoner fournit le tableau suivant pour la répartition des électrons sur les quatre premières couches d'un atome lourd :

	$n = 1.$	$n = 2.$				$n = 3.$				
	$l = 0,$	$l = 0,$	$l = 1,$	$l = 1,$	$l = 0,$	$l = 1,$	$l = 1,$	$l = 2,$	$l = 2,$	$l = 2,$
	$j = \frac{1}{2}.$	$j = \frac{1}{2}.$	$j = \frac{1}{2}.$	$j = \frac{3}{2}.$	$j = \frac{1}{2}.$	$j = \frac{1}{2}.$	$j = \frac{3}{2}.$	$j = \frac{3}{2}.$	$j = \frac{5}{2}.$	$j = \frac{5}{2}.$
Niveaux.....	K	L _I	L _{II}	L _{III}	M _I	M _{II}	M _{III}	M _{IV}	M _V	
Nombre d'électrons par niveau = $2j + 1$	2	2	2	4	2	2	4	4	6	
Nombre d'électrons par couche = $2n^2$	2	8			18					
		$n = 4.$								
	$l = 0,$	$l = 1,$	$l = 1,$	$l = 2,$	$l = 2,$	$l = 3,$	$l = 3,$	$l = 3,$	$l = 3,$	$l = 3.$
	$j = \frac{1}{2}.$	$j = \frac{1}{2}.$	$j = \frac{3}{2}.$	$j = \frac{3}{2}.$	$j = \frac{5}{2}.$	$j = \frac{5}{2}.$	$j = \frac{5}{2}.$	$j = \frac{7}{2}.$	$j = \frac{7}{2}.$	$j = \frac{7}{2}.$
Niveaux.....	N _I	N _{II}	N _{III}	N _{IV}	N _V	N _{VI}	N _{VII}	N _{VIII}	N _{IX}	
Nombre d'électrons par niveau = $2j + 1$	2	2	4	4	6	6	6	8	8	
Nombre d'électrons par couche = $2n^2$	32									

Nous n'analyserons pas ici comment en étudiant le développement progressif des systèmes de niveaux au cours de la série de Mendéléeff, on arrive à faire correspondre les étapes de ce développement avec les propriétés chimiques, optiques et magnétiques des éléments. C'est là un sujet d'un haut intérêt, mais ici notre but essentiel est d'interpréter la règle de Stoner.

2. Le principe d'exclusion de Pauli. — Considérons un atome et un de ses niveaux d'énergie électronique $W(n, l, j)$. Nous avons vu que, si nous soumettons l'atome à l'action d'un champ magnétique \mathbf{H} , le niveau en question se décompose en $2j + 1$ niveaux d'énergie

$$W_{II}(n, l, j) = W(n, l, j) + m g \frac{eH}{4\pi m_0 c},$$

où m peut prendre $2j + 1$ valeurs $-j, -(j-1), \dots, (j-1), j$, le nombre quantique j étant toujours un demi-entier de la série $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$. Si le champ \mathbf{H} tend progressivement vers zéro, les $2j + 1$ valeurs de W_{\parallel} tendent vers W , de sorte que pour un champ nul, il y a dégénérescence, les $2j + 1$ niveaux distincts en présence du champ étant venus se confondre. On peut donc dire que, pour caractériser un niveau, il faut en réalité quatre nombres quantiques n, l, j, m , mais qu'en l'absence de champs extérieurs, les niveaux qui ne diffèrent que par les valeurs de m sont confondus. Pour retrouver la règle de Stoner, il suffit donc d'admettre le principe suivant dû à M. Pauli :

PRINCIPE DE PAULI. — *Sur chaque niveau simple caractérisé par les quatre nombres quantiques n, l, j, m , il ne peut y avoir plus d'un électron.*

En effet, si l'on suppose l'atome plongé dans un champ magnétique, les $2j + 1$ niveaux correspondant aux mêmes valeurs de n, l, j et aux $(2j + 1)$ valeurs possibles de m sont distincts et, d'après le principe de Pauli, portent chacun un électron s'ils sont tous saturés. Si ensuite le champ tend vers zéro, les $2j + 1$ niveaux en question ont, à la limite, la même énergie et l'on peut admettre qu'en l'absence de champ il n'y a plus qu'un seul niveau d'énergie $W(n, l, j)$ portant $2j + 1$ électrons s'il est saturé. On retombe donc sur la règle de Stoner.

On peut encore énoncer le principe de Pauli en disant que la présence d'un électron sur un niveau non dégénéré *exclut* la présence de tout autre électron sur ce niveau. D'où le nom de principe d'exclusion qui lui est souvent donné.

Il est intéressant de voir comment le principe de Pauli s'applique dans un cas tout à fait différent de celui des électrons dans l'atome : le cas d'un gaz d'électrons contenu dans une enceinte à parois parfaitement réfléchissantes pour ces corpuscules. Si l'on néglige les interactions, le mouvement des électrons est rectiligne et uniforme et, en Mécanique ondulatoire, on associera à ces électrons des ondes planes monochromatiques. Mais ces ondes seront quantifiées, car elles doivent être nulles sur les parois et, par suite, doivent avoir la forme d'ondes stationnaires définies par les dimensions de l'enceinte. Par exemple, si deux parois de l'enceinte sont planes, parallèles et distantes de D , il pourra exister des ondes stationnaires dont la demi-longueur d'onde sera un sous-multiple de D ($D = n \frac{\lambda}{2}$).

Il existera donc un ensemble d'ondes Ψ stationnaires correspondant aux mouvements quantifiés possibles des électrons dans l'enceinte. Nous verrons dans le chapitre XVI, en étudiant les statistiques quantiques, qu'au point de vue de la Mécanique statistique du gaz d'électrons, l'existence de ces états stationnaires est équivalente à l'existence de cellules d'extension-en-phase de grandeur h^3 . Le principe de Pauli nous apprend alors qu'à chaque onde stationnaire (ou ce qui revient au même à chaque cellule d'extension-en-phase) peuvent être associés au maximum *deux* électrons. Pourquoi deux ? Parce que, pour caractériser entièrement l'état stationnaire d'un électron dans l'enceinte, il ne suffit pas de se donner l'onde stationnaire à laquelle il est associé : il faut encore se donner la direction de son spin. Pour cela, on choisira une direction de référence Δ et l'on se donnera la composante suivant Δ du spin de l'électron considéré. Or, cette composante a deux valeurs possibles $\pm \frac{h}{4\pi}$, et, par suite, à une onde Ψ stationnaire donnée correspondent deux états non dégénérés de l'électron : une onde stationnaire peut donc porter deux électrons dont les valeurs du spin sont opposées. Nous verrons que la statistique quantique obtenue en prenant le principe de Pauli ainsi précisé pour base, *la statistique de Fermi-Dirac*, est bien celle qui paraît s'appliquer aux assemblées d'électrons et donner la clef de certaines difficultés de la théorie des métaux.

Nous avons maintenant à chercher la traduction mathématique du principe de Pauli, mais, pour pouvoir le faire, il nous faut d'abord connaître les grandes lignes de la Mécanique ondulatoire (non relativiste) des systèmes de corpuscules dont nous n'avons pas encore parlé.

3. La Mécanique ondulatoire des systèmes de corpuscules. — Jusqu'à présent, en Mécanique ondulatoire, nous avons considéré uniquement le mouvement du corpuscule dans un champ de force donné. Un problème plus général est celui du mouvement d'un système de corpuscules qui interagissent les uns avec les autres. Schrödinger a indiqué la façon dont on devait appliquer la Mécanique ondulatoire à ce cas, en négligeant toutefois les corrections de Relativité.

Pour obtenir cette Mécanique ondulatoire non relativiste des systèmes, il faut partir de l'idée que l'on doit envisager l'espace de configuration à $3N$ dimensions formé à l'aide des $3N$ coordonnées rectangulaires des N corpuscules constituant le système et associer au mouvement du système la propagation d'une onde dans cet espace de configuration.

Pour que le raccord avec la Mécanique classique puisse se faire, il faut que l'Optique géométrique de cette propagation d'ondes nous ramène à la théorie de Jacobi.

Pour obtenir la propagation de l'onde Ψ dans l'espace de configuration, le moyen le plus simple est d'employer le procédé automatique qui a réussi dans le cas d'un seul corpuscule : on remplacera donc dans l'expression hamiltonienne classique de l'énergie du système les coordonnées x_k, y_k, z_k du $k^{\text{ième}}$ constituant par les opérateurs « multiplication par x_k », « multiplication par y_k », « multiplication par z_k » et les moments de Lagrange $p_{x_k}, p_{y_k}, p_{z_k}$ par les opérateurs

$$-\frac{h}{2\pi i} \frac{\partial}{\partial x_k}, -\frac{h}{2\pi i} \frac{\partial}{\partial y_k}, -\frac{h}{2\pi i} \frac{\partial}{\partial z_k}.$$

On obtient ainsi un opérateur hamiltonien

$$H\left(x, y, z, \dots, x_N, y_N, z_N, -\frac{h}{2\pi i} \frac{\partial}{\partial x_1}, \dots, -\frac{h}{2\pi i} \frac{\partial}{\partial z_N}, t\right).$$

En appliquant cet opérateur à la fonction d'ondes $\Psi(x_1, \dots, z_N, t)$ et en égalant le résultat à $\frac{h}{2\pi i} \frac{\partial \Psi}{\partial t}$, on obtient l'équation de propagation

$$\frac{h}{2\pi i} \frac{\partial \Psi}{\partial t} = H\left(x_1, \dots, z_N, -\frac{h}{2\pi i} \frac{\partial}{\partial x_1}, \dots, -\frac{h}{2\pi i} \frac{\partial}{\partial z_N}, t\right) \Psi.$$

Les coordonnées x_1, \dots, z_N étant des coordonnées rectangulaires, l'expression hamiltonienne classique de l'énergie s'écrit

$$H(x_1, \dots, z_N, p_{x_1}, \dots, p_{z_N}, t) = \sum_1^N \frac{1}{2m_k} (p_{x_k}^2 + p_{y_k}^2 + p_{z_k}^2) + V(x_1, \dots, z_N, t).$$

L'énergie potentielle $V(x_1, \dots, z_N, t)$ est formée de deux sortes de termes :

1° des termes d'actions mutuelles entre les corpuscules de la forme

$$V_{ij} \left(\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \right);$$

2° des termes exprimant l'action du champ extérieur éventuellement appliqué au système, termes qui sont de la forme $V_i(x_i, y_i, z_i, t)$.

En procédant comme il a été indiqué ci-dessus, on trouve donc comme équation de propagation dans l'espace de configuration

$$\sum_1^N \frac{1}{2m_k} \left(\frac{\partial^2 \Psi}{\partial x_k^2} + \frac{\partial^2 \Psi}{\partial y_k^2} + \frac{\partial^2 \Psi}{\partial z_k^2} \right) - \frac{8\pi^2}{h^2} V(x_1, \dots, z_N, t) \Psi = \frac{4\pi i}{h} \frac{\partial \Psi}{\partial t}.$$

S'il n'y a qu'un seul corpuscule ($N = 1$), l'espace de configuration se réduit à l'espace physique ordinaire et l'on retombe sur l'équation d'ondes antérieurement étudiée.

Étudions le cas particulier où le champ extérieur, s'il existe, est constant. Alors V est indépendant du temps et le système est conservatif. On peut donc considérer des solutions monochromatiques dépendant du temps par le facteur $e^{\frac{2\pi i}{h}Et}$ et obéissant à l'équation de Schrödinger

$$\sum_1^N \frac{1}{m_k} \left(\frac{\partial^2 \Psi}{\partial x_k^2} + \frac{\partial^2 \Psi}{\partial y_k^2} + \frac{\partial^2 \Psi}{\partial z_k^2} \right) + \frac{8\pi^2}{h^2} [E - V(x_1, \dots, z_N)] \Psi = 0.$$

Dans une région de l'espace de configuration où V varie lentement à l'échelle de la longueur d'onde, l'Optique géométrique sera applicable à la propagation de l'onde Ψ et l'on pourra écrire celle-ci sous la forme

$$\Psi = \alpha e^{\frac{2\pi i}{h}[Et - S_1(x_1, \dots, z_N)]},$$

où α est une fonction lentement variable dont les dérivées sont négligeables par rapport à celles de S_1 . En substituant cette forme du Ψ dans l'équation d'ondes on trouve

$$\sum_1^N \frac{1}{2m_k} \left[\left(\frac{\partial S_1}{\partial x_k} \right)^2 + \left(\frac{\partial S_1}{\partial y_k} \right)^2 + \left(\frac{\partial S_1}{\partial z_k} \right)^2 \right] + V(x_1, \dots, z_N) = E,$$

ce qui est l'équation de Jacobi de la Mécanique classique des systèmes. On voit donc bien qu'ici, comme dans le cas d'un seul corpuscule, la Dynamique classique correspond à l'approximation de l'Optique géométrique.

Un cas intéressant à envisager est celui où les corpuscules sont sans interaction les uns avec les autres : on doit alors pouvoir les considérer isolément. Or la fonction V se réduit alors à $\sum_1^N V_k(x_k, y_k, z_k, t)$ et l'équation d'ondes dans l'espace de configuration s'écrit

$$\frac{4\pi i}{h} \frac{\partial \Psi}{\partial t} = \sum_1^N \left[\frac{1}{m_k} \left(\frac{\partial^2 \Psi}{\partial x_k^2} + \frac{\partial^2 \Psi}{\partial y_k^2} + \frac{\partial^2 \Psi}{\partial z_k^2} \right) - \frac{8\pi^2}{h^2} V_k(x_k, y_k, z_k, t) \right] \Psi.$$

Si nous posons

$$\Psi(x_1, \dots, z_N, t) = \Psi_1(x_1, y_1, z_1, t) \Psi_2(x_2, y_2, z_2, t) \dots \Psi_N(x_N, y_N, z_N, t),$$

l'équation précédente se décompose en N équations indépendantes de la forme

$$\frac{1}{m_k} \left(\frac{\partial^2 \Psi}{\partial x_k^2} + \frac{\partial^2 \Psi}{\partial y_k^2} + \frac{\partial^2 \Psi}{\partial z_k^2} \right) - \frac{8\pi^2}{h^2} V_k(x_k, y_k, z_k, t) \Psi_k = \frac{4\pi i}{h} \frac{\partial \Psi_k}{\partial t}$$

et ce résultat exprime précisément que l'on peut considérer chaque corpuscule isolément. La fonction d'onde du système dans l'espace de configuration est donc alors le *produit* des fonctions d'onde de chaque corpuscule considéré isolément dans l'espace à trois dimensions.

Il est aisé de transposer le principe de localisation dans la Mécanique ondulatoire des systèmes. On doit ici l'énoncer comme il suit :

Si l'état d'un système est représenté par la fonction d'onde $\Psi(x_1, \dots, z_N, t)$ dans l'espace de configuration, la probabilité pour qu'une mesure faite au temps t permette de localiser le point figuratif du système dans l'élément de volume $d\tau = dx_1 dy_1 \dots dz_N$ de l'espace de configuration est mesurée par

$$\Psi(x_1, \dots, z_N, t) \Psi^*(x_1, \dots, z_N, t) d\tau.$$

S'il y a un seul corpuscule ($N = 1$), nous retombons sur le principe de localisation sous la forme étudiée antérieurement. Si les N corpuscules du système sont sans interactions les uns sur les autres, la fonction Ψ du système sera le produit des fonctions individuelles Ψ_k des constituants et nous aurons

$$\begin{aligned} & |\Psi(x_1, \dots, z_N, t)|^2 d\tau \\ &= |\Psi_1(x_1, y_1, z_1, t)|^2 dx_1 dy_1 dz_1 \dots |\Psi_N(x_N, y_N, z_N, t)|^2 dx_N dy_N dz_N. \end{aligned}$$

Ce résultat est en accord avec le théorème des probabilités composées, car, les constituants ne réagissant pas les uns sur les autres, leurs présences dans les différents éléments de volume de l'espace physique sont des événements indépendants. Nous apercevons alors pourquoi la fonction Ψ du système devait, dans ce cas, être le produit des fonctions d'onde individuelles des constituants.

On démontre que l'intégrale $\int_{\mathfrak{R}^N} |\Psi|^2 d\tau$ est constante au cours du temps, ce qui permet de normer d'une façon permanente la fonction Ψ en la posant égale à 1.

Le principe de décomposition spectrale s'énonce pour un système comme pour un corpuscule unique. En particulier, si le système est conservatif, l'onde Ψ peut toujours être décomposée en une série d'ondes planes monochromatiques et l'intensité de chaque composante

spectrale donne la probabilité pour qu'une mesure permette d'attribuer au système l'énergie correspondante.

En étudiant la représentation d'un train d'ondes Ψ dans l'espace de configuration par une intégrale de Fourier, on retrouve les relations d'incertitude sous la forme

$$\Delta q_i \Delta p_i \dots h.$$

Bref, tous les résultats essentiels obtenus par la Mécanique ondulatoire du corpuscule unique se transposent aisément dans la Mécanique ondulatoire des systèmes de corpuscules, mais la propagation de l'onde Ψ d'un système dans l'espace de configuration a un caractère très abstrait et montre clairement combien les ondes de la Mécanique ondulatoire sont loin d'avoir la signification concrète des ondes de l'ancienne Physique.

4. Cas des systèmes formés de corpuscules de même nature. — Le cas des systèmes formés de corpuscules de même nature est très important. Il comprend, en particulier, tous les atomes qui, quand on considère les noyaux comme de simples centres de forces, sont des systèmes de ce genre puisqu'ils sont formés d'électrons. Ce qui caractérise ces systèmes, c'est que si l'on y considère un certain état quantique, l'état n'est en réalité aucunement modifié si l'on y permute le rôle de deux des particules constituantes. On doit donc admettre que toutes les grandeurs observables telles que le $|\Psi|^2$ doivent être insensibles à une permutation quelconque du rôle de deux constituants. Ce postulat conduit à restreindre la forme possible des fonctions d'onde. Comme les interactions entre deux corpuscules sont toujours des fonctions symétriques de leurs coordonnées, si l'on a trouvé une solution de l'équation d'ondes $\Psi(x_1, y_1, z_1, \dots, x_i, y_i, z_i, \dots, x_k, y_k, z_k, \dots, t)$, la fonction $\Psi(x_1, y_1, z_1, \dots, x_k, y_k, z_k, \dots, x_i, y_i, z_i, \dots, t)$ obtenue en permutant le rôle des particules i et k sera aussi une solution ainsi que toute combinaison linéaire de la forme

$$a\Psi(x_1, y_1, z_1, \dots, x_i, y_i, z_i, \dots, x_k, y_k, z_k, \dots, t) \\ + b\Psi(x_1, y_1, z_1, \dots, x_k, y_k, z_k, \dots, x_i, y_i, z_i, \dots, t).$$

Si nous écrivons alors que $|\Psi|^2$ doit être insensible à une permutation des particules i et k , nous sommes conduit à poser $b = \pm a$ et nous pourrions ensuite vérifier que, si cette condition est réalisée, toutes les autres grandeurs observables sont aussi insensibles à la permutation de i et de k .

Pour énoncer ce résultat sous une forme plus générale, introduisons les définitions suivantes : une fonction d'onde sera dite « symétrique » si, quand on y intervertit le rôle de deux particules de même nature, elle ne change pas de valeur; une fonction d'onde sera dite « antisymétrique » si, quand on y intervertit le rôle de deux particules de même nature, elle garde la même valeur absolue, mais en changeant de signe. Ceci posé, le résultat général auquel on arrive en exprimant que toutes les grandeurs observables doivent être insensibles à la permutation de deux particules de même nature peut s'énoncer ainsi : *la fonction d'onde Ψ d'un système contenant des particules de même nature doit être, soit symétrique, soit antisymétrique pour tout couple de ces particules.*

On peut démontrer qu'il existe toujours une infinité d'états symétriques et d'états antisymétriques possibles pour un système de particules identiques. En effet, soit $\Psi(1, 2, \dots, N, t)$ une solution de l'équation d'ondes du système contenant N particules (i représente l'ensemble des trois coordonnées x_i, y_i, z_i). Soit \mathfrak{P} une permutation quelconque des N nombres $1, 2, \dots, N$. La fonction $\mathfrak{P}\Psi(1, 2, \dots, N, t)$ est la fonction obtenue en effectuant la permutation \mathfrak{P} sur les arguments de la fonction $\Psi(1, 2, \dots, N, t)$. Il est alors évident que la fonction $\sum_{\mathfrak{P}} \mathfrak{P}\Psi(1, 2, \dots, N, t)$ où la somme $\sum_{\mathfrak{P}}$ porte sur toutes les permutations \mathfrak{P} possibles, y compris la permutation identique, est une fonction symétrique et l'existence d'une infinité de fonctions symétriques se trouve ainsi démontrée. D'autre part, soit la fonction $\sum_{\mathfrak{P}} \pm \mathfrak{P}\Psi(1, 2, \dots, N, t)$ où l'on prend le signe $+$ devant les permutations *paires* (c'est-à-dire celles qui dérivent de la suite $1, \dots, N$ par un nombre *pair* d'échanges simples) et le signe $-$ devant les permutations *impaires* (c'est-à-dire celles qui dérivent de la suite $1, 2, \dots, N$ par un nombre *impair* d'échanges simples) : la fonction obtenue est visiblement antisymétrique, ce qui démontre l'existence d'une infinité d'états antisymétriques.

Donc, parmi les états possibles *a priori* pour le système, se trouvent une infinité d'états symétriques et une infinité d'états antisymétriques. On doit admettre que, pour les systèmes à particules de nature identique, seuls ces états sont réalisés dans la nature. Ce qui achève de justifier cette hypothèse, c'est le théorème suivant que l'on démontre en s'appuyant sur la symétrie des fonctions qui expriment les interactions :

Un système qui est initialement dans un état symétrique ne peut subir de transition que vers un autre état symétrique et un système qui est initialement dans un état antisymétrique ne peut subir de transition que vers un autre état antisymétrique.

Ainsi, les états symétriques, d'une part, les états antisymétriques, d'autre part, forment des ensembles fermés et ne peuvent se combiner qu'entre eux.

On peut donc très bien supposer que, suivant la nature physique des particules envisagées, une seule de ces catégories d'états se trouve réalisée dans la nature car, si à un instant donné une seule de ces catégories existe, l'autre n'apparaîtra jamais. La Mécanique ondulatoire n'impose pas l'existence d'une seule catégorie d'états (symétriques ou antisymétriques) pour chaque genre de particules, mais elle est compatible avec le principe supplémentaire suivant lequel, pour chaque genre de particules, existent exclusivement, soit les états symétriques, soit les états antisymétriques. L'expérience a montré qu'il en est bien ainsi.

Suivant que l'on suppose seule possible l'une ou l'autre de ces deux catégories d'états, on obtiendra pour le genre de particules considéré des statistiques très différentes. Si l'on suppose que les états symétriques sont seuls possibles, il n'y a aucune raison pour qu'on ne puisse pas avoir un nombre quelconque de particules dans le même état individuel et l'application des méthodes de la statistique quantique conduit alors à la statistique de Bose-Einstein. Les photons, les particules α , certains noyaux d'atome suivent cette statistique et sont donc des particules à états symétriques : on les nomme parfois des « Bosons ». Par contre, si l'on suppose seuls possibles les états antisymétriques, il ne pourra y avoir plus d'un corpuscule dans chaque état individuel. En effet, si nous avons deux corpuscules dans un même état antisymétrique, en permutant leur rôle, il devrait y avoir changement de signe de la fonction d'onde qui est, par hypothèse, antisymétrique, mais il ne devrait aussi y avoir aucun changement puisqu'on n'a fait que permuter deux corpuscules identiques occupant des états individuels identiques : la fonction d'onde, devant à la fois rester la même et changer de signe, est nécessairement nulle et ce fait signifie que cet état ne peut exister. On retrouve ainsi la règle d'exclusion de Pauli et, en appliquant les méthodes de la statistique quantique, on parvient ainsi à la statistique de Fermi-Dirac. Le principe de Pauli est donc seulement valable pour les particules à états antisymétriques. On trouve que les électrons,

les protons, les neutrons et certains noyaux d'atome sont des particules de ce genre et on leur donne parfois le nom générique de « Fermions ».

On pourrait donc croire que les considérations qui précèdent peuvent être appliquées directement aux électrons et fournir un énoncé mathématique du principe de Pauli. Il n'en est rien parce que les électrons sont des corpuscules doués de spin et qu'il est nécessaire de tenir compte du spin pour définir correctement la symétrie et l'antisymétrie des fonctions d'onde tandis que, jusqu'à présent dans ce chapitre, nous n'avons pas introduit le spin. D'ailleurs, l'énoncé du principe de Pauli pour les électrons doit faire intervenir, nous l'avons vu, les nombres quantiques j et m qui sont relatifs au moment de rotation total comprenant le spin, d'où l'impossibilité de retrouver correctement ce principe sans faire intervenir le spin.

§. **Introduction du spin en Mécanique ondulatoire des systèmes. Énoncé rigoureux du principe de Pauli.** — Jusqu'ici nous n'avons développé la Mécanique ondulatoire des systèmes que pour des corpuscules sans spin. Pour obtenir rigoureusement une Mécanique ondulatoire des systèmes de corpuscules doués de spin, il faudrait développer une théorie qui généralise la Mécanique ondulatoire relativiste de l'électron due à Dirac. Mais cette théorie comporterait nécessairement de grandes complications à cause du caractère « retardé » des interactions dans toute théorie relativiste et il faut se contenter d'un procédé approximatif. Ce qui permet de construire une théorie approximative, c'est que, du moins dans les conditions ordinaires, l'existence du spin n'introduit que de légers écarts par rapport à ce qui se passerait s'il n'existait pas : le spin n'introduit, en somme, que des structures fines au sens de Sommerfeld. Ceci nous permet de supposer que la situation d'un système d'électrons, en ce qui concerne la position et le mouvement de ses constituants, est très approximativement décrite par l'onde Ψ du système calculée à l'aide de l'équation de Schrödinger. On peut alors admettre que, pour tenir compte du spin, il suffit de multiplier cette fonction d'onde par une fonction représentant l'état des spins des constituants. La fonction d'onde Φ représentant entièrement l'état du système sera donc de la forme

$$\Phi(x_1, \dots, x_N, \sigma_1, \dots, \sigma_N, t) = \Psi(x_1, \dots, x_N, t) \zeta(\sigma_1, \dots, \sigma_N),$$

où $\Psi(x_1, \dots, x_N, t)$ est une solution de l'équation de Schrödinger et où $\sigma_1, \dots, \sigma_N$ représentent les spins des N électrons du système

projetés sur une direction de référence et susceptibles d'avoir les deux valeurs $\pm \frac{h}{4\pi}$.

Le principe de Pauli s'énoncera alors en disant :

Les seuls états possibles pour un système de N électrons sont ceux dont l'onde Φ est antisymétrique pour l'ensemble des quatre variables d'espace et de spin de toute paire d'électrons.

En d'autres termes, les fonctions Φ représentant les états du système réellement réalisables dans la nature sont telles que si l'on permute dans Φ les variables x_i, y_i, z_i, σ_i et x_j, y_j, z_j, σ_j pour une paire quelconque d'indices i et j , la fonction Φ change seulement de signe sans changer de valeur absolue.

Telle est l'expression mathématique rigoureuse du principe de Pauli.

6. Théorie du spectre de l'hélium (Heisenberg). — Le principe de Pauli est fortement confirmé par l'existence de la saturation des niveaux d'énergie atomiques, puisqu'il conduit à la règle de Stoner qui traduit les faits expérimentaux : il l'est aussi par le succès de la statistique de Fermi-Dirac, par exemple dans la théorie électronique des métaux. Mais une autre preuve très remarquable en a été apportée par la théorie du spectre de l'hélium due à Heisenberg. L'atome d'hélium fournit, en effet, un exemple très simple de système formé des corpuscules identiques, puisqu'il comprend deux électrons seulement soumis à l'action du champ coulombien du noyau et qu'en première approximation, celui-ci peut être considéré comme un simple centre de force.

Avant d'aborder le côté théorique du problème voyons ce que nous apprend l'expérience sur le spectre de l'hélium. Un fait fondamental est que ce spectre se décompose, à première vue, en deux catégories de raies complètement distinctes et formant comme deux spectres juxtaposés que l'on a nommés « spectre du parhélium » et « spectre de l'orthohélium ». Ce qui permet de dire que ces deux spectres sont complètement distincts, c'est que, si l'on dresse le tableau des niveaux (ou termes spectraux) correspondant aux raies du parhélium et à celles de l'orthohélium, il n'y a pas, du moins en première approximation quand on néglige certaines raies très faibles, de combinaisons entre les deux catégories de termes. Néanmoins, on ne peut pas dire, comme on avait cru tout d'abord pouvoir le faire, qu'il y a deux sortes d'hélium à atomes différents, le parhélium et l'orthohélium, car

les *mêmes* atomes d'hélium se montrent capables, suivant les circonstances, d'émettre l'un ou l'autre spectre. Il y avait là une circonstance mystérieuse qui n'a pu être interprétée que par la Mécanique ondulatoire.

Si l'on examine plus attentivement les spectres du parhélium et de l'orthohélium, on constate que l'on peut faire correspondre chaque niveau de l'un à un niveau de l'autre, mais que chaque niveau de l'orthohélium est un peu plus profond que son homologue du parhélium. Si l'on représente les terme du parhélium et de l'orthohélium, on obtient le schéma suivant de niveaux :

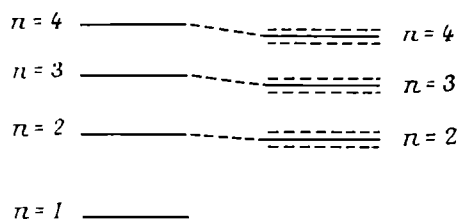


Fig. 30.

Un fait très remarquable visible sur ce schéma est que le niveau fondamental ($n = 1$) du parhélium n'a pas d'homologue dans le spectre de l'orthohélium.

Une étude plus approfondie a conduit à découvrir une structure fine du spectre de l'orthohélium, chaque niveau de ce spectre devant être considéré comme triple, ainsi que nous l'avons indiqué sous la figure par des traits pointillés. Il en résulte que le spectre du parhélium est un spectre de singulets, mais que le spectre de l'orthohélium est un spectre de triplets.

Tels sont les principaux faits expérimentaux, assez singuliers au premier abord, dont la théorie avait à rendre compte.

Pour faire la théorie du spectre de l'hélium, on remarquera d'abord que, d'après la règle de Stoner, l'atome d'hélium dans son état normal non ionisé doit avoir ses deux électrons sur le niveau fondamental $n = 1$. Par suite d'une perturbation extérieure, l'un de ces électrons peut être porté sur un niveau supérieur caractérisé par l'indice n et les raies du spectre usuel de l'hélium doivent être dues aux transitions de cet électron excité vers des états de moindre énergie. Le cas où les deux électrons se trouveraient excités à la fois doit être considéré comme très rare et ne donnant lieu qu'à des raies observées dans des circonstances très exceptionnelles. Donc les niveaux du spectre usuel de l'hélium

s'obtiendront en calculant les fonctions d'onde qui correspondent à un électron dans l'état fondamental et à l'autre électron dans l'état n . Nous les désignerons par $\Psi_{1,n}$.

On peut d'abord faire une théorie incomplète sans tenir compte du spin. On a alors à résoudre une équation de Schrödinger à six variables : $x_1, y_1, z_1, x_2, y_2, z_2$. On trouve des solutions symétriques $\Psi_{1,n}^{(s)}$ et des solutions antisymétriques $\Psi_{1,n}^{(a)}$ correspondant à des valeurs propres $E_n^{(s)}$ et $E_n^{(a)}$. Le calcul de ces valeurs propres fait par Heisenberg indique que $E_n^{(a)}$ est toujours un peu plus faible que $E_n^{(s)}$ pour une même valeur de n . Comme il n'y a pas d'intercombinaisons entre les états symétriques et les états antisymétriques, on a donc deux spectres indépendants, comme l'exige l'expérience, et les valeurs relatives de $E_n^{(s)}$ et de $E_n^{(a)}$ imposent de faire correspondre le spectre du parhélium aux états symétriques et le spectre de l'orthohélium aux états antisymétriques. Cette attribution est immédiatement confirmée d'une manière très remarquable par le fait suivant : la fonction propre antisymétrique $\Psi_{1,n}^{(a)}$ étant nulle si $n = 1$, car alors les deux électrons sont dans le même état individuel, le niveau fondamental $n = 1$ doit manquer dans le spectre de l'orthohélium et nous savons qu'il en est bien ainsi. On pourrait objecter que les états symétriques sont à écarter d'après le principe de Pauli, ce qui ferait tomber les interprétations précédentes, mais cette objection ne vaut pas parce que le principe de Pauli ne peut valablement s'appliquer que si l'on a tenu compte du spin et nous ne l'avons pas fait jusqu'à présent.

Pour compléter la théorie, il nous faut donc nécessairement introduire le spin de l'électron. Nous le ferons, comme nous l'avons indiqué plus haut, en multipliant la fonction d'onde Ψ des coordonnées, solution de l'équation de Schrödinger, par une fonction représentant l'état de spin des deux électrons, de façon à obtenir par ce produit une fonction d'onde représentant approximativement l'état du système, compte tenu du spin. Le spin de chacun des électrons pouvant avoir les deux valeurs $+\frac{h}{4\pi}$ et $-\frac{h}{4\pi}$, nous conviendrons de faire correspondre à la première de ces valeurs la lettre a et à la seconde la lettre b . Nous dirons alors que la fonction représentant le spin du système est $a(1)b(2)$ si le premier électron a le spin $\frac{h}{4\pi}$ et le second spin $-\frac{h}{4\pi}$, qu'elle est $a(1)a(2)$ si les deux électrons ont le spin $\frac{h}{4\pi}$, etc.

Heisenberg a montré que l'on est amené à introduire quatre états de spin pour le système représentés par les quatre fonctions

$$\begin{aligned} & a(1)a(2), \quad b(1)b(2), \quad a(1)b(2) + a(2)b(1), \\ & \quad \quad \quad a(1)b(2) - a(2)b(1), \end{aligned}$$

les trois fonctions de la première ligne représentant des états de spin symétriques, la fonction de la deuxième ligne un état de spin anti-symétrique.

Pour obtenir les fonctions d'onde $\Phi_{1,n}$ du système correspondant à un nombre n donné, on doit multiplier chacune des deux fonctions $\Psi_{1,n}^{(s)}$ et $\Psi_{1,n}^{(a)}$ précédemment utilisées par chacune des quatre fonctions de spin que nous venons d'écrire. On obtient ainsi les huit fonctions $\Phi_{1,n}$ suivantes :

$$\begin{aligned} \Psi_{1,n}^{(s)} a(1)a(2), \quad \Psi_{1,n}^{(s)} b(1)b(2), \quad \Psi_{1,n}^{(s)} [a(1)b(2) + a(2)b(1)], \\ \Psi_{1,n}^{(a)} [a(1)b(2) - a(2)b(1)] \end{aligned}$$

et

$$\begin{aligned} \Psi_{1,n}^{(a)} a(1)a(2), \quad \Psi_{1,n}^{(a)} b(1)b(2), \quad \Psi_{1,n}^{(a)} [a(1)b(2) + a(2)b(1)], \\ \Psi_{1,n}^{(s)} [a(1)b(2) - a(2)b(1)]. \end{aligned}$$

Les quatre premières sont symétriques, les quatre dernières anti-symétriques par rapport à l'ensemble des variables de configuration et des spins. D'après le principe de Pauli, nous devons admettre que, seules, les quatre dernières représentent des états réellement existant dans la nature.

Pour un nombre n donné, nous trouvons donc les quatre états suivants :

$$\begin{array}{ll} \Psi_{1,n}^{(s)} [a(1)b(2) - a(2)b(1)] & \text{avec l'énergie } E_n^{(s)} + \epsilon, \\ \Psi_n^{(a)} a(1)a(2) & \text{avec l'énergie } E_n^{(a)} + \eta_1, \\ \Psi_n^{(a)} b(1)b(2) & \text{avec l'énergie } E_n^{(a)} + \eta_2, \\ \Psi_n^{(a)} [a(1)b(2) + a(2)b(1)] & \text{avec l'énergie } E_n^{(a)} + \eta_3. \end{array}$$

En effet, l'intervention du spin modifie très légèrement les valeurs propres de la théorie sans spin : d'où l'apparition, dans l'expression des énergies, des très petites constantes ϵ , η_1 , η_2 , η_3 . Ces corrections étant très petites, on a toujours le résultat que la valeur de l'énergie correspondant à la fonction $\Psi_{1,n}^{(s)}$ est un peu supérieure aux énergies des états en $\Psi_{1,n}^{(a)}$. On est donc encore amené à faire correspondre l'état unique en $\Psi_{1,n}^{(s)}$ au spectre du parhélium et les trois états très voisins en $\Psi_{1,n}^{(a)}$ au spectre de l'orthohélium.

Comme $\Psi_{1,1}^{(a)}$ est nul, on a toujours l'interprétation de l'absence

du niveau fondamental $n = 1$ dans le spectre de l'orthohélium. Mais ici nous avons une précision supplémentaire, car η_1, η_2, η_3 étant de petites constantes très légèrement différentes, les niveaux de l'orthohélium doivent être triples et tous les faits expérimentaux représentés sur le schéma donné ci-dessus se trouvent ainsi expliqués.

Nous avons obtenu par cette théorie une très bonne vérification du principe de Pauli, car si l'on supposait, par exemple, qu'il faut pour les électrons conserver les états symétriques par rapport aux coordonnées et aux spins et rejeter les états antisymétriques, on serait conduit à prévoir que les raies de l'orthohélium sont des singulets et celles du parahélium des triplets, ce qui serait l'inverse de la constatation expérimentale. Le principe de Pauli se trouve ainsi en quelque sorte inscrit dans la structure fine du spectre de l'hélium.

La théorie dont nous venons de donner un exposé qualitatif est susceptible de vérifications numériques. Heisenberg et Unsöld ont pu calculer exactement de cette façon la valeur des termes spectraux de l'hélium et de son potentiel d'ionisation.

Nous devons enfin noter qu'il existe, en réalité, dans le spectre de l'hélium des raies très faibles correspondant à des combinaisons entre les niveaux du parahélium et ceux de l'orthohélium (Jacquinot, 1937), ce qui prouve en passant que les deux sortes de niveaux existent bien dans le même atome. Heisenberg a pu expliquer la présence de ces raies de la façon suivante. La démonstration que les états $\Psi^{(s)}$ et $\Psi^{(a)}$ ne peuvent se combiner n'est rigoureuse que lorsqu'on néglige le spin : si l'on tient compte du spin, ce sont les fonctions Φ symétriques et les fonctions Φ antisymétriques qui ne peuvent se combiner, ce qui permet de ne conserver pour l'électron que ces dernières. Mais comme en introduisant le spin on introduit des interactions entre moments magnétiques propres qui ne dépendent pas seulement de la distance des électrons, mais aussi de l'orientation de leurs moments, on ne peut plus démontrer que les états en $\Psi^{(s)}$ et en $\Psi^{(a)}$ ne se combinent pas. Les transitions entre les niveaux du parahélium et ceux de l'orthohélium sont donc possibles, mais comme leurs probabilités sont de l'ordre des corrections de spin, les raies correspondantes sont très faibles.

7. Orthohydrogène et parahydrogène. — La théorie des spectres moléculaires de la molécule H_2 , quand on fixe son attention sur le rôle des deux protons noyaux des deux atomes H , amène à introduire des considérations tout à fait analogues à celles que nous venons

de rencontrer en étudiant le spectre de l'hélium, et cela parce que les protons comme les électrons sont des Fermions.

Le spectre de rotation de la molécule H_2 est dû aux transitions de la molécule entre états de rotation quantifiés. Dans ces états de rotation, la molécule H_2 se comporte sensiblement comme une haltère dont les noyaux H formeraient les masses terminales, les deux électrons de la molécule H_2 jouant ici un rôle négligeable. Le système des deux noyaux H admet des états stationnaires, les valeurs propres de l'énergie correspondant aux niveaux du spectre de rotation et les fonctions propres décrivant les états de rotation quantifiés. Par analogie avec la théorie du spectre de l'hélium, on voit alors aisément que les niveaux de rotation comprendront une série de niveaux simples du type para avec fonction propre symétrique d'espace et une série de niveaux triples du type ortho avec fonction propre antisymétrique d'espace. La transition d'un niveau para à un niveau ortho et inversement ne pourra être due qu'aux interactions de spin entre les protons; mais le moment magnétique du proton étant environ 600 fois plus petit que celui de l'électron, les interactions de spin entre protons doivent être beaucoup plus faibles encore que les interactions de spin entre électrons. Bref, pour l'émission du spectre de rotation de la molécule H_2 , tout se passera comme s'il existait deux sortes d'hydrogène, l'orthohydrogène et le parahydrogène : la séparation du spectre ortho et du spectre para sera, en ce cas, plus complète encore que dans le cas du parhélium et de l'orthohélium.

Voici une conséquence curieuse de ce qui précède. Lorsqu'après une excitation suivie par l'émission de raies du spectre de rotation, certaines molécules H_2 sont revenues à l'état d'énergie minimum de l'orthohydrogène, elles ne pourront ensuite revenir que très difficilement à l'état fondamental du parahydrogène (qui correspond au minimum absolu de l'énergie), puisqu'une telle transition ne se produit qu'avec une probabilité extrêmement faible. L'état d'énergie minimum de l'orthohydrogène présente donc une métastabilité très marquée. En maintenant l'hydrogène à une très basse température où pratiquement la totalité des molécules H_2 devrait se trouver dans l'état de minimum absolu de l'énergie, il faut en réalité près d'une semaine pour que toutes les molécules reviennent à cet état à cause de la grande métastabilité du niveau inférieur de l'orthohydrogène.

L'existence des deux spectres moléculaires presque entièrement séparés de l'ortho et du parahydrogène se manifeste aussi d'une autre

manière. Si nous excitons brutalement par action d'une forte température des molécules H_2 , nous verrons apparaître le spectre de rotation au complet avec ses raies ortho et para, les premières étant, toutes choses égales d'ailleurs, trois fois plus fortes que les secondes car le poids des niveaux de l'orthohydrogène est triple de celui des niveaux du parahydrogène. Mais commençons par maintenir de l'hydrogène à très basse température pendant un temps très long de façon à amener pratiquement toutes les molécules dans l'état fondamental du parahydrogène, puis portons l'hydrogène ainsi préparé dans un tube de Plücker : l'excitation électrique portera presque toujours les molécules H_2 de l'état fondamental du parahydrogène à un autre état du parahydrogène en raison de la difficulté des transitions para \rightarrow ortho. On n'obtiendra donc tout d'abord que les raies du parahydrogène, celles de l'orthohydrogène qui devraient normalement être trois fois plus intenses étant absentes. Peu à peu cependant, les raies de l'orthohydrogène feront leur apparition au fur et à mesure que des molécules H_2 ayant subi une transition para \rightarrow ortho deviendront capables de les émettre.

La difficulté des transitions para \rightarrow ortho et des transitions inverses a aussi comme conséquence que des équilibres thermodynamiques ne peuvent s'établir rapidement qu'entre des molécules H_2 dans des états para ou entre des molécules H_2 dans des états ortho. Il en résulte que dans des mesures faites rapidement à température pas trop basse telles que les mesures de chaleur spécifiques, l'hydrogène doit se comporter comme un mélange de deux gaz différents, l'orthohydrogène et le parahydrogène, mélangés dans la proportion de 3 à 1 et dont l'un ne serait susceptible que d'états antisymétriques d'espace et l'autre d'états symétriques d'espace. Cette remarque a permis à Heisenberg d'expliquer quantitativement les anomalies apparentes qui avaient été observées par Dennison pour les chaleurs spécifiques de l'hydrogène.

Tout comme le parahélium et l'orthohélium, le parahydrogène et l'orthohydrogène ne sont pas des éléments différents, mais correspondent seulement à deux catégories de niveaux quantifiés, difficilement combinables entre elles, d'un même système quantifié qui est l'atome He dans un cas et la molécule H_2 dans l'autre.

8. Applications diverses. — La Mécanique ondulatoire des systèmes de corpuscules a permis, en introduisant la notion nouvelle d'énergie d'échange, de rendre compte de la stabilité des molécules homopolaires et, en particulier, de faire une belle théorie de la liaison des deux atomes H dans la molécule H_2 (Heitler et London).

Si l'on veut appliquer la Mécanique ondulatoire des systèmes à des atomes plus compliqués que celui d'hélium, on tombe dans de grandes difficultés mathématiques et il a fallu des méthodes puissantes comme celles de la théorie des Groupes pour parvenir à des résultats généraux.

Néanmoins, des méthodes d'approximation ingénieuses ont permis d'aborder les questions relatives à l'interaction des atomes dans les molécules, à l'interprétation de la notion chimique de valence, à la structure et aux propriétés des molécules complexes de la Chimie organique, aux phénomènes de Mésomérie, etc. Il y a là toute une science nouvelle, la Chimie théorique quantique, qui est en plein développement, mais dont nous ne pouvons entreprendre ici l'exposé. Nous ne pouvons que renvoyer aux Ouvrages spéciaux [5], [6], [7], [8].

BIBLIOGRAPHIE.

- [1] N. BOHR, *Les spectres et la structure de l'atome*, Hermann, Paris, 1923,
- [2] E. SCHRÖDINGER, *Mémoires sur la Mécanique ondulatoire*, déjà cité.
- [3] LOUIS DE BROGLIE, *La Mécanique ondulatoire des systèmes de corpuscules*, Gauthiers-Villars, Paris, 2^e édition, 1950.
- [4] LOUIS DE BROGLIE, *De la Mécanique ondulatoire à la théorie du noyau*, t. I, Hermann, Paris, 1943.
- [5] H. WEYL, *Gruppentheorie und quantenmechanik*, Hirzel, Leipzig, 1928.
- [6] L. PAULING et E. B. WILSON, *Introduction to quantum-mechanics*, Mac-Graw and Hill, Londres, 1935.
- [7] P. et R. DAUDEL, *Les apports de la Mécanique ondulatoire à l'étude de la molécule*, Paris, 1950.
- [8] R. DAUDEL, *Les fondements de la Chimie théorique* (à paraître chez Gauthier-Villars dans cette collection).

CHAPITRE XVI.

LES STATISTIQUES QUANTIQUES.

1. Les cellules d'extension-en-phase dans la théorie des Quanta. — Dans le chapitre IV, nous avons exposé les principes de la Mécanique statistique en introduisant l'hypothèse que l'extension-en-phase était subdivisée en cellules de grandeur α et nous avons obtenu ainsi une expression de l'entropie où la constante de l'entropie dépend de $\log \alpha$. Or nous avons remarqué que dans la Mécanique statistique classique, tous les points de l'extension-en-phase représentant des états possibles, il est logique d'admettre que α est nulle et, par suite, de faire tendre α vers zéro dans toutes les formules obtenues. On tombe alors sur la difficulté que la constante de l'entropie devrait être infinie. La théorie des Quanta a levé cette difficulté en attribuant à α une valeur très petite, mais finie.

En effet, dès le début de ses recherches sur les quanta, Planck avait remarqué que la théorie des Quanta conduit à subdiviser l'extension-en-phase d'une molécule en cellules finies de grandeur h^3 . Au point de vue des dimensions, cette hypothèse est évidemment admissible car, le produit $dx dp_x$ ayant comme h les dimensions d'une action ML^2T^{-1} , l'élément d'extension-en-phase pour une molécule a bien les dimensions de h^3 . Le quantum h jouant dans la théorie quantique le rôle d'une sorte d'atome d'action, il est donc naturel de prendre la grandeur des cellules d'extension-en-phase égale à h^3 .

On peut justifier cette hypothèse d'une façon plus précise en raisonnant comme il suit. Considérons le cas d'un mouvement périodique quantifié défini par une seule coordonnée q . La condition de quanta adoptée par Planck pour les mouvements quantifiés était $\oint p dq = nh$: elle exprime que la trajectoire fermée décrite par le point figuratif dans

L'extension-en-phase à deux dimensions (q, p) limite une aire intérieure égale à nh .

Puisque n est un entier, cela revient à dire que l'extension-en-phase est divisée en cellules élémentaires de grandeur h . D'après les conditions de Wilson-Sommerfeld, pour un mouvement périodique à n coordonnées, la même conclusion est exacte pour chacun des plans $q_i p_i$ et, comme l'extension-en-phase a alors $2n$ dimensions, il en résulte que cette extension est divisée en cellules élémentaires de volume h^n . Pour l'extension-en-phase d'une molécule qui est à six dimensions, les cellules élémentaires ont bien le volume h^3 .

Dès ses débuts, la Mécanique ondulatoire a fourni une interprétation

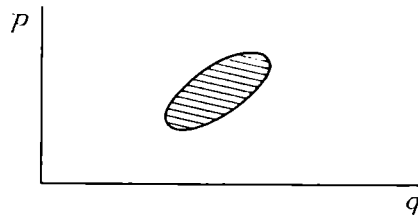


Fig 31.

très intéressante de la division apparente de l'extension-en-phase d'une molécule en cellules élémentaires de volume h^3 . Dire que l'extension-en-phase est divisée en cellules de volume h^3 , cela revient en effet à dire que l'intervalle de quantité de mouvement ($p, p + dp$) correspond pour la molécule à un nombre d'états stables possibles égal à $\frac{4\pi p^2 dp}{h^3} \mathcal{V}$ (voir chap. IV), puisqu'ici $a = h^3$. Or la Mécanique ondulatoire explique très bien ce fait.

En effet, nous rappelons d'abord que, si l'on calcule le nombre des ondes stationnaires qui peuvent exister dans une enceinte de volume \mathcal{V} sous la condition d'être nulles aux limites de l'enceinte, on trouve qu'il y en a un nombre

$$(1) \quad dn = \frac{4\pi v^2 dv}{UV^2} \mathcal{V}$$

correspondant à l'intervalle de fréquence ($\nu, \nu + d\nu$). V est ici la vitesse de phase des ondes considérées pour la fréquence ν et U la vitesse de groupe définie par la formule de Rayleigh (voir chap. IX, § 6)

$$U = \frac{d}{d\nu} \frac{1}{\lambda} = \frac{d}{d\nu} \frac{v}{V} = \frac{1}{V} - \frac{v}{V^2} \frac{dV}{d\nu}$$

qui diffère de V si V dépend de ν . Si les ondes considérées sont les ondes lumineuses et si l'enceinte est vide, on a $V = U = c$ et l'on trouve simplement

$$dn = \frac{4\pi\nu^2 d\nu}{c^3} \mathfrak{V},$$

résultat qu'il faut encore doubler pour tenir compte des deux états de polarisation possibles de la Lumière. Nous retrouvons ainsi la formule due à Jeans que nous avons rencontrée au chapitre V, paragraphe 4. La démonstration de la formule générale (1) se trouve dans le livre de Léon Brillouin sur les Statistiques quantiques [1] à la page 36.

Mais, en Mécanique ondulatoire, à un mouvement rectiligne uniforme d'une molécule avec une quantité de mouvement p , est associée la propagation d'une onde Ψ de longueur d'onde

$$\lambda = \frac{h}{p}.$$

On a donc

$$p = \frac{h}{\lambda} = h \frac{\nu}{V}.$$

Nous pouvons donc écrire la formule (1) pour les ondes associées à la molécule sous la forme

$$\frac{dn}{\mathfrak{V}} = \frac{4\pi}{h^3} \left(\frac{h\nu}{V}\right)^2 \frac{h d\nu}{U} = \frac{4\pi}{h^3} \left(\frac{h\nu}{V}\right)^2 h d\left(\frac{\nu}{V}\right) = \frac{4\pi p^2 dp}{h^3}.$$

Or, en Mécanique ondulatoire, les seuls mouvements stables possibles sont ceux qui correspondent à des ondes stationnaires : c'est le principe même de la quantification des mouvements dans la nouvelle Mécanique. Le nombre dn représente donc le nombre des états possibles pour la molécule dans l'enceinte qui correspondent à l'intervalle $(p, p + dp)$. Or pour ce nombre, la Mécanique statistique nous fournit la valeur $\frac{4\pi p^2 dp}{a} \mathfrak{V}$, où a est la grandeur de la cellule d'extension-en-phase. On voit donc bien que tout doit se passer comme si l'extension-en-phase était divisée en cellules élémentaires de volume h^3 et cela parce que les états possibles sont associés à des ondes stationnaires pouvant s'établir dans l'enceinte envisagée.

L'ancienne théorie des Quanta avait donc posé $a = h^3$ et montré qu'en conservant la méthode statistique de Boltzmann-Gibbs, on retrouvait les formules classiques de la Thermodynamique statistique, mais cette fois avec une valeur finie de la constante de l'entropie, valeur qui était entièrement calculable pour un gaz monoatomique parfait à partir

de la valeur numérique de h . Stern avait indiqué comment on pouvait vérifier la valeur de la constante de l'entropie par l'étude de la loi des tensions de vapeurs saturantes. L'accord obtenu a été satisfaisant.

Mais le développement des théories quantiques a conduit à penser qu'il ne suffisait pas de compléter les méthodes statistiques classiques en posant $a = h^3$, mais qu'il fallait leur apporter des modifications plus profondes que nous allons maintenant examiner.

2. Nouvelles méthodes statistiques. — La Mécanique statistique doit avant tout faire le dénombrement des répartitions élémentaires qui réalisent une certaine répartition globale des constituants d'un système complexe entre les diverses énergies. Dans la méthode classique de Boltzmann, ce dénombrement doit se faire en considérant comme différentes deux répartitions élémentaires dont la seule différence est que le rôle de deux particules de même nature y est permuté. En d'autres termes, on accordait aux particules une individualité suffisante pour qu'il faille distinguer le cas où la $i^{\text{ème}}$ particule est dans un état a et la $k^{\text{ème}}$ dans un état b et le cas où la $i^{\text{ème}}$ particule est dans l'état b et la $k^{\text{ème}}$ dans l'état a . Or le développement de la Mécanique ondulatoire des systèmes de particules de même nature a montré qu'il n'y avait pas lieu, pour un tel système, de considérer comme différents deux états qui ne diffèrent que par la permutation du rôle de deux de ces particules. Pour faire la statistique d'un tel système, on doit donc compter les complexions en tenant compte de ce fait, c'est-à-dire en considérant comme identiques deux complexions qui ne diffèrent que par la permutation de deux particules de même nature. En faisant le dénombrement de cette façon, on trouvera évidemment des nombres de complexions inférieurs à ceux que donnait la méthode classique et toutes les formules en seront modifiées.

Mais, de plus, il y a lieu de distinguer deux cas très différents. Nous avons vu, en effet, que les particules se divisent en deux catégories : les unes (bosons) pour lesquelles la fonction d'onde est toujours symétrique, les autres (fermions) pour lesquelles elle est toujours antisymétrique. Nous avons remarqué que, pour un système formé de particules identiques de la première catégorie, il pouvait y avoir un nombre quelconque de particules dans le même état tandis que, pour un système formé de particules de la seconde catégorie, il ne peut y avoir plus d'une particule par état.

Soit alors un système formé d'un très grand nombre N de particules de même nature. Nous voulons trouver le nombre de répartitions élé-

mentaires qui réalisent une répartition globale donnée $(n_1; \dots; n_i; \dots)$ entre les divers intervalles d'énergie $(E_1, E_1 + dE_1; \dots; E_i, E_i + dE_i; \dots)$. Nous savons évaluer le nombre g_i de cellules d'extension-en-phase afférentes à chaque intervalle, maintenant que nous savons que la grandeur de ces cellules est h^3 . Nous supposons que les dE_i sont des infiniment petits physiques, c'est-à-dire qu'ils sont suffisamment petits pour permettre de remplacer les sommations sur les intervalles par des intégrations, mais qu'ils sont cependant suffisamment grands pour que les g_i soient très grands : ces deux suppositions ne sont pas contradictoires à cause de l'extrême petitesse des cellules h^3 .

Dès lors, si les particules considérées sont des bosons, nous aurons à calculer le nombre de manières dont on peut répartir les n_i particules du $i^{\text{ème}}$ intervalle entre les g_i cellules de cet intervalle sans nous préoccuper de l'individualité des particules et sans limiter le nombre des particules qui peuvent se trouver dans une même cellule. Nous obtiendrons donc le nombre des répartitions de n_i particules entre les g_i cellules en prenant de toutes les façons possibles n_i des g_i cellules, chaque cellule pouvant être prise plusieurs fois. Le nombre des répartitions cherché est donc le nombre des combinaisons *avec* répétition de g_i objets pris n_i à n_i qui est, d'après l'analyse combinatoire, égal à $\frac{(n_i + g_i - 1)!}{n_i! (g_i - 1)!}$. En partant de là, nous construirons une statistique quantique qu'on appelle la statistique de Bose-Einstein.

Si, au contraire, nos particules sont des fermions, nous aurons à calculer le nombre de manières dont on peut répartir les n_i particules du $i^{\text{ème}}$ intervalle entre les g_i cellules de cet intervalle sans nous préoccuper de l'individualité des particules, mais en limitant à une le nombre des particules par cellule. Nous obtiendrons donc le nombre de répartitions cherché en prenant de toutes les façons possibles n_i des g_i cellules, chaque cellule ne pouvant maintenant être prise qu'une fois (ce qui exige d'ailleurs $n_i \leq g_i$). Le nombre des répétitions cherché est donc le nombre des combinaisons *sans* répétition de g_i objets pris n_i à n_i : une formule élémentaire d'analyse combinatoire nous apprend qu'il est égal à $\frac{g_i!}{n_i! (g_i - n_i)!}$. Partant de là, nous construirons une autre statistique quantique qu'on nomme la statistique de Fermi-Dirac.

Nous allons maintenant effectivement développer ces deux statistiques.

3. La statistique de Bose-Einstein. — Considérons un système formé d'un très grand nombre N de particules identiques pour lesquelles les

états symétriques sont seuls possibles. Nous avons vu que le nombre des répartitions des n_i particules du $i^{\text{ième}}$ intervalle d'énergie entre les g_i cellules d'extension-en-phase de cet intervalle est donné par le nombre des combinaisons avec répétition de g_i objets pris n_i à n_i . Ce nombre est égal à $\frac{(n_i + g_i - 1)!}{n_i! (g_i - 1)!}$.

La démonstration de cette formule par l'analyse combinatoire est un peu longue et ne se trouve pas dans les traités élémentaires. Nous allons en donner une ingénieuse démonstration imagée qui a été indiquée par M. Fermi. Représentons les n_i particules du $i^{\text{ième}}$ intervalle d'énergie par des cercles et alignons ces n_i cercles. Traçons ensuite $g_i - 1$ traits verticaux séparant ces cercles en groupes suivant le schéma suivant :

$$ooo | oo || oooo | \dots \dots \dots oo | oo | o.$$

Ce schéma représente une des distributions possibles des n_i particules entre les g_i cellules, car nos $g_i - 1$ traits verticaux répartissent les cercles en g_i sections et nous pouvons supposer que les cercles de la première section représentent des particules placées dans la première cellule, etc. Or, nous obtiendrons tous les schémas du type ci-dessus en permutant de toutes les façons possibles les n_i cercles et les $g_i - 1$ traits verticaux, ce qui fait au total $n_i + g_i - 1$ objets. Il y a $(n_i + g_i - 1)!$ permutations de cette sorte. Mais parmi les $(n_i + g_i - 1)!$ schémas ainsi obtenus, tous ceux qui ne diffèrent que par la permutation de deux cercles ou par la permutation de deux traits représentent la même répartition des n_i particules identiques entre les g_i cellules puisqu'on n'attribue plus d'individualité aux particules. Donc le nombre des répartitions distinctes est bien $\frac{(n_i + g_i - 1)!}{n_i! (g_i - 1)!}$ et, comme on suppose g_i très grand par rapport à l'unité, on peut écrire approximativement $\frac{(n_i + g_i)!}{n_i! g_i!}$. Le nombre des manières de réaliser la répartition globale $(n_1; \dots; n_i; \dots)$ entre les intervalles d'énergie est donc

$$P = \prod_i \frac{(n_i + g_i)!}{n_i! g_i!}$$

et, en employant la formule de Stirling, on trouve

$$\begin{aligned} \log P &= \sum_i [\log(n_i + g_i)! - \log n_i! - \log g_i!] \\ &= \sum_i [(n_i + g_i)(\log(n_i + g_i) - 1) - n_i(\log n_i - 1) - g_i(\log g_i - 1)] \\ &= \sum_i [(n_i + g_i) \log(n_i + g_i) - g_i \log g_i - n_i \log n_i]. \end{aligned}$$

Pour déterminer l'état le plus probable, seul réalisé en pratique, écrivons que $\delta \log P$ est nul pour une variation des n_i , compte tenu des deux conditions

$$\sum_i \delta n_i = 0, \quad \sum_i E_i \delta n_i = 0.$$

En introduisant des multiplicateurs de Lagrange α et β , on écrit

$$0 = \delta \log P + \alpha \sum_i \delta n_i + \beta \sum_i E_i \delta n_i = \sum_i \delta n_i [\log(n_i + g_i) - \log n_i + \alpha + \beta E_i]$$

et, les δn_i pouvant maintenant être pris arbitrairement, chaque crochet de la somme doit être nul, ce qui donne aisément

$$(2) \quad n_i = \frac{g_i}{e^{\alpha + \beta E_i - 1}}.$$

En raisonnant comme nous l'avons fait en Mécanique statistique classique au chapitre IV, paragraphe 5, on trouve que $\beta = \frac{1}{kT}$. Quant à α , sa valeur est déterminée par la condition $\sum_i n_i = N$.

Supposons que notre système de particules identiques forme un gaz monoatomique parfait (gaz d'Einstein). Alors on doit poser

$$g_i = \frac{4\pi p^2 dp}{h^3} \varrho = \frac{2\pi(2m)^{\frac{3}{2}} \sqrt{E} dE}{h^3} \varrho$$

et l'on trouve pour le nombre des atomes dont l'énergie appartient à l'intervalle $(E, E + dE)$:

$$dn_E = \frac{2\pi(2m)^{\frac{3}{2}}}{h^3} \varrho \frac{\sqrt{E} dE}{e^{\alpha + \frac{E}{kT} - 1}}.$$

Si e^α est très grand, on pourra négliger l'unité au dénominateur et l'on retrouvera la loi de distribution classique de Maxwell. Si, au contraire, e^α est petit devant l'unité, on aura une loi très différente de celle de Maxwell : on dit alors que le gaz est « dégénéré ». Or e^α est déterminé par l'équation

$$\int dn = N.$$

En faisant les calculs, on peut montrer que l'ordre de grandeur de e^α dépend du paramètre sans dimensions

$$D = \frac{N h^3}{\varrho (2\pi m k T)^{\frac{3}{2}}}$$

nommé « paramètre de dégénérescence ». Si D est petit devant l'unité, e^{α} est grand et la statistique classique est très approximativement valable. Si D est grand devant l'unité, le gaz est dégénéré et ses propriétés peuvent différer beaucoup des propriétés classiques.

Il est facile de voir que D est toujours très petit pour les gaz matériels dans les conditions usuelles. On ne doit donc pas s'attendre à pouvoir mettre aisément en évidence des écarts par rapport aux propriétés classiques provenant de la dégénérescence.

Si l'on calcule l'énergie totale U du gaz et sa pression p en supposant D assez petit pour qu'on puisse négliger son carré, on trouve

$$U = \frac{3}{2} N k T \left(1 - \frac{D}{2^{\frac{3}{2}}} \right), \quad p = \frac{N k T}{\mathcal{V}} \left(1 - \frac{D}{2^{\frac{3}{2}}} \right).$$

Le gaz doit donc, pour un volume donné, avoir une pression plus faible que ne l'indique la loi de Mariotte-Gay-Lussac. Les cas les plus favorables (hélium à 4° absolus) donnent des écarts de quelques pour-cent par rapport à la loi classique, mais ces écarts, observables en principe, peuvent être masqués par les écarts provenant de ce qu'aucun gaz n'est parfait. Néanmoins, les propriétés étranges de l'hélium à très basse température paraissent apporter la confirmation que ce gaz obéit à la statistique de Bose-Einstein (les noyaux He ou particules α sont des bosons).

4. Application de la statistique de Bose-Einstein au rayonnement noir. — La statistique de Bose-Einstein reçoit une vérification importante dans l'étude du rayonnement noir. Du point de vue de la théorie des Photons, le rayonnement noir est assimilable à un « gaz de photons ». Il est d'ailleurs aisé de calculer la valeur des g_i pour les photons. Les photons dont l'énergie est comprise entre $h\nu$ et $h(\nu + d\nu)$ ont des quantités de mouvement comprises entre $\frac{h\nu}{c}$ et $\frac{h(\nu + d\nu)}{c}$; on a donc

$$p = \frac{h\nu}{c}, \quad dp = \frac{h d\nu}{c}, \quad \frac{4\pi p^2 dp}{h^3} = \frac{4\pi\nu^2 d\nu}{c^3}.$$

Il faut encore doubler ce nombre pour tenir compte des deux états de polarisation possibles du photon et l'on obtient finalement

$$g_i = \frac{8\pi\nu^2 d\nu}{c^3} \mathcal{V}$$

et c'est précisément le nombre des ondes stationnaires de fréquences comprises entre ν et $\nu + d\nu$ qui peuvent exister dans l'enceinte \mathcal{V} .

Nous avons maintenant à appliquer aux photons, qui sont des bosons, la formule (2), mais ici il faut faire un peu attention parce qu'en présence de matière le rayonnement noir peut gagner ou perdre des photons si la matière présente en émet ou en absorbe. L'équilibre du rayonnement noir en présence de la matière n'est donc pas absolument comparable à celui d'un gaz enfermé dans une enceinte, mais plutôt à celui d'une vapeur saturante en présence de sa phase condensée. Il en résulte que, le nombre des photons pouvant varier, on ne doit pas imposer au gaz de photons la condition $\sum_i \delta n_i = 0$ et, si nous reprenons la démonstration de la formule donnant les n_i , nous devons poser $\alpha = 0$. On obtient donc

$$n_i = \frac{g_i}{\frac{E_i}{e^{kT}} - 1}$$

et l'on en conclut que le nombre des photons dans le rayonnement noir qui correspond à l'intervalle de fréquence $(\nu, \nu + d\nu)$ est donné par

$$dn_\nu = \frac{8\pi\nu^2 d\nu}{c^3} \frac{1}{\frac{h\nu}{e^{kT}} - 1}$$

par unité de volume. D'où pour la densité spectrale

$$\rho(\nu) d\nu = dn_\nu h\nu = \frac{8\pi h\nu^3}{c^3} \frac{1}{\frac{h\nu}{e^{kT}} - 1} d\nu.$$

On retrouve bien ainsi la loi de Planck, ce qui montre que les photons sont bien des particules à états symétriques obéissant à la statistique de Bose-Einstein. La statistique classique nous aurait conduit à la formule

$$\rho(\nu) d\nu = \frac{8\pi h\nu^3}{c^3} e^{-\frac{h\nu}{kT}} d\nu$$

ayant la forme de la loi empirique de Wien qui n'est valable que pour les grandes valeurs de $\frac{\nu}{T}$. Quant à la statistique de Fermi-Dirac, elle nous aurait donné $+1$ au lieu de -1 au dénominateur, ce qui serait tout à fait incompatible avec l'expérience.

§. La statistique de Fermi-Dirac. — Considérons un système formé d'un grand nombre N de particules identiques ne possédant que des

états antisymétriques et soumises au principe de Pauli (Fermions). Le nombre des répartitions des n_i particules du $i^{\text{ième}}$ intervalle d'énergie entre les g_i cellules d'extension-en-phase de cet intervalle est donné par le nombre des combinaisons *sans* répétition de g_i objets pris n_i à n_i . Ce nombre est égal à $\frac{g_i!}{n_i!(g_i - n_i)!}$ avec la condition $n_i \leq g_i$. La démonstration de cette formule est classique et se trouve dans les traités élémentaires.

Le nombre des répartitions élémentaires réalisant une répartition globale $(n_1; \dots; n_i; \dots)$ entre les intervalles d'énergie est donc

$$P = \prod_i \frac{g_i!}{n_i!(g_i - n_i)!}.$$

et la formule de Stirling nous donne

$$\begin{aligned} \log P &= \sum_i [g_i(\log g_i - 1) - n_i(\log n_i - 1) - (g_i - n_i)(\log(g_i - n_i) - 1)] \\ &= \sum_i [g_i \log g_i - n_i \log n_i - (g_i - n_i) \log(g_i - n_i)]. \end{aligned}$$

Pour déterminer l'état le plus probable, nous devons toujours écrire que $\delta \log P$ est nul pour une variation des n_i soumise aux conditions

$$\sum_i \delta n_i = 0 \quad \text{et} \quad \sum_i E_i \delta n_i = 0.$$

L'introduction des multiplicateurs de Lagrange nous donne

$$0 = \delta \log P + \alpha \sum_i \delta n_i + \beta \sum_i E_i \delta n_i = \sum_i \delta n_i [\log(g_i - n_i) - \log n_i + \alpha + \beta E_i].$$

Ceci devant être vérifié quels que soient les δn_i , tous les crochets de la somme doivent être nuls, ce qui nous donne aisément

$$(3) \quad n_i = \frac{g_i}{e^{\alpha + \beta E_i + 1}}.$$

La formule (3) a une forme analogue à celle de la formule (2), mais avec la différence essentielle qu'il y a $+1$ au lieu de -1 au dénominateur. On démontre, toujours par le même genre de raisonnement que dans les cas précédents, que $\beta = \frac{1}{kT}$. Quant à la constante α , elle peut se déterminer par la condition $\sum_i n_i = N$.

Le quotient $\frac{n_i}{g_i}$ nous donne le nombre moyen de particules par cellule pour le $i^{\text{ème}}$ intervalle d'énergie. D'après (3), ce nombre est toujours inférieur à l'unité, ce qui est bien en accord avec le principe de Pauli.

Nous pouvons supposer que notre système de particules à états anti-symétriques forme un gaz monoatomique parfait (gaz de Fermi). Si le spin des particules a s valeurs possibles, il faut multiplier par s la valeur de g_i calculée au paragraphe 3 et l'on a

$$g_i = s \frac{2\pi(2m)^{\frac{3}{2}} \sqrt{E} dE}{h^3} \mathcal{Q}.$$

L'introduction du nombre entier s correspond à l'introduction du facteur 2 dans le cas des photons. Pour les électrons, il faut poser $s = 2$. Le nombre des particules pour l'intervalle d'énergie $(E, E + dE)$ est donc

$$dn = \frac{2\pi s(2m)^{\frac{3}{2}}}{h^3} \mathcal{Q} \frac{\sqrt{E} dE}{e^{\frac{E}{kT} + 1}}.$$

Si $e^{\frac{E}{kT}}$ est très grand devant l'unité, on retrouve la loi de Maxwell et les résultats de la statistique classique. Si, au contraire, $e^{\frac{E}{kT}}$ est petit devant l'unité, le gaz est « dégénéré » et ses propriétés s'écartent beaucoup des propriétés classiques. En calculant $e^{\frac{E}{kT}}$ par la condition $\int dn = N$, on peut encore montrer que l'ordre de grandeur de $e^{\frac{E}{kT}}$ dépend de la valeur du « paramètre de dégénérescence » dont l'expression est toujours

$$D = \frac{N h^3}{\mathcal{Q} (2\pi m k T)^{\frac{3}{2}}}.$$

Pour les gaz dans les conditions usuelles, D est toujours petit devant l'unité et l'on ne peut avoir que de très petits écarts par rapport aux lois classiques. Par exemple, on trouve pour la pression aux termes d'ordre supérieur à D près

$$p = \frac{N k T}{\mathcal{Q}} \left(1 + \frac{D}{2^{\frac{5}{2}}} \right).$$

La pression pour un volume donné est donc ici plus grande que ne le prévoit la loi de Mariotte-Gay-Lussac et le gaz de Fermi s'écarte de la loi classique en sens inverse du gaz d'Einstein. L'écart, toujours assez petit (8 % au maximum), est masqué par les écarts dus au caractère non parfait des gaz réels.

6. **Le gaz de Fermi dans un état fortement dégénéré.** — En vue des applications à la théorie des électrons libres dans les métaux, il est intéressant d'étudier les propriétés d'un gaz de Fermi fortement dégénéré.

Nous considérons donc un gaz de Fermi pour lequel D est très grand devant l'unité. On peut alors calculer e^z , puis l'énergie totale $U = \int E dn$ en développant les formules suivant les puissances de $\frac{1}{D}$ qui est très petit et en s'arrêtant au premier terme en $\frac{1}{D}$. On trouve alors pour U l'expression

$$(4) \quad U = \frac{3}{40} \left(\frac{6}{\pi} \right)^{\frac{2}{3}} \frac{h^2 n^{\frac{2}{3}}}{m s^{\frac{3}{2}}} + \frac{2^{\frac{1}{2}} \pi^{\frac{8}{3}}}{3^{\frac{2}{3}}} \frac{m s^{\frac{2}{3}} k^2 T^2}{n^{\frac{2}{3}} h^2} + \dots,$$

où $n = \frac{N}{\Omega}$ et où le second terme est de l'ordre de $\left(\frac{1}{D} \right)^{\frac{1}{3}}$. Pour $T = 0$, on obtient pour U la valeur non nulle (énergie au zéro absolu)

$$(5) \quad U_0 = \frac{3}{40} \left(\frac{6}{\pi} \right)^{\frac{2}{3}} \frac{h^2 n^{\frac{2}{3}}}{m s^{\frac{3}{2}}} N.$$

Dans la Mécanique statistique classique où $U = \frac{3}{2} N k T$ et aussi dans celle de Bose-Einstein, U tend vers zéro avec T . Ici, il en va différemment car, le paramètre de dégénérescence de tout gaz de Fermi tendant toujours vers l'infini quand T tend vers zéro, on peut dire que, pour $T = 0$, le gaz de Fermi garde toujours une énergie résiduelle non nulle.

Cette divergence entre les différentes statistiques est aisée à expliquer. En statistique classique et en statistique de Bose-Einstein, rien ne s'oppose à ce que toutes les particules viennent occuper la cellule d'énergie nulle. Au zéro absolu où l'agitation thermique disparaît, les particules obéissant à ces statistiques viendront donc toutes occuper la cellule d'énergie nulle et il n'y aura pas d'énergie résiduelle pour $T = 0$. Mais, en statistique de Fermi-Dirac, il ne peut y avoir plus d'une particule par cellule. Donc, même au zéro absolu, les particules ne peuvent pas toutes venir se placer dans la cellule d'énergie nulle et elles viendront seulement se serrer autour de l'origine dans les cellules les plus voisines de la cellule d'énergie nulle à raison d'une par cellule. Montrons qu'il en résulte pour l'énergie résiduelle U_0 la valeur (5).

Si nous considérons dans l'extension-en-moments (constituée à l'aide des variables rectangulaires p_x, p_y, p_z) une sphère de rayon p entourant

l'origine des coordonnées, le volume de cette sphère sera $\frac{4\pi}{3} p^3$, l'extension-en-phase correspondante est $\frac{4\pi}{3} p^3 \mathcal{V}$ et le nombre des cellules de cette extension, compte tenu du nombre s des valeurs possibles du spin, est $\frac{4\pi s p^3 \mathcal{V}}{3 h^3}$. Au zéro absolu, les particules du gaz de Fermi occuperont toutes les cellules correspondant dans l'espace des moments à l'intérieur d'une sphère de rayon p_M telle que

$$\frac{4\pi s p_M^3 \mathcal{V}}{3 h^3} = N,$$

d'où

$$p_M = \sqrt[3]{\frac{3 N h^3}{4 \pi s \mathcal{V}}}.$$

L'énergie résiduelle du gaz au zéro absolu sera donc (puisqu'il y aura une particule d'énergie $\frac{p^2}{2m}$ par cellule)

$$U_0 = \int_0^{p_M} \frac{4\pi s p^2 dp}{h^3} \frac{p^2}{2m} \mathcal{V}.$$

On trouve tout de suite

$$U_0 = \frac{2\pi s \mathcal{V}}{5 m h^3} p_M^5 = \frac{2\pi s \mathcal{V}}{5 m h^3} \left(\frac{3 N h^3}{4 \pi s \mathcal{V}} \right)^{\frac{5}{3}} = \frac{3}{40} \left(\frac{6}{\pi} \right)^{\frac{2}{3}} \frac{h^2 n^{\frac{2}{3}}}{m s^{\frac{2}{3}}} N.$$

Nous retrouvons bien l'expression (5).

En nous reportant maintenant à la formule (4), nous voyons que, pour une température supérieure au zéro absolu, mais néanmoins assez faible pour que $D \gg 1$, l'énergie augmente parce que l'agitation thermique tend à faire sortir les particules de la sphère de rayon p_M et ceci explique la présence du terme en T^2 dans la formule (4).

La chaleur spécifique à volume constant d'un gaz de Fermi fortement dégénéré s'obtient à partir de (4) en écrivant

$$(6) \quad c = \frac{1}{\mathcal{V}} \frac{\partial U}{\partial T} = \frac{2^{\frac{1}{3}} \pi^{\frac{8}{3}}}{3^{\frac{2}{3}}} \frac{m s^{\frac{2}{3}} k^2 T}{n^{\frac{2}{3}} h^2} n.$$

Comme nous allons le voir, cette formule lève une grande difficulté qui se présentait dans la théorie de la conductibilité des métaux.

7. La théorie électronique des métaux et la statistique de Fermi-Dirac. — Le fait que les métaux soient conducteurs de l'électricité

conduit à penser qu'il existe dans ces corps des électrons libres ou presque libres qui sont susceptibles de prendre sous l'action d'un champ électrique un mouvement d'ensemble uniforme constituant le courant électrique dans le conducteur. Drude, puis Lorentz ont développé une théorie des électrons libres dans les métaux dans laquelle ils considéraient les électrons libres comme formant un gaz soumis aux lois de la Mécanique statistique classique. Cette théorie a fourni une explication, satisfaisante dans l'ensemble, non seulement de la conductibilité électrique, mais de la conductibilité calorifique dans les métaux. En particulier, elle rend compte du fait que le rapport des deux conductibilités est sensiblement le même pour tous les métaux à une température donnée (loi de Wiedemann-Frantz). Néanmoins, le coefficient numérique de cette loi de Wiedemann-Frantz n'a pas la valeur prévue par la théorie rigoureuse de Lorentz.

L'hypothèse qu'il existe dans les métaux des électrons libres, bien que paraissant nécessaire pour expliquer la conductibilité électrique et donnant dans l'ensemble de bons résultats, se heurtait avec les idées classiques à une grave objection. En effet, soit $n^{(a)}$ le nombre des atomes du métal par unité de volume. Ces atomes, étant susceptibles d'osciller autour d'une position d'équilibre, prennent dans l'état d'équilibre thermique à la température absolue T d'après la Mécanique statistique classique une énergie moyenne égale à $3kT$ par atome, soit au total $3n^{(a)}kT$. La chaleur spécifique du métal par unité de volume est donc

$$c^{(a)} = \frac{\partial U^{(a)}}{\partial T} = 3n^{(a)}k.$$

Rapportée à l'atome-gramme, la chaleur spécifique est donc $3\mathcal{N}k = 3\mathcal{R}$, \mathcal{N} étant le nombre d'Avogadro. Si l'on calcule cette chaleur spécifique atomique en calories, on trouve 6 environ : c'est la loi de Dulong et Petit qui est bien vérifiée par les métaux.

Tout semble donc d'abord aller très bien, mais ici se présente la difficulté annoncée. La chaleur spécifique que nous venons de calculer est due uniquement au mouvement d'oscillation des atomes. S'il y a dans le métal des électrons libres, il faut évidemment y ajouter la chaleur spécifique qui provient du mouvement de ces électrons. Si nous désignons par $n^{(e)}$ le nombre des électrons libres par unité de volume, chacun doit prendre en moyenne d'après la Mécanique statistique classique une énergie $\frac{3}{2}kT$. Or le développement de la théorie électronique des métaux avait montré que l'on devait considérer $n^{(e)}$ comme étant de

l'ordre $n^{(a)}$. Alors la chaleur spécifique d'origine électronique sera égale par unité de volume à

$$c^{(e)} = \frac{\partial U^{(e)}}{\partial T} = \frac{3}{2} n^{(e)} k$$

et sera du même ordre que $c^{(a)}$. La chaleur spécifique totale égale à $c^{(a)} + c^{(e)}$ sera supérieure d'environ 50 % à $c^{(a)}$. Ceci est en complète contradiction avec l'expérience puisque la loi de Dulong et Petit est bien vérifiée.

On a pu lever cette grave difficulté à l'aide de la statistique de Fermi-Dirac. En effet, on sait maintenant que les électrons obéissent au principe de Pauli et suivent la statistique de Fermi-Dirac. De plus, si l'on calcule la valeur du paramètre de dégénérescence pour le gaz des électrons libres dans un métal, paramètre dont l'expression s'écrit ici

$$D = \frac{n^{(e)} h^3}{(2\pi m_e kT)^{\frac{3}{2}}},$$

on le trouve de l'ordre de 2 500. Le gaz d'électrons en question est donc fortement dégénéré. La forte valeur trouvée pour D vient de ce que la masse m_e de l'électron est beaucoup plus petite que celle des atomes (1850 fois plus petite que celle de l'atome le plus léger).

Dès lors, s'il est permis pour les atomes d'employer la Mécanique statistique classique et de poser

$$c^{(a)} = 3 n^{(a)} k$$

pour les électrons, il faut utiliser les formules valables en statistique de Fermi-Dirac pour les gaz fortement dégénérés et poser pour la chaleur spécifique par unité de volume

$$c^{(e)} = \frac{2^{\frac{1}{2}} \pi^{\frac{8}{3}}}{3^{\frac{2}{3}}} \frac{m_e (2)^{\frac{2}{3}} k^2 T}{h^2} n^{(e) \frac{1}{3}}$$

en faisant $s = 2$, $n = n^{(e)}$ dans la formule (6).

En comparant les valeurs précédentes de $c^{(a)}$ et de $c^{(e)}$, on voit alors que $c^{(e)}$ est une très petite fraction, de l'ordre du centième, de $c^{(a)}$. Tout se passe donc comme si les électrons libres, tout en assurant la conductibilité électrique et calorifique du métal, ne contribuaient pratiquement pas à sa chaleur spécifique. La raison en est que ces électrons forment un gaz dégénéré et se trouvent, par suite, dans un état très particulier, inconnu des conceptions classiques.

Divers auteurs (Sommerfeld, Léon Brillouin, F. Bloch, Peierls) ont

repris la théorie électronique des métaux sur la base de la Mécanique ondulatoire et de la statistique de Fermi-Dirac. On a pu ainsi retrouver les résultats de Drude et Lorentz en les améliorant sur certains points : ainsi le nouveau coefficient de la loi de Wiedemann-Frantz est en accord avec l'expérience. Toute la théorie des conducteurs et des semi-conducteurs est devenue aujourd'hui un vaste champ d'applications importantes pour la Mécanique ondulatoire et les statistiques quantiques.

BIBLIOGRAPHIE.

- [1] LÉON BRILLOUIN, *Les statistiques quantiques et leurs applications* (2 vol.), Presses universitaires, Paris, 1930.
- [2] FRANCIS PERRIN, *Mécanique statistique quantique*, Gauthier-Villars, Paris, 1939.
- [3] H. A. LORENTZ, *Theory of electrons*, Teubner, Leipzig, 1916, p. 53 et suiv.
- [4] R. PEIERLS, *Electronen theorie der Metalle*, Springer, Berlin, 1932.

FIN.

ERRATUM

—

Page 1, ligne 2, au lieu de John Clerk Maxwell, lire James Clerk Maxwell.

INDEX ALPHABÉTIQUE DES AUTEURS ET DES MATIÈRES

	Pages.
Atome d'hydrogène.....	133, 194
— quantifié (théorie de Bohr).....	110
Balmer (formule de).....	109, 120, 123, 241
Bohr, loi des fréquences.....	101, 111, 136, 191, 199, 203, 205, 215
— postulat des états stationnaires quantifiés.....	111
— théorie de l'atome quantifié.....	110
Boltzmann (constante de).....	63, 64, 71, 90
— (loi de Stefan et).....	80
— (relation de).....	63, 67, 70
Born (principe de).....	171
Chaleurs spécifiques (loi des).....	74
— (théories quantiques des).....	90, 293
Composition des vitesses relativistes.....	32
Compton (effet).....	102
Cone de l'avenir.....	39
Cone de lumière.....	39
Cone du passé.....	39
Constante de structure fine.....	119, 256
Cordes vibrantes.....	184
Correspondance (principe de).....	127, 135, 203, 214
Densité spatiale d'énergie.....	4
Dirac (théorie de).....	250
— (fonction δ de).....	228
— (matrices de).....	252
Doppler.....	82
Dulong et Petit (loi de).....	75, 91
Dynamique relativiste du point matériel.....	44
Dynamique relativiste du corpuscule électrisé.....	54
Effet photoélectrique.....	95
Ehrenfest (théorème d').....	168
Einstein (relations d').....	97, 99, 146, 156
Energie rayonnée.....	203, 214
Equations canoniques.....	205
Equations de Dirac.....	252, 254
Equipartition de l'énergie (théorème de l').....	72
Espace-temps.....	37
Etats (somme d').....	66

	Pages.
Extension en phase.....	60, 69, 279
Fermi (gaz fortement dégénéré de).....	200
Fizeau (expérience de).....	33
— (formule de).....	34
Flux d'énergie électromagnétique.....	4, 8
Fonction propre (définition).....	187
Galilée (transformation de).....	19
Gaz parfait.....	68
Gibbs (loi de répartition canonique de).....	65, 88
Hamilton (équations de).....	130
— (principe de).....	48
Hamiltonien (opérateur).....	208, 225
Heisenberg (matrices de).....	211
— (mécanique quantique de).....	201
— (microscope de).....	175
— (relations d'incertitude de).....	172, 230
Hermite (polynomes de).....	193
Hypothèse ergodique.....	63
Hypothèse d'Uhlenbeck et Goudsmit.....	244
Impulsion d'Univers.....	51, 56, 148
Intégrales premières en Mécanique ondulatoire.....	232
Jacobi (équations de).....	150
— (fonction de).....	150
— (fonction raccourcie de).....	150, 181
Jeans (loi de distribution spectrale de Rayleigh et).....	84, 281
Kronecker (symboles de).....	187
Lagrange (fonction de).....	45, 46, 54
Landé (formule de).....	257
Laplace (fonctions sphériques de).....	191, 195
Larmor (théorème de).....	134, 142, 197
Legendre (polynomes de).....	191
Ligne d'Univers.....	40, 51
Liouville (théorème de).....	60, 67
Lorentz (contraction de).....	24
— (équations de).....	6, 53
— (force de).....	7, 55
— (relation de).....	9
— (théorie des électrons de).....	6
— (transformation de).....	19, 21
Lorentz-Lorenz (formule de).....	16
Magnéton de Bohr.....	242, 245, 255
Masse longitudinale.....	56
Masse propre (ou au repos).....	48
Masse transversale.....	56
Matrices (intervention en Mécanique quantique des).....	203, 232
Matrice diagonale.....	207
Matrice hermitique (ou hermitienne).....	203
Maxwell (équations de).....	2, 52
— (loi de distribution des énergies de).....	67, 70, 86
— (relation de).....	5, 15
— (théorie de).....	1
Mécanique ondulatoire (formule fondamentale de la).....	157, 159

	Pages.
Métaux (théorie électronique des).....	291
Michelson.....	19, 20
Moment de rotation propre.....	255
Moment dipolaire.....	11, 12, 136, 142, 202, 215
Moment quadrupolaire.....	11
Opérateur complet.....	224
Opérateur hermitique.....	209
Opérateur incomplet.....	224
Optique géométrique.....	153, 158, 181
Orthogonalité des fonctions propres.....	187, 209
Orthohélium.....	272
Orthohydrogène.....	275
Oscillateur harmonique.....	192
Paradoxe de réciprocité.....	30
Parahydrogène.....	275
Paramètre de dégénérescence.....	285, 281, 293
Parhélium.....	272
Photons (hypothèse des).....	99, 104
Planck (constante de).....	90, 111
— (loi de répartition spectrale de).....	87
Polarisabilité.....	15
Polarisation (phénomènes de).....	13
— (tenseur).....	52
Potentiels avancés.....	10
Potentiel d'Univers.....	53, 54
Potentiel scalaire.....	8
Potentiels électromagnétiques.....	8
Potentiels retardés.....	8, 10
Potentiel vecteur.....	8
Poynting (vecteur radiant de).....	4, 8, 13
Principe de combinaison (Ritz).....	110, 111
Principe de décomposition spectrale.....	169, 226, 228
Principe de Fermat.....	154
Principe de Maupertuis.....	151
Principe de quantification.....	225
Principe des interférences (ou de localisation).....	165, 229
Principe d'exclusion de Pauli.....	261, 270
Probabilité de transition.....	139
Quantification dans l'espace.....	233, 243
Quantification en Mécanique Ondulatoire.....	183
Ralentissement des horloges.....	26
Raman (effet).....	101
Rayleigh (loi de distribution spectrale de Jeans et).....	84, 281
— (théorème de).....	160, 280
Rayonnement noir.....	77, 287
Règles de sélection.....	141, 216, 243
Relations de commutation.....	205
Rotateur plan.....	188
Rotateur sphérique.....	190
Rydberg (constante de).....	109, 113, 120, 256
Schrödinger (équations de).....	158, 160, 163, 183, 208

	Pages.
Sommerfeld (conditions de quantification de).....	116, 117, 129, 182, 280
— (formules de).....	119, 123, 256
— (théorie de).....	116
Spectre continu.....	187
Spectre de l'hélium.....	271
Spectre de raies.....	187
Spin.....	124, 245, 270
Spin de l'électron.....	245, 248
Statistiques de Bose-Einstein.....	283
Statistiques de Fermi-Dirac.....	271, 287, 291
Stirling (formule de).....	69, 283
Stoner (règle de).....	260
Susceptibilité.....	15
Systèmes (Mécanique Ondulatoire des).....	263
Système complet.....	188, 225
Tenseur d'espace-temps.....	43, 44
Valeur moyenne.....	214, 230
Valeur propre (définition).....	187
Valeur propre dégénérée.....	210
Variables angulaires.....	128, 132, 142
Variables d'action.....	132, 142
Vecteur d'espace-temps.....	42, 43
Vecteur du genre espace.....	39
Vecteur du genre temps.....	39
Vitesse de groupe.....	160
Wien (loi du déplacement de).....	81
— (loi empirique de distribution de).....	86
Zeeman (effet).....	18, 134, 143, 197, 242

TABLE DES MATIÈRES.

	Pages.
PRÉFACE.....	V
PROGRAMME DU CERTIFICAT D'ÉTUDES SUPÉRIEURES DE THÉORIES PHYSIQUES.....	VII

CHAPITRE I.

RÉSUMÉ DE LA THÉORIE DE MAXWELL ET DE LA THÉORIE DES ÉLECTRONS.

1. Grandes lignes de la Théorie de Maxwell ..	1
2. Théorie électromagnétique de la Lumière	4
3. Théorie des Électrons (H. A. Lorentz).....	6
4. Les potentiels électromagnétiques. Formule des potentiels retardés.....	8
5. Calcul du potentiel créé par une distribution statique.....	10
6. Distribution non statique. Rayonnement.....	12
7. Phénomènes de polarisation. Théorie de la dispersion.....	13
8. Succès et échecs de la théorie de Lorentz.....	18

CHAPITRE II.

LE PRINCIPE DE RELATIVITÉ.

1. La transformation de Lorentz.....	19
2. Signification physique de la transformation de Lorentz.....	22
3. La contraction de Lorentz.....	24
4. Le ralentissement des horloges.....	26
5. Mesure de la vitesse de la lumière par deux observateurs galiléens.....	29
6. Paradoxe de la réciprocité.....	30
7. Formule de composition des vitesses.....	32
8. Interprétation de l'expérience de Fizeau par la cinématique relativiste.....	33

CHAPITRE III.

COMPLÉMENTS SUR LA THÉORIE DE LA RELATIVITÉ RESTREINTE.

1. L'Espace-temps.....	37
2. Démonstration géométrique dans l'espace-temps des relations de Lorentz.....	40
3. Vecteurs et tenseurs d'espace-temps.....	42
4. La Dynamique relativiste du point matériel.....	44
5. L'Électromagnétisme relativiste.....	52
6. Dynamique relativiste du corpuscule électrisé dans un champ électromagnétique.....	54
7. Masse transversale et masse longitudinale.....	56
8. Vérifications expérimentales de la théorie de la Relativité restreinte.....	58

CHAPITRE IV.

LA MÉCANIQUE STATISTIQUE CLASSIQUE.

	Pages.
1. But de la Mécanique statistique classique.....	59
2. Extension-en-phase. Théorème de Liouville.....	60
3. Entropie et probabilité. Relation de Boltzmann.....	63
4. La loi de répartition canonique de Gibbs.....	65
5. Application au cas d'un gaz parfait.....	68
6. Théorème de l'équipartition de l'énergie.....	72
7. Applications du théorème de l'équipartition de l'énergie.....	74

CHAPITRE V.

LA THÉORIE DU RAYONNEMENT NOIR.

1. Généralités sur le problème du Rayonnement noir.....	77
2. Loi de Stefan-Boltzmann.....	80
3. Loi du déplacement de Wien.....	81
4. La loi de distribution spectrale de Rayleigh-Jeans.....	84
5. Échec de la loi de Rayleigh.....	85
6. La loi de répartition spectrale de Planck.....	87
7. La théorie quantique des chaleurs spécifiques (Einstein, Debye).....	90

CHAPITRE VI.

LA STRUCTURE CORPUSCULAIRE DE LA LUMIÈRE. LES PHOTONS.

1. Aperçu historique sur les théories de la Lumière.....	93
2. L'effet photoélectrique.....	95
3. L'hypothèse des photons et ses succès.....	99
4. L'Effet Compton.....	102
5. Difficultés soulevées par la théorie des photons.....	104
6. Caractère probabiliste de la synthèse onde-corpuscule.....	106

CHAPITRE VII.

LA THÉORIE QUANTIQUE DE L'ATOME DE BOHR-SOMMERFELD.

1. La formule de Balmer, les termes spectraux et le principe de combinaison....	109
2. La théorie de l'atome quantifié de Bohr.....	110
3. Succès et insuffisances de la théorie de Bohr.....	113
4. La théorie de Sommerfeld.....	116
5. Succès et insuffisances de la théorie de Sommerfeld.....	121

CHAPITRE VIII.

LE PRINCIPE DE CORRESPONDANCE.

1. Objet du principe de correspondance.....	127
2. Les variables angulaires.....	128
3. Exemples de quantification par la méthode des variables angulaires.....	133

	Pages.
4. Le principe de correspondance.....	135
5. Démonstration des règles de sélection par le principe de correspondance.....	141

CHAPITRE IX.

IDÉES DE BASE ET ÉQUATIONS FONDAMENTALES DE LA MÉCANIQUE ONDULATOIRE.

1. Point de départ.....	145
2. La théorie de Jacobi en Mécanique analytique.....	148
3. Rappel de notions fondamentales de la théorie des ondes.....	152
4. Nouvelle manière d'accéder à la Mécanique ondulatoire.....	155
5. Remarques importantes et généralisation au cas des champs non permanents..	158
6. Le théorème de la vitesse de groupe.....	160

CHAPITRE X.

LA SIGNIFICATION PHYSIQUE DE LA MÉCANIQUE ONDULATOIRE.

1. Remarques générales.....	163
2. Principe de localisation ou principe des interférences.....	165
3. Théorème d'Ehresfest.....	168
4. Principe de décomposition spectrale ou principe de Born.....	169
5. Les relations d'incertitude d'Heisenberg.....	172
6. Évolution de l'onde Ψ . Importance de la mesure.....	173
7. Le microscope d'Heisenberg.....	175
8. Passage d'un corpuscule à travers une fente percée dans un écran.....	177
9. Le raccord avec la Mécanique classique.....	179

CHAPITRE XI.

APPLICATIONS DE LA MÉCANIQUE ONDULATOIRE A LA QUANTIFICATION.

1. Première interprétation de la quantification par la Mécanique ondulatoire.....	181
2. La théorie de la quantification en Mécanique ondulatoire.....	183
3. Définition générale et propriétés des valeurs et fonctions propres.....	186
4. Le rotateur plan.....	188
5. Le rotateur sphérique.....	190
6. L'oscillateur harmonique.....	192
7. L'atome d'hydrogène.....	194
8. L'Effet Zeeman en Mécanique ondulatoire.....	197
9. Les méthodes de perturbation.....	199

CHAPITRE XII.

MÉCANIQUE QUANTIQUE D'HEISENBERG ET PRINCIPE DE CORRESPONDANCE.

1. Point de départ de la Mécanique quantique d'Heisenberg.....	201
2. Équations canoniques et relations de commutation.....	205
3. Retour sur certains points de la Mécanique ondulatoire.....	207
4. Interprétation des matrices d'Heisenberg en Mécanique ondulatoire.....	211
5. Le principe de correspondance dans la nouvelle Mécanique.....	214
6. Démonstration des règles de sélection.....	216
7. La formule de Kramers-Heisenberg.....	217

CHAPITRE XIII.

L'INTERPRÉTATION PROBABILISTE DE LA MÉCANIQUE ONDULATOIRE.

	Pages.
1. Idées générales.....	219
2. Grandeurs et opérateurs en Mécanique ondulatoire.....	223
3. Principes généraux de l'interprétation probabiliste de la Mécanique ondulatoire.....	225
4. Application des principes généraux.....	227
5. Mesure simultanée de deux grandeurs.....	229
6. Valeur moyenne d'une grandeur en Mécanique ondulatoire.....	230
7. Les intégrales premières en Mécanique ondulatoire.....	232
8. La quantification dans l'espace.....	233

CHAPITRE XIV.

LE SPIN DE L'ÉLECTRON. LA THÉORIE DE DIRAC.

1. Nécessité d'introduire un élément supplémentaire dans la définition de l'électron.....	239
2. L'hypothèse d'Uhlenbeck et Goudsmit (1925). Le spin de l'électron.....	244
3. Le spin de l'électron en Mécanique ondulatoire.....	248
4. La théorie de Dirac.....	250
5. Caractéristiques diverses de la théorie de Dirac.....	253
6. Succès et insuffisances de la théorie de Dirac.....	255

CHAPITRE XV.

LE PRINCIPLE DE PAULI ET LA MÉCANIQUE ONDULATOIRE DES SYSTÈMES DE CORPUSCULES.

1. La répartition des électrons dans l'atome.....	259
2. Le principe d'exclusion de Pauli.....	261
3. La Mécanique ondulatoire des systèmes de corpuscules.....	263
4. Cas des systèmes formés de corpuscules de même nature.....	267
5. Introduction du spin en Mécanique ondulatoire des systèmes. Énoncé rigoureux du principe de Pauli.....	270
6. Théorie du spectre de l'hélium (Heisenberg).....	271
7. Orthohydrogène et parahydrogène.....	275
8. Applications diverses.....	277

CHAPITRE XVI.

LES STATISTIQUES QUANTIQUES.

1. Les cellules d'extension-en-phase dans la théorie des Quanta.....	279
2. Nouvelles méthodes statistiques.....	282
3. La statistique de Bose-Einstein.....	283
4. Application de la statistique de Bose-Einstein au Rayonnement noir.....	286
5. La statistique de Fermi-Dirac.....	287
6. Le gaz de Fermi dans un état fortement dégénéré.....	290
7. La théorie électronique des métaux et la statistique de Fermi-Dirac.....	291
INDEX ALPHABÉTIQUE DES AUTEURS ET DES MATIÈRES.....	295
TABLE DES MATIÈRES.....	299