

ANNALES DE L'I. H. P.

W. HEITLER

La théorie quantique des forces de valence

Annales de l'I. H. P., tome 4, n° 2 (1933), p. 237-272

[<http://www.numdam.org/item?id=AIHP_1933__4_2_237_0>](http://www.numdam.org/item?id=AIHP_1933__4_2_237_0)

© Gauthier-Villars, 1933, tous droits réservés.

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

NUMDAM

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

La théorie quantique des forces de valence

PAR

W. HEITLER

I

Introduction

Le sujet de nos conférences sera la théorie de la *liaison chimique homéopolaire*. Ce problème était insoluble en physique classique, pour deux raisons. D'abord, on ne pouvait pas trouver des *forces* capables d'engendrer l'attraction de *deux atomes* neutres, tels que deux atomes d'hydrogène par exemple. Nous ne connaissons en physique classique que les forces de gravitation et les forces électriques et magnétiques. Les premières, ainsi que les forces magnétiques sont beaucoup trop petites pour expliquer les attractions que nous révèle la chimie. Considérons donc les forces électriques s'exerçant entre deux atomes. Si ces derniers sont neutres, chacun étant composé de différentes charges, par exemple d'un noyau positif et d'électrons chargés négativement, on peut développer l'énergie potentielle mutuelle des deux atomes en série de la forme suivante :

$$(1) \quad \Phi(R) = \frac{(\vec{\mu}_1, \vec{\mu}_2)}{R^3} + \dots$$

L'énergie de COULOMB $\frac{e_1 e_2}{R}$ est nulle, parce que les deux atomes sont neutres ; μ^1 est le moment *dipolaire* de l'atome, qui s'écrit

$$\vec{\mu} = \sum \vec{r}_i e_i.$$

où e_i représentent les différentes charges de l'atome et $\vec{r}_i$ leurs distances vectorielles au centre.

Le premier membre de cette série est encore très petit, égal à zéro en première approximation. En effet, si l'on suppose en première approximation que le mouvement des électrons du premier atome n'est pas changé par la présence du deuxième atome, μ_1 est un vecteur qui peut prendre toutes les directions de l'espace avec une égale probabilité. On comprend alors que la valeur moyenne de $(\vec{\mu}_1, \vec{\mu}_2)$ soit nulle. C'est donc seulement le changement du mouvement des électrons du premier atome causé par la présence du deuxième, qui rend Φ différent de zéro. On appelle ce phénomène la *polarisation*. Les termes suivants de la série sont encore plus petits. Seule la polarisation produit une certaine interaction entre les deux atomes. Des calculs de ce genre ont été effectués par DEBYE. Il trouva, que les deux atomes s'attirent l'un l'autre par suite de la polarisation, mais que cette attraction est encore trop petite, inférieure à l'énergie de dissociation de la molécule H_2 , par exemple.

Ce n'est là, qu'une première difficulté. Les forces de la chimie sont caractérisées par la notion de valence. Considérons de nouveau deux atomes d'hydrogène. Nous savons, que les forces chimiques produisent une attraction ; mais nous savons aussi, qu'un troisième atome d'hydrogène ne sera plus attiré par l'ensemble des deux autres. L'atome d'hydrogène possède une valence égale à un : il ne peut attirer qu'un seul atome d'hydrogène, tout autre atome étant au contraire repoussé. Si l'on admettait pour un instant, que l'attraction des deux atomes était due à la polarisation, ou à une autre force classique du même genre, il serait impossible de comprendre pourquoi un troisième atome ne serait pas également attiré. Les forces de la physique classique sont soumises à la loi de l'*addition des vecteurs*, c'est-à-dire qu'une force qui agit entre un premier et un second atome doit également agir entre le second et un troisième. Aucune force de la physique classique ne possède cette propriété de *saturation*, qui est si caractéristique de toute la chimie.

C'est seulement la *théorie des quanta* qui permet de résoudre ce problème. Nous allons voir comment la mécanique quantique nous fournit les forces désirées. Quelles sont donc les différences entre la mécanique quantique d'une part et la mécanique et l'électrodynamique classique

d'autre part ? Elles ne consistent pas dans le fait que les forces entre particules élémentaires sont changées. Entre particules élémentaires, c'est-à-dire entre les électrons et les noyaux, agissent en théorie quantique les mêmes forces qu'en théorie classique, c'est-à-dire les forces de COULOMB. Mais c'est la *cinématique* des particules, qui est essentiellement modifiée.

Un électron ne décrit plus — comme en mécanique classique — une trajectoire définie dans l'espace ; son mouvement est décrit par une certaine fonction d'onde Ψ déterminée par l'équation bien connue de SCHRÖDINGER

$$\Delta \Psi_n + \frac{8\pi^2 m}{h^2} (E_n - V) \Psi_n = 0,$$

qui conduit en général aux états quantiques discrets E_n . Cette fonction Ψ_n nous donne seulement la probabilité de présence de l'électron au point x, y, z . Elle est l'expression du fait, que l'électron possède une certaine nature ondulatoire, abstraction faite de sa nature de particule. L'équation de SCHRÖDINGER détermine cette probabilité, lorsque l'électron se meut sous l'influence d'une force, dérivant d'un potentiel V , dont l'expression est la même qu'en théorie classique.

Mais cette différence n'est pas la seule entre la cinématique classique et la cinématique quantique. D'après la cinématique nouvelle, un électron lié à un atome semble posséder un certain moment de quantité de mouvement.

$$(2) \quad s = \frac{1}{2} \frac{h}{2\pi}$$

indépendamment du moment dû à la rotation de l'électron autour du noyau. Nous appelons ce moment le *spin*. Le spin a la propriété suivante : il est impossible d'attribuer en même temps des valeurs définies à ses trois composantes s_x, s_y, s_z ; on ne peut le faire que pour l'une d'entre elles, s_x par exemple, tandis que pour les deux autres on ne peut donner que des probabilités relatives aux différentes valeurs possibles. Ces dernières sont seulement au nombre de deux, à savoir $+\frac{1}{2}$ ou $-\frac{1}{2}$. Nous pouvons désigner un état, où l'électron possède une valeur $s_z = +\frac{1}{2}$ par une flèche dirigée vers le haut $\uparrow$, et

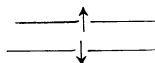
un état, où l'électron a la valeur $s_z = -\frac{1}{2}$, par une flèche dirigée vers le bas $\downarrow$.

Enfin, nous trouvons entre les mécaniques quantique et classique des différences fondamentales si nous considérons le cas de plusieurs électrons. Même si nous négligeons les forces d'interaction entre électrons, c'est-à-dire si nous négligeons le potentiel de COULOMB $\frac{e^2}{r_{ik}}$ où, r_{ik} représente la distance entre deux électrons i et k , il subsistera toujours une très grande *influence* d'un électron donné sur les autres. Considérons le cas de deux électrons.

Si l'on néglige l'interaction de COULOMB $\frac{e^2}{r_{12}}$, chaque électron se trouve en un certain état Ψ_n qui est un de ceux, dans lequel un électron pourrait se trouver, si l'autre n'existait pas. Les deux électrons se trouvent donc dans deux états

$$\Psi_a, \quad \Psi_b,$$

avec les énergies E_a et E_b . Pour décrire complètement l'état du système, il est nécessaire de se donner aussi l'état de spin de chaque électron, c'est-à-dire, la direction de sa composante s_z . Un état du système est donc complètement déterminé, si nous connaissons Ψ_a, Ψ_b



et les deux directions de spin correspondantes.

Naturellement, il est possible, que les deux électrons se trouvent dans le même état : $a = b$. Mais alors, une loi fondamentale de la cinématique quantique exige que : Si les deux électrons se trouvent dans le même état $a = b$, il est nécessaire que leurs directions du spin soient opposées. Le seul état qui puisse se rencontrer dans la nature est représenté par le symbole $\uparrow\downarrow$. Cette loi est connue sous le nom de *principe d'exclusion de Pauli*.

Jusqu'ici nous avons négligé les forces d'interaction entre les deux électrons. Si nous introduisons maintenant ces forces, nous ne pourrions plus parler qu'avec une certaine approximation de l'état de chaque électron. Mais si l'on étend à ce cas le principe de PAULI, il est légitime d'admettre l'apparition d'influences très grandes entre les

spins de deux électrons, même quand ces derniers se trouvent dans deux états différents. Une conséquence du principe d'exclusion, généralisé au cas de l'interaction des deux électrons est qu'il existe un *couplage* très fort entre les deux spins, qui sont forcés de se placer soit tous les deux dans la même direction ou soit dans deux directions opposées.

$$\uparrow\uparrow \quad \text{ou} \quad \downarrow\downarrow.$$

A ce fort couplage correspond un grand changement de l'énergie du système. Nous pouvons définir l'énergie de l'interaction en formant la différence

$$\varepsilon = E - (E_a + E_b)$$

où E représente l'énergie exacte du système, $E_a + E_b$ représentent l'énergie que le système aurait sans interaction entre les électrons. Cette énergie d'interaction n'aura pas la même valeur suivant que les spins seront parallèles ou antiparallèles.

Au lieu d'une seule énergie $E_a + E_b$ sans interaction des deux électrons, on obtient maintenant deux énergies différentes

$$E_a + E_b \longrightarrow \begin{cases} \xrightarrow{\varepsilon_p} \uparrow\uparrow \\ \xrightarrow{\varepsilon_n} \downarrow\downarrow \end{cases}$$

pour les deux configurations des spins : $E_a + E_b + \varepsilon_p$ et $E_a + E_b + \varepsilon_n$. Si le couplage entre les deux spins n'existait pas, l'énergie de l'interaction ε serait simplement — d'après la théorie des perturbations — la moyenne du potentiel de COULOMB $\frac{e_2}{r_{12}}$. Il s'y ajoute maintenant l'énergie de couplage entre les deux spins, que nous appelons A . Un calcul détaillé nous donne alors

$$\begin{aligned} \varepsilon_p &= C + A \\ \varepsilon_n &= C - A. \end{aligned}$$

Cette énergie du couplage A est du même ordre de grandeur, que C . C et A dépendent naturellement de la structure du champ dans lequel se meuvent les deux électrons, c'est-à-dire de la structure de l'atome.

Le fait que nous venons d'exposer est le point de départ de la théorie de la liaison homopolaire. J'essaierai tout d'abord d'esquisser les points de vue généraux et les résultats principaux de la théorie, sans

calculs détaillés, et d'une manière intuitive. Je développerai ensuite la théorie exacte dans les chapitres suivants.

Considérons le cas le plus simple de deux atomes, ayant chacun un électron. Soit R la distance entre les deux noyaux. L'énergie du système sans interaction des électrons sera la somme de l'énergie du premier atome et du second atome, et l'énergie d'interaction s'écrira : $\varepsilon = C \pm A$. Ces deux quantités C et A dépendent de la structure du système, c'est-à-dire dans notre cas de la distance R des noyaux. Si, par exemple, cette distance est *infinie*, C et A doivent être nuls parce que il ne subsiste alors aucune interaction entre les deux atomes.

On peut alors interpréter $\varepsilon(R)$ comme l'énergie potentielle de l'interaction des deux atomes. La force, qui agit entre eux sera $-\frac{\partial \varepsilon(R)}{\partial R}$.

Nous avons vu que l'expression de l'interaction des deux atomes en mécanique quantique est :

$$(3) \quad \varepsilon = C \pm A$$

On a calculé ces deux quantités en fonction de R pour quelques atomes simples et l'on a trouvé que le couplage entre les spins, A , qui est même plus grand que C , est du même ordre de grandeur que l'énergie de dissociation de la molécule. Il s'ensuit que c'est seulement pour *une seule* configuration des spins, que l'interaction des deux atomes est telle que les deux atomes s'attirent. Si par exemple le signe de A est négatif (et le calcul montre qu'il en est vraiment ainsi), l'énergie d'interaction est négative lorsque les deux spins sont antiparallèles $\uparrow\downarrow$.

L'énergie diminue alors en même temps que la distance à partir de $R = \infty$, c'est-à-dire, que les deux atomes s'attirent. Au contraire, si les deux spins sont parallèles, les deux atomes se repoussent. Si l'on trace les courbes de C , A , en fonction de R , on obtient pour $\varepsilon(R) = C \pm A$ les deux courbes de la figure 1.

On voit que pour l'une de ces courbes, ε passe par un minimum : les deux atomes peuvent former une molécule ; si l'on calcule les constantes de cette molécule, par exemple la distance d'équilibre et l'énergie de dissociation, on obtient des valeurs voisines des valeurs expérimentales. La courbe pointillée correspond à l'expérience.

Si nous considérons maintenant l'interaction de deux atomes ayant plusieurs électrons, nous n'avons qu'à généraliser les résultats obtenus plus haut.

LA THÉORIE QUANTIQUE DES FORCES DE VALENCE

Soient donc deux atomes A et B à plusieurs électrons. Supposons que le premier se trouve dans un état dans lequel n_a électrons seulement aient leurs spins parallèles, c'est-à-dire un état dans lequel le moment de spin total soit $s_a = \frac{n_a}{2}$.

Le second atome B aura un spin $s_b = \frac{n_b}{2}$. Supposons d'autre part s_a plus grand que s_b . Dans ce cas, si les deux atomes commencent à réagir l'un sur l'autre, les deux spins s_a et s_b s'accouplent. Le modèle de l'addition des vecteurs, qui est une conséquence générale de la mécanique quantique, nous apprend comment ce couplage a lieu. Il

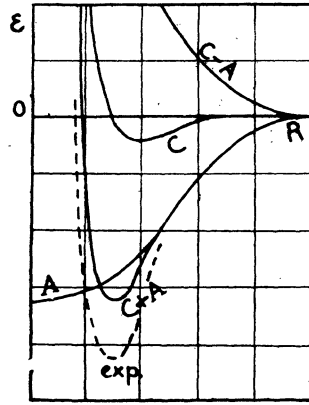


FIG. 1.

est possible que le vecteur s_a prenne la même direction que s_b , de sorte que le spin résultant du système soit $s = s_a + s_b$. Mais le spin s_b peut également prendre toutes les directions permises, telles que le spin résultant s possède des valeurs inférieures, différant d'un nombre entier de $s_a + s_b$, et cela jusqu'à ce que s_b ait une direction opposée à s_a . Les valeurs possibles de s sont alors

$$(4) \quad s = s_a + s_b, \quad s_a + s_b - 1 \dots s_a - s_b.$$

On peut exprimer ce fait d'une autre manière : la valeur la plus grande de s est $s_a + s_b$: tous les électrons, les n_a électrons de a et les n_b de b , ont leur spin dirigé dans la même direction. Si donc le spin résultant s est inférieur d'une unité, $s = s_a + s_b - 1$, on peut dire

que deux électrons, l'un de l'atome a et l'autre de l'atome b ont opposé leurs spins, $\uparrow\downarrow$, de sorte qu'ils ne contribuent plus en rien au spin résultant. En général, on peut définir un nombre p de paires d'électrons à spins antiparallèles, tel que le spin résultant soit

$$s = s_a + s_b - p.$$

p peut prendre les valeurs

$$p = 0, 1, \dots, n_b.$$

Dorénavant, c'est toujours du nombre p que nous parlerons, p représentant le nombre de paires d'électrons.

Dans chacun de ces différents cas nous aurons une énergie différente pour l'interaction de deux atomes. Le *couplage* entre les spins des atomes sera différent pour les différentes valeurs de p . La théorie conduit à un résultat général :

$$(5) \quad \varepsilon_p = C + [p - (n_a - p)(n_b - p)] A,$$

où C et A sont deux fonctions qui dépendent seulement de la distance R et de la structure des atomes. Dans la plupart des cas leur dépendance de R est la même que dans le cas de deux atomes ayant un seul électron chacun. En particulier, le signe de A est négatif. On voit alors que la liaison entre les deux atomes est maximum si p est aussi *grand* que possible. L'énergie d'interaction est dans ce cas

$$\varepsilon_p = C + n_b A;$$

si p décroît, la liaison devient plus lâche, et enfin, si $p = 0$ nous avons

$$\varepsilon = C - n_a p_b A,$$

c'est-à-dire les deux atomes se repoussent.

Le résultat de cette théorie est parfaitement d'accord avec les idées que G. N. LEWIS a développées sur la liaison chimique. LEWIS, en discutant le matériel empirique de la chimie, eut l'idée, que chaque liaison homopolaire chimique est produite par une paire d'électrons, un électron de chaque atome. Nous voyons maintenant, que les paires de LEWIS ne sont autre chose que les paires d'électrons à spins opposés de la mécanique quantique. Notre nombre p de paires d'électrons est donc le nombre des traits de valence qui joignent

les deux atomes. $n_a - p$ est égal au nombre des traits libres, qui viennent de l'atome a ; le nombre n_a d'électrons à spins parallèles est la valence de l'atome a .

Si l'un des deux atomes possède un spin $s_a = 0$, tous ses électrons ont saturé leurs spins. Une liaison chimique avec un tel atome n'est pas possible. La théorie conduit au résultat suivant :

$$\varepsilon = C - n_b A$$

L'énergie de couplage des spins indique une répulsion, en accord avec les faits connus.

L'idée de la valence a trouvé ses applications les plus fécondes dans le cas de plusieurs atomes, surtout en chimie organique. Il nous faut donc étendre notre théorie à ce cas.

Considérons une série d'atomes $a, b, c...$ de valences respectives $n_a, n_b, n_c...$ Le problème de la liaison de ces atomes sera résolu si nous parvenons à calculer l'énergie d'interaction de ces atomes, qui sera maintenant une fonction de toutes les distances R_{ab} (des atomes a et b), $R_{ac}...$ etc. Si nous connaissons cette fonction, nous connaissons les configurations d'équilibre possibles du système considéré et nous saurons quels sont les atomes qui peuvent former une molécule.

D'après les résultats que nous avons obtenu dans la théorie de deux atomes, nous pouvons prévoir à peu près le fait suivant : deux atomes A et B s'attireront s'il est possible de former une paire d'électrons avec spins anti-parallèles entre eux.

Nous développerons la théorie pour ce cas dans les chapitres suivants, et nous verrons qu'elle est en général d'accord avec cette idée fondamentale. Mais d'autre part, les circonstances énergétiques sont très compliquées et nous verrons que la théorie permet d'interpréter beaucoup de points de détail, par exemple les phénomènes d'activation, qui ne peuvent pas être compris avec la simple idée des traits de valence ou des paires d'électrons.

II

Les invariants de spin et les états purs de valence

1. Pour expliquer les phénomènes de la liaison homopolaire, il est nécessaire de calculer l'énergie d'interaction de plusieurs atomes. Nous allons développer maintenant la théorie générale de cette interaction.

Supposons que nous ayons une série d'atomes A, B, C, etc. et que chacun d'eux se trouve dans un certain état quantique, par exemple dans l'état *fondamental*, défini par sa fonction d'onde. Soit Ψ_a la fonction d'onde de l'atome A, etc. Lorsqu'on calcule l'interaction des atomes, il n'est pas permis en général de prendre chaque atome dans un seul état *défini*. Nous savons, que, dans le problème de l'interaction, *tous* les états quantiques de chaque atome interviennent. Cependant, si l'énergie d'interaction est petite par rapport des différences d'énergie des divers états, c'est-à-dire par rapport à l'énergie d'excitation de chaque atome, il suffira, pour calculer l'interaction, de ne considérer que l'état fondamental de chaque atome. Dans cette hypothèse, il ne faut néanmoins pas perdre de vue les autres états voisins de l'état considéré, s'il y en a. Par exemple, dans la théorie des valences dirigées de SLATER et PAULING la considération de plusieurs états est essentielle.

Soit donc Ψ_a la fonction d'onde de l'atome A.

Pour simplifier le problème autant que possible, nous ne considérons que les électrons de A dont les spins ne sont pas saturés ; en effet, nous avons vu que les électrons qui ont saturé mutuellement leurs spins à l'intérieur d'un atome, ne peuvent contribuer à aucune liaison avec un autre atome. Dans ce cas on peut toujours écrire cette fonction comme le produit d'une fonction d'onde u_a qui ne dépend que des coordonnées des électrons de A et d'une fonction de spin φ_a

$$(6) \quad \Psi_a = u_a \varphi_a.$$

La fonction Ψ_a ne dépend donc que des électrons ayant leurs spins parallèles. Les électrons des *couches internes* ne jouent qu'un rôle né-

gligeable pour toute l'énergie d'interaction. De plus, nous supposons que l'état de l'atome u_a n'est pas dégénéré, c'est-à-dire qu'il s'agit d'un état S (abstraction faite de la dégénérescence de spin).

Soit n_a le nombre des électrons ayant leurs spins parallèles dans l'atome A. u_a sera alors une fonction des coordonnées de ces n_a électrons, φ_a une fonction des n_a variables de spin. Le spin résultant de A a pour valeur $s_a = \frac{n_a}{2}$; il en résulte, comme nous allons le démontrer, que φ_a est une fonction symétrique des n_a variables de spin.

En effet, les composants de s_a dans une direction déterminée (axe des z) peuvent prendre les valeurs

$$(6) \quad \sigma_a = s_a, s_{a-1}, \dots, -s_a$$

c'est-à-dire l'état de spin s_a est toujours dégénéré et tel qu'il y a $2s_a + 1$ différentes fonctions de spin qui correspondent aux différentes composantes σ_a .

Si nous désignons par $\alpha(1)$ ou par $\beta(1)$ la fonction de spin du premier électron suivant que ce spin est dirigé vers le haut ou vers le bas, c'est-à-dire suivant que la composante $\sigma(1) = +\frac{1}{2}$ ou $-\frac{1}{2}$, il n'y aura qu'une seule fonction qui, pour les n_a électrons, donnera une composante $\sigma_a = s_a$ à savoir

$$(7) \quad \varphi_{na} = \alpha(1)\alpha(2) \dots \alpha(n_a).$$

Elle est symétrique. Faisons tourner maintenant les axes de coordonnées : φ_{na} se transforme en une autre fonction φ_{ra} , avec une autre composante σ_a . Mais la symétrie de la fonction ne peut pas être changée. Il en résulte, que toutes les autres fonctions φ_{ra} restent symétriques. Alors il est facile d'écrire toutes les fonctions de spin : ce sont les fonctions symétriques, construites à l'aide de α et β

$$(8) \quad \varphi_{ra} = \text{sym } \alpha(1) \dots \alpha(n_a - r_a) \beta(n_a - r_a + 1) \dots \beta(n_a)$$

où r_a représente le nombre des électrons dont la fonction de spin est β , c'est-à-dire dont la composante de spin est $-\frac{1}{2}$. Le signe « sym » signifie, que l'on doit rendre symétrique la fonction (8), en échangeant les arguments et formant des combinaisons linéaires. La composante σ_a du spin résultant est alors

$$(9) \quad \sigma_a = \frac{n_a}{2} - r_a.$$

parce que σ_a est égal à la somme des composantes de chaque électron.

Notre fonction de spin φ_{r_a} est donc toujours symétrique. Nous savons que la fonction entière de l'atome Ψ_a doit être antisymétrique, à cause du principe de PAULI. L'antisymétrie de la fonction d'onde est une autre manière plus générale de formuler le principe d'exclusion, dont nous avons parlé dans le premier chapitre. Si Ψ_a est antisymétrique, et φ_{r_a} symétrique, la fonction u_a doit être antisymétrique. La fonction de SCHRÖDINGER d'un atome, pour lequel tous les électrons considérés ont des spins parallèles, est donc antisymétrique en ces électrons.

2.— Pour calculer l'énergie d'interaction des atomes, il faut d'abord former toutes les fonctions propres qui décrivent le système sans interaction. Si l'on a un système composé d'un certain nombre de parties sans interaction mutuelle, les fonctions d'onde du système entier sont le produit des fonctions d'onde des parties. Nous obtenons donc pour le système entier

$$(10) \quad \Psi = u_a \varphi_{r_a} u_b \varphi_{r_b} \dots$$

Cette fonction est déjà antisymétrique pour les électrons de l'atome A, pour ceux de l'atome B, etc. Mais pour satisfaire au principe de PAULI, il faut qu'elle soit antisymétrique pour *tous* les électrons. On peut y arriver en formant, par exemple, toutes les fonctions qui résultent de la première par échange des électrons de différents atomes de A et B. Nous allons représenter par P une permutation qui échange des électrons d'atomes différents ; leur nombre est

$$\frac{(n_a + n_b + \dots)!}{n_a! n_b! \dots}$$

On obtient alors une fonction antisymétrique en *tous* les électrons, en faisant agir tous les P sur la fonction (10), et en combinant d'une manière adéquate les différentes fonctions ainsi obtenues.

Soit γ_P un nombre égal à + 1 ou — 1 suivant que P est une permutation paire ou impaire. Il suffit, pour obtenir une fonction antisymétrique, d'écrire

$$(11) \quad \Psi_{r_a r_b} = c \cdot \sum_P \gamma_P \Psi u_a u_b \dots \varphi_{r_a} \varphi_{r_b} \dots,$$

où c représente un facteur de normalisation.

LA THÉORIE QUANTIQUE DES FORCES DE VALENCE

Nous avons formé ainsi des fonctions d'onde antisymétriques qui décrivent le système sans interaction. Evidemment il y en a plusieurs de ce genre, parce qu'on peut prendre pour chaque atome différentes fonctions de spin, par exemple pour l'atome A les fonctions $\varphi_0, \varphi_{r_a}, \varphi_{n_a}$, etc. ; c'est-à-dire que le nombre r_a des électrons de A qui sont dirigés vers le bas peut prendre les différentes valeurs $0, \dots, r_a, \dots, n_a$. Il y a donc au total $(n_a + 1)(n_b + 1) \dots$ fonctions d'onde antisymétriques $\Psi_{r_a r_b}$, distinctes.

Le système sans interaction est donc dégénéré et le degré de dégénérescence est égal

$$(n_a + 1)(n_b + 1) \dots$$

D'après les règles de la théorie des perturbations d'un système dégénéré il faudrait former la matrice de la fonction d'interaction

$$(12) \quad H_{r_a r_b \dots, r_a' r_b' \dots} = \int \Psi_{r_a r_b}^* H \Psi_{r_a' r_b' \dots}$$

où H représente la fonction d'interaction entre les atomes, c'est-à-dire l'interaction coulombienne entre toutes les charges de l'atome A et toutes les charges des autres atomes, etc. Une fois cette matrice connue, il faut chercher ses valeurs propres en résolvant l'équation

$$(13) \quad |H_{r, r'} - \varepsilon| = 0.$$

A chaque racine de (13) correspond une fonction d'onde *qui est une combinaison linéaire des différentes fonctions d'onde* $\Psi_{r_a r_b}$.

Les différentes racines de cette équation représentent les énergies d'interaction. Il est possible de simplifier beaucoup ce problème.

Si nos atomes A, B... avec leurs spins forment un système commun, les différents états obtenus pour le système total se distingueront par la valeur de l'énergie d'interaction et celle du spin résultant s . A chaque racine de (13) correspond une valeur s du spin résultant. En particulier, dans le cas de deux atomes il n'y a qu'une seule racine pour chaque valeur possible de s . Mais dans le cas de plusieurs atomes, à chaque valeur de s correspondent, en général, plusieurs énergies d'interaction et, par suite, plusieurs fonctions d'onde.

Nous allons maintenant former des fonctions d'onde correspondant aux différentes valeurs de s avant de calculer les racines ε . Cela est possible, parce que le spin résultant est caractérisé par ses propriétés

au cours de rotations. Si nous avons formé deux fonctions correspondant à deux valeurs différentes s et s' , de s , tous les éléments de la matrice $H_{s,s'} = \int \Psi_s^* H \Psi_{s'}$ sont nuls, parce que deux fonctions avec différents spins résultants ne forment jamais de combinaison. Il suffit donc de former une matrice où n'entrent que les différentes fonctions d'onde possédant le même spin résultant s . Il en résulte qu'en formant les combinaisons linéaires, qui correspondent à une certaine valeur définie de s , nous abaissons le degré de dégénérescence.

Avant d'arriver à former ces combinaisons, il faut parler de quelques faits mathématiques relatifs aux propriétés des fonctions de spin pendant les rotations du système de coordonnées.

3. — Considérons dans l'espace deux vecteurs orthogonaux de même longueur x, y, z et X, Y, Z , tels que :

$$\begin{aligned} x^2 + y^2 + z^2 &= X^2 + Y^2 + Z^2 = 1 \\ xX + yY + zZ &= 0. \end{aligned}$$

Si nous faisons tourner le système de coordonnées, les deux vecteurs subissent une transformation linéaire.

$$(14) \quad \begin{cases} x' = \text{lin}(x \dots) \\ X' = \text{lin}(X \dots). \end{cases}$$

Formons maintenant un nouveau vecteur complexe de composantes

$$\begin{aligned} \xi &= x + iX \\ \eta &= y + iY \\ \zeta &= z + iZ. \end{aligned}$$

Ce nouveau vecteur est tel que :

$$\xi^2 + \eta^2 + \zeta^2 = 0.$$

Si l'on pose

$$(15) \quad \begin{cases} A_1 = \sqrt{\xi + i\eta}, & A_1^2 = \xi + i\eta \\ A_2 = \sqrt{-\xi + i\eta}, & A_2^2 = -\xi + i\eta \end{cases}$$

il sera très aisé de démontrer que ces deux quantités A_1, A_2 sont les deux composants d'un vecteur complexe à deux dimensions, et su-

bissent une *transformation linéaire*, quand on fait tourner le système de coordonnées dans l'espace.

Formons en effet

$$A_1 A_2 = \sqrt{\xi^2 + \tau^2} = \zeta$$

et, en désignant par des accents les grandeurs transformées par rotation des coordonnées,

$$\begin{aligned} A_1' A_2' &= \zeta' = \text{Lin}(\xi, \tau, \zeta) \\ &= \text{Lin}(A_1^2, A_2^2, A_1 A_2) \end{aligned}$$

D'autre part on a

$$\begin{aligned} A_1' &= \sqrt{\text{lin}(A_1^2, A_2^2, A_1 A_2)} \\ A_2' &= \sqrt{\text{lin}'(A_1^2, A_2^2, A_1 A_2)} \end{aligned}$$

Or

$$A_1' A_2' = \sqrt{\text{lin} \cdot \text{lin}'} = \text{Lin}.$$

Si donc la racine d'un produit de deux fonctions linéaires est elle-même une fonction linéaire, il faut qu'il soit possible d'extraire aussi la racine de $\text{lin}(A_1^2, \dots)$ c'est-à-dire, que

$$A_1' = \text{lin}(A_1, A_2) = a_{11}A_1 + a_{12}A_2$$

et

$$(16) \quad A_2' = \text{lin}'(A_1, A_2) = a_{21}A_1 + a_{22}A_2.$$

Il est très simple de démontrer maintenant, que cette transformation est orthogonale dans le sens complexe (unitaire), c'est-à-dire que l'on a

$$(17) \quad a_{11}a_{12}^* + a_{21}a_{22}^* = 0$$

et que son déterminant est égal à un,

$$(18) \quad a_{11}a_{22} - a_{12}a_{21} = 1.$$

En résumé : A_1, A_2 forment un vecteur complexe A à deux dimensions. A chaque rotation dans l'espace correspond une transformation unitaire à deux dimensions. On peut exprimer les coefficients de la transformation $a_{11}, a_{12}, \dots$ en fonction des angles d'Euler φ, χ, ψ , de la rotation dans l'espace, ou plus exactement :

$a_{11}, \dots$, etc., sont des fonctions ambiguës des angles φ, χ, ψ (ambiguës, parce que le signe de la racine (14), etc., n'est pas déterminé).

On appelle cette transformation une *représentation* à deux dimensions de la rotation à trois dimensions.

4. — Si nous considérons de nouveau nos deux fonctions de spin α et β , nous savons, que ces deux fonctions subissent une transformation linéaire orthogonale au sens complexe et de déterminant égal à 1, lorsqu'on fait tourner le système des coordonnées dans l'espace. Et c'est justement la transformation (16) du vecteur A, qui est aussi la transformation de α , β , parce que on peut démontrer qu'il n'existe qu'une seule transformation à deux dimensions de ce type. Nous avons donc la correspondance

$$\alpha \sim A_1, \quad \beta \sim A_2$$

ou le signe $\sim$ veut dire : « se transforme comme ».

Considérons maintenant les fonctions de spin de nos atomes. Nous avons vu que ces fonctions de spin sont toutes symétriques ; elles ont, par exemple, la forme (8) pour un atome à n_a électrons.

Par conséquent, si pour chaque électron α et β se transforment comme A_1 et A_2 nos fonctions de spin symétriques se transformeront simplement comme les produits

$$(19) \quad \begin{cases} \varphi_0 \sim A_1^{n_a} \\ \dots\dots\dots \\ \varphi_{r_a} \sim A_1^{n_a - r_a} A_2^{r_a} \\ \dots\dots\dots \\ \varphi_{n_a} \sim A_2^{n_a} \end{cases}$$

Lorsque A_1 et A_2 subissent une transformation linéaire et unitaire, les produits symétriques

$$A_1^{n_a}, \quad \dots, \quad A_1^{n_a - r_a} A_2^{r_a}, \quad \dots, \quad A_2^{n_a}$$

subissent également une transformation linéaire, qui dépend des angles de rotation dans l'espace. Mais pour la rendre orthogonale, il faut introduire certains facteurs de normalisation dans chaque membre. Il faut donc que les quantités (19) soient les composantes d'un vecteur de longueur 1 dans un espace de $n + 1$ dimensions, c'est-à-dire que

$$c_1^2 |A_1^{n_a}|^2 + c_1^2 |A_1^{n_a - 1} A_2|^2 + \dots + c_{n_a}^2 |A_2^{n_a}|^2 = 1.$$

Et puisque nous savons que

$$|A_1|^2 + |A_2|^2 = 1$$

(17) et (18) nous apprennent qu'il faut poser $C_r = \sqrt{\binom{n_a}{r}}$

Nous pouvons alors remplacer nos fonctions de spin, dans la mesure où nous considérons leurs propriétés relatives aux rotations, par les grandeurs

$$(20) \quad \varphi_{r_a} = \sqrt{\binom{n_a}{r_a}} A_1^{n_a - r_a} A_2^{r_a}.$$

Les $n_a + 1$ fonctions de spin qui correspondent à la valeur du spin s_a subissent une *transformation* qui est *caractéristique de la valeur s_a* . Si $s = 0$ il n'y a qu'une seule fonction de spin, qui ne se transforme pas du tout : elle est *invariante*.

Retournons maintenant à notre problème de l'interaction de plusieurs atomes. Nous avons représenté chacune des fonctions propres des atomes A, B, comme le produit d'une fonction $u_a, u_b \dots$ des coordonnées des électrons de chaque atome, et d'une fonction des spins $\varphi_{r_a}, \varphi_{r_b} \dots$.

Si nous représentons la fonction de spin d'un atome par un vecteur A_1, A_2 , il faut introduire un tel vecteur pour chaque atome ; nous le désignerons par exemple par B_1, B_2 pour l'atome B. B_1 et B_2 se transforment exactement A_1 et A_2 . Les fonctions de spin et l'atome A seront alors représentées par

$$\varphi_{r_a} = \sqrt{\binom{n_a}{r_a}} A_1^{n_a - r_a} A_2^{r_a}$$

et celles de l'atome B par

$$(21) \quad \varphi_{r_b} = \sqrt{\binom{n_b}{r_b}} B_1^{n_b - r_b} B_2^{r_b}$$

La fonction $\Psi_{r_a r_b}$ se transforme comme le produit $\varphi_{r_a} \varphi_{r_b} \dots$ c'est-à-dire :

$$(22) \quad \Psi_{r_a r_b} \dots \propto A_1^{n_a - r_a} A_2^{r_a} B_1^{n_b - r_b} B_2^{r_b} \dots$$

Si nous cherchons maintenant les fonctions d'onde, qui correspondent à une certaine valeur s , il suffira de chercher des combinaisons linéaires des fonctions $\Psi_{r_a r_b} \dots$, qui se transforment suivant la transformation à $2s + 1$ dimensions, qui correspond à la valeur s du spin résultant.

Le cas le plus important est celui où $s = 0$, parce que nous devons nous attendre à obtenir une liaison entre les atomes, quand les spins des différents atomes se saturent au maximum. Il ne reste plus alors aucune valence libre dans le système total.

Nous ne parlerons pour le moment que du cas $s = 0$. On passe d'ailleurs, par une généralisation très simple, aux autres cas $s > 0$.

Quand $s = 0$, la fonction d'onde ne change pas au cours d'une rotation. Une fonction d'onde avec un spin $s = 0$ est un invariant de rotation.

Pour obtenir les fonctions d'onde pour lesquelles $s = 0$, il suffit de former toutes les combinaisons linéaires des fonctions $\Psi_{r_a r_b}$ qui sont invariantes. C'est un problème très simple. La fonction $\Psi_{r_a r_b}$ se transforme d'après (22).

Quelles sont donc les invariants de rotation ? Si nous considérons les deux vecteurs complexes A, B, nous obtenons immédiatement un invariant, en formant le déterminant.

$$(23) \quad [AB] = A_1 B_2 - A_2 B_1.$$

Cette expression est invariante, parce que le déterminant de la matrice de la transformation est égal à l'unité

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = 1.$$

Cependant, cette fonction $[AB]$ n'est pas encore une combinaison linéaire des fonctions $\Psi_{r_a r_b}$.

Pour cela, les composantes de A devraient apparaître partout à la puissance n_a , parce qu'il en est ainsi dans toutes les fonctions $\Psi_{r_a r_b}$. Mais formons des produits de la forme

$$(24) \quad [AB]^{p_{ab}} \cdot [AC]^{p_{ac}} [BC]^{p_{bc}} \dots$$

où

$$p_{ab} + p_{ac} + p_{ad} + \dots = n_a.$$

Ces produits sont invariants (parce que le produit de deux invariants est invariant). Alors, si nous satisfaisons aux conditions

$$(25) \quad \begin{cases} p_{ab} + p_{ac} + \dots = n_a \\ p_{ab} + p_{bc} + \dots = n_b \\ \dots \dots \dots \end{cases}$$

nous obtiendrons un invariant qui est une combinaison linéaire des fonctions $\Psi_{r_a r_b}$. Nous désignerons un tel invariant par

$$\Psi_{p_{ab} p_{ac}} \dots$$

En formant ainsi les fonctions (24) $\Psi_{p_{ab} p_{ac}}$ sous les conditions (25) nous obtenons des fonctions avec un spin total $s = 0$. Et nous les obtenons toutes. Ces fonctions (24) constituent le point de départ du calcul de la perturbation entre atomes.

Il faut remarquer que ces fonctions ne sont pas linéairement indépendantes. On démontre à partir de (23), l'identité

$$(26) \quad [AB][CD] + [AC][DB] + [AD][BC] = 0.$$

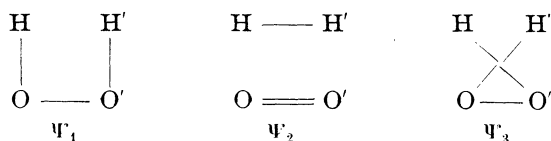
d'où l'on déduit des relations analogues entre les différentes fonctions $\Psi_{p_{ab} p_{ac}} \dots$

5. — Ces fonctions ont une signification physique ou chimique très simple et intuitive. Former une parenthèse telle que

$$[AB] = A_1 B_2 - A_1 B_2 = \gamma(1)\gamma(2) - \gamma(1)\gamma(2),$$

signifie constituer une paire d'électrons qui ont saturé leur spin, avec un électron de l'atome A et un de B. Si nous avons formé une fonction $[AB]^{p_{ab}} [AC]^{p_{ac}} \dots$ nous pouvons appeler p_{ab} le nombre des paires d'électrons reliant A et B, ou — comme nous savons qu'une paire d'électrons à spins antiparallèles signifie un trait de valence, — le nombre de traits de valence reliant A et B. De même p_{ac} sera le nombre de traits de valence reliant A et C, etc. Une fonction $\Psi_{p_{ab} p_{ac}}$ représente donc une certaine distribution des traits de valence entre les atomes. Naturellement il existe en général plusieurs distributions possibles de ce genre.

Il faut remarquer qu'une telle distribution des traits de valence existe pour chaque position relative des atomes (nous n'avons encore jamais parlé de la position des atomes dans l'espace). Par exemple, si nous considérons deux atomes à deux électrons de valence (oxygène) et deux atomes d'hydrogène, on trouve les distributions suivantes des traits de valence :



W. HEITLER

Chaque distribution représente une fonction d'onde, que nous appelons un *état pur de valence*. Ces trois états purs de valence existent pour chaque position des atomes dans l'espace.

En effet il existe une relation linéaire entre ces trois états. D'après (30) nous avons

$$[\text{OO}'] + [\text{OO}'] [\text{HH}'] + [\text{OH}] [\text{H'O}'] + [\text{OH}'] [\text{O'H}] = 0$$

ou

$$\Psi_2 + \Psi_1 + \Psi_3 = 0.$$

Une telle relation ne peut pas s'exprimer dans l'image des traits de valence.

III

L'énergie d'interaction

1. — Nous allons maintenant calculer l'énergie de l'interaction entre ces atomes. Dans ce but, il faut faire agir l'opérateur d'interaction H sur toutes les fonctions $\Psi_{p_{ab}} \Psi_{p_{ac}} \dots$. Cet opérateur H représente l'énergie mutuelle de Coulomb toutes des charges de l'atome A et celles de B, de C, etc. Mais H ne contient pas les variables de spin. D'après les règles générales de la mécanique quantique, l'énergie ε de l'interaction est donnée par :

$$H\Psi = \varepsilon\Psi.$$

Multiplions cette équation par $u_a^* u_b^* \dots$ et intégrons sur tout l'espace des coordonnées de tous les électrons de A, de B, etc. Nous obtenons alors

$$\begin{aligned} (27) \quad \sum_P r_P P_{p_{ab} p_{ac}} \dots \int u_a^* u_b^* \dots H P u_a u_b \dots \\ = \varepsilon \sum_P r_P P_{p_{ab} p_{ac}} \dots \int u_a^* u_b^* \dots P u_a u_b. \end{aligned}$$

Nous considérons d'abord les intégrales et calculerons ensuite l'influence des permutations P sur les fonctions de spin.

P représente toutes les permutations entre électrons d'atomes différents ; il y en a de différents types.

a) d'abord la permutation identique E qui ne change rien. L'intégrale du premier membre de (27)

$$(28) \quad \int |u_a u_b u_c \dots|^2 H = C.$$

représente alors simplement l'interaction de Coulomb des charges représentées par les fonctions de SCHRÖDINGER $|u_a|^2$, etc. Nous appelons cet intégrale « l'intégrale de COULOMB ».

L'intégrale du second membre est, pour la permutation E, égale à 1. car les fonctions u_a , etc., sont normalisées.

b) P représente une permutation de deux électrons, par exemple d'un électron de A avec un électron de B. Nous appellerons une telle permutation une transposition T_{ab} . L'intégrale, que nous désignerons par (AB)

$$(29) \quad \int u_a^* u_b^* \dots H T_{ab} u_a u_b \dots = AB,$$

ne dépend pas du fait que tel ou tel électron de A et de B sont échangés. Nous appelons cette intégrale « l'intégrale d'échange ». C'est la quantité fondamentale de la théorie. C'est justement la quantité que le premier chapitre nous a fait connaître sous le nom « énergie de couplage des vecteurs de spin », et que nous avons désignée par A.

L'intégrale correspondante du second membre de (27)

$$\int u_a^* u_b^* \dots T_{ab} u_a u_b \dots$$

est très petite. Si les fonctions d'onde u_a et u_b de différents atomes étaient exactement orthogonales, cet intégrale serait nulle. Les fonctions u_a et u_b ne sont pas exactement orthogonales, et pour cette raison l'intégrale est très petite. Nous pouvons la négliger.

c) Si nous considérons maintenant des permutations supérieures, où plus de deux électrons sont échangés, les intégrales correspondantes sont aussi très petites. Cela tient à ce que la fonction d'interaction H

est une somme de termes, dont chacun ne contient que l'interaction de *deux électrons*. A cause de la presque-orthogonalité des fonctions u_a, u_b , etc., toutes nos intégrales seront petites, et nous les négligerons.

D'autre part :

$$(30) \quad \left\{ \begin{array}{l} \text{Si } P = E, \text{ notre symbole } r_p = 1 \\ \text{et comme } P = i_{ab} \text{ est impair, } r_p = -1. \end{array} \right.$$

Nous obtenons donc finalement l'équation

$$(31) \quad \varepsilon \varphi_p = \varphi_p \cdot C - \sum_{a,b} (AB) \sum_{Tab} T_{ab} \varphi_p = H \varphi_p$$

qui nous donne ε .

La somme $\sum_{Tab}$ est étendue à toutes les transpositions entre électrons des atomes différents A et B, et la $\sum_{a,b}$ est étendue à toutes les paires de deux atomes.

2. — Il reste à calculer l'opérateur

$$\sum_{Tab} T_{ab} \varphi_{p_{ab} p_a} \dots$$

On peut mettre en évidence très simplement l'effet de l'opérateur sur la fonction de spin $\varphi_{p_{ab} p_{ac}}$.

Considérons la fonction

$$(32) \quad \varphi = [AB]^{p_{ab}} [AX]^{p_{ax}} [BY]^{p_{by}} \dots$$

Cette fonction représente un état, dans lequel tous les électrons de A par exemple, sont liés avec des électrons d'autres atomes (X, Y) pour former des paires. Quelques électrons de A sont liés avec des électrons de X : le nombre de ces électrons est p_{ax} ; d'autres sont liés avec des électrons d'autres atomes. De même, il y a par exemple p_{by} électrons de B, qui sont liés avec des électrons de Y, etc. Si donc nous échangeons successivement tous les électrons de A avec tous les électrons de B, il faudra échanger aussi chaque fois un de ces p_{ax} électrons de A, qui sont liés avec X, avec un des électrons de B, qui sont liés

avec Y. L'effet d'une telle permutation est qu'un électron de A est lié maintenant avec Y, tandis qu'un électron de B est lié avec X, au lieu d'une liaison entre B et Y. Il y a p_{ax}, p_{by} permutations de ce genre parce qu'il y a p_{ax} liaisons entre A et X et, par suite, p_{ax} électrons liés avec X. Ces permutations donnent lieu alors à la transformation

$$(33) \quad \varphi \rightarrow \varphi \cdot \frac{[AY][BX]}{[AX][BY]} \cdot p_{ax} p_{by}.$$

Il y a naturellement aussi des électrons de A qui sont liés avec d'autres atomes, et nous voyons, que nous avons à faire une sommation relative à tous les atomes X, Y :

$$(34) \quad \sum_{XY} p_{ax} p_{by} \cdot \varphi p_{ax} p_{ay} \cdots \frac{[AY][BX]}{[AX][BY]}.$$

Nous devons encore examiner l'effet d'une permutation qui échange deux électrons de A et B formant eux-mêmes une paire. Il y a p_{ab} paires de ce genre. Si nous considérons la fonction de spin qui représente une telle liaison, $[AB] = A_1 B_2 - A_2 B_1 = \alpha(1)\beta(2) - \alpha(2)\beta(1)$ nous voyons, qu'elle est antisymétrique entre les deux électrons : l'effet d'une permutation du type considéré est de changer son signe. Comme il y a p_{ab} paires entre A et B, il existe aussi p_{ab} permutations de ce genre et nous obtenons au total :

$$(35) \quad t_{ab}\varphi = \sum_{T_{ab}} T_{ab}\varphi = -p_{ab}\varphi + \sum_{XY} p_{ax} p_{by} \varphi \frac{[AY][BX]}{[AX][BY]}.$$

Le terme de droite contient, au lieu d'une liaison AX et BY, deux autres liaisons AY et BX. Mais le nombre des liaisons qui sortent de A, B, X, Y est chaque fois le même que dans la fonction φ , c'est-à-dire que $\frac{[AY][BX]}{[AX][BY]}$ représente aussi un invariant de notre problème. Nous voyons ainsi que l'opérateur $t_{ab} = \sum_{T_{ab}} T_{ab}$, agissant sur φ produit, une combinaison linéaire d'autres invariants φ .

3. — Considérons maintenant l'opérateur $H = C - \sum_{a, b} (AB)t_{ab}$;

cet opérateur produit aussi une combinaison linéaire d'invariants. Le nombre des invariants indépendants étant égal à f , choisissons une série $\varphi_1 \dots \varphi_f$ d'invariants indépendants. L'effet de $H\varphi$ peut être développé aussi suivant cette série de fonctions indépendantes, sous forme de combinaison linéaire de $\varphi_1 \dots \varphi_f$. Nous les appelons une *base*.

Les équations

$$(36) \quad H\varphi_p = \varepsilon \varphi_p$$

forment un système d'équations linéaires et homogènes pour les variables φ_p . Ces équations ne peuvent être résolues que si leur déterminant est égal à zéro. Nous obtenons ainsi une équation du f -ième degré pour notre énergie d'interaction ε . H est une fonction de l'intégrale de COULOMB C et de toutes les intégrales d'échange (AB) ; ε sera par conséquent aussi une fonction des intégrales d'échange (AB). ε dépendra donc de la structure des atomes et de leurs distances.

Une discussion détaillée montre que l'intégrale d'échange (AB) dépend essentiellement de la structure de A et B et de leur distance, mais presque pas des autres atomes. Toutes ces intégrales (AB), (AC), seront du même ordre de grandeur et dépendront des distances R_{ab} , R_{ac} à peu près de la même manière. Si nous avons calculé ε en fonction de ces intégrales, — cette fonction ne dépendra que des nombres de valence de tous les atomes, on peut facilement discuter l'interaction des atomes en fonction de leur position.

4. — Pour les applications pratiques de la théorie, il est très important d'avoir une méthode générale, pour construire d'une manière unique une série d'états purs de valence indépendants. RUMER nous a montré que cela est facile.

Considérons la relation linéaire :

$$[AB][CD] + [AC][DB] + [AD][BC] = 0.$$

Si nous disposons les quatre atomes en carré, nous pouvons représenter cette équation de la manière suivante :

$$(37) \quad \left\{ \begin{array}{ccc} A \longrightarrow B & A & D \\ & + \downarrow & \downarrow \\ C \longrightarrow D & C & B \end{array} \right. = \begin{array}{cc} A & B \\ \swarrow & \searrow \\ C & D \end{array}$$

Puisque $[AB] = -[BA]$ il faut munir d'une flèche chaque trait

de valence. On voit qu'il est possible d'exprimer un état pur de valence, où deux traits de valence sont croisés linéairement, par des états purs de valence, où le croisement n'existe pas.

De cette remarque résulte une méthode générale très simple permettant de construire une base d'états purs de valence indépendants. Disposons tous les atomes en cercle. Naturellement, cela n'a rien à faire avec la position réelle des atomes dans l'espace, qui peut être quelconque. Formons alors tous les états purs de valence où il n'existe aucun croisement entre deux traits de valence. On peut démontrer facilement que ces états de valence sont linéairement indépendants, et on obtient ainsi une série définie univoquement d'états purs de valence indépendants.

5. — Nous n'avons parlé que du cas où toutes les valences sont saturées, c'est-à-dire, où le spin résultant est égal à zéro. Une généralisation très simple étend la théorie au cas où le spin résultant s est supérieur à zéro. Il suffit d'introduire un nouvel atome L , correspondant à l'espace libre, de manière à l'accrocher aux spins libres et de construire tous les invariants

$$(38) \quad \varpi_{p_{al}} \cdots = [AL]^{p_{al}} [AB]^{p_{ab}} [BL]^{p_{bl}}$$

où $[AL]$ représente maintenant un trait de valence qui sort de A et n'est pas saturé. La somme

$$p_{al} + p_{bl} + \cdots$$

représente naturellement le nombre de traits de valence libres, c'est-à-dire

$$(39) \quad p_{al} + p_{bl} + \cdots = 2s.$$

IV

Applications

1. — Nous allons maintenant faire des applications de notre théorie générale. Considérons le cas le plus simple de deux atomes. Soient A

Pour discuter ce résultat, il faut connaître les intégrales C et (AB) . Ces deux intégrales ne dépendent que de la structure des atomes et de leur distance. On les a calculées pour quelques paires d'atomes, par exemple H_2 , Li_2 , LiH , etc. et on a trouvé, que dans tous ces cas, (AB) est négatif et plus grand que C pour des distances supérieures à celle de l'équilibre. C'est aussi négatif aux grandes distances et infini positif aux petites, à cause de la répulsion infinie des noyaux. Pour le cas de deux atomes avec un électron, par exemple pour les alcalins, le nombre des traits de valence p peut être égal à 1 ou à zéro, et l'énergie sera $\varepsilon = C \pm (AB)$. C'est justement la même formule que nous avons déjà rencontré dans le premier chapitre, où nous avons appelé (AB) l'énergie de couplage entre les deux vecteurs de spin. Si nous traçons la courbe de ces deux énergies en fonction de R nous obtenons le résultat de la figure 1 : l'énergie $\varepsilon = C + (AB)$ est négative et possède un minimum qui correspond à la formation d'une molécule. Nous voyons donc que la formation d'une molécule a lieu d'accord avec les données de la chimie.

En général, la loi de variation des intégrales C et (AB) est la même, quand les deux atomes ne possèdent aucun moment de quantité de mouvement, c'est-à-dire quand le nombre des quanta $l = 0$. Notre formule générale de l'énergie d'interaction nous montre alors que l'énergie négative est minimum quand le nombre des traits de valence est aussi grand que possible. Dans ce cas, elle s'écrit

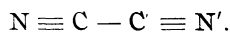
$$(43) \quad \varepsilon = C + p(AB)$$

ce qui correspond au niveau le plus bas de la molécule. Dans le cas $p = 0$, l'énergie a pour valeur

$$(44) \quad \varepsilon = C - n_a n_b (AB)$$

et les deux atomes se repoussent. Tous ces résultats sont naturellement d'accord avec ceux obtenus en chimie.

2. — Nous allons traiter maintenant un cas simple de molécule à plusieurs atomes. Considérons par exemple la molécule C_2N_2 avec la formule chimique



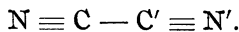
Il faut calculer, en fonction de leur position mutuelle, l'énergie d'interaction de quatre atomes C, C', N, N' avec un spin résultant $O = 1$. Dans ce but, il faut former d'abord les invariants, c'est-à-dire les états purs de valence. Mais dans ce cas, il y a plusieurs invariants. Parmi eux, nous choisirons une certaine série d'invariants, linéairement indépendants. D'après les règles générales esquissées dans le chapitre III. 4, nous obtenons une telle série en disposant pour un moment les quatre atomes en cercle et formant tous les états purs de valence sans croisement de traits de valence.

$$(45) \quad \left\{ \begin{array}{ll} \varphi_1 = \begin{array}{cc} \text{C} \rightarrow \text{C}' \\ \downarrow \downarrow \downarrow \downarrow \\ \text{N} \quad \text{N}' \end{array} & \varphi_2 = \begin{array}{cc} \text{C} \rightarrow \text{C}' \\ \downarrow \downarrow \downarrow \downarrow \\ \text{N} \rightarrow \text{N}' \end{array} \\ \varphi_3 = \begin{array}{cc} \text{C} \rightleftharpoons \text{C}' \\ \downarrow \quad \downarrow \\ \text{N} \rightleftharpoons \text{N}' \end{array} & \varphi_4 = \begin{array}{cc} \text{C} \rightleftharpoons \text{C}' \\ \downarrow \quad \downarrow \\ \text{N} \rightleftharpoons \text{N}' \end{array} \end{array} \right.$$

Nous obtenons ainsi 4 invariants indépendants. L'énergie d'interaction est donc donnée par un déterminant du 4^e ordre. Pour faire ces calculs, il faut d'abord évaluer l'action des opérateurs $t_{CC'}$, t_{CN} , $t_{C'N'}$, etc., sur ces quatre fonctions. D'après les règles générales de notre théorie, pour obtenir par exemple $t_{CC'}\varphi_1$, il faut échanger tous les traits de valence sortants de C et C'. L'équation (35) donne, en tenant compte de (37),

$$t_{CC'}\varphi_1 = -\varphi_1 + 9 \cdot \overbrace{\text{N} = \text{C} - \text{C}' = \text{N}'}^{\text{N} = \text{C} - \text{C}' = \text{N}'} = 8\varphi_1 - 9\varphi_2,$$

Pour simplifier le problème, nous supposons que les quatre atomes se trouvent en ligne conformément à la formule chimique



Admettons d'abord que les distances R_{CN} et $R_{C'N'}$ soient égales. Les distances $R_{NN'}$, $R_{CN'}$, $R_{C'N}$ sont plus grandes que R_{CC} et $R_{CN} = R_{C'N'}$. Les intégrales d'échange (C'N) etc., sont alors petites comparées avec (CC'), (CN) = (C'N'). Nous les négligerons donc, ne prenant en considération que les intégrales d'échange entre atomes voisins. Seules sont différentes de zéro les intégrales suivantes :

$C, (CC'), (CN) = (C'N')$. La fonction de HAMILTON de notre problème est alors d'après (31) et (35)

$$H = -t_{CC'}(CC') - (t_{CN} + t_{C'N'})((C'N') + (CN)).$$

Si nous calculons l'action des deux opérateurs $t_{CC'}$ et $t_{CN} + t_{C'N'}$ sur les quatre fonctions pures de valence et si nous écrivons pour abréger

$$(CC') = \alpha, \quad (CN) = \beta, \quad \varepsilon - C = \varepsilon$$

nous obtenons l'équation séculaire suivante pour l'énergie d'interaction.

$$(46) \begin{vmatrix} 8\alpha - 6\beta + \varepsilon & -9\alpha & 0 & 0 \\ -4\beta & 4\alpha + \varepsilon & -4\alpha & 0 \\ 0 & -12\beta & -2\alpha + 10\beta + \varepsilon & -\alpha \\ 0 & 0 & -24\beta & -4\alpha + 24\beta + \varepsilon \end{vmatrix} = 0$$

qui est une équation algébrique du 4^e ordre pour ε . Cette énergie d'interaction ε ne dépend que d'une seule variable, à savoir du rapport $z = \frac{\alpha}{\beta} = \frac{(CC')}{(CN)}$. La figure 2 montre la courbe de $\frac{\varepsilon - C}{(CN)}$ en fonction de

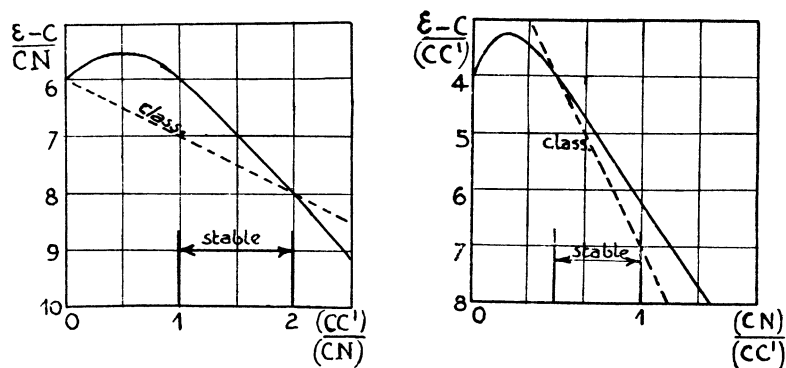


FIG. 2.

$z = \frac{(CC')}{(CN)}$ et $\frac{\varepsilon - C}{(CC')}$ en fonction de $\frac{(CN)}{(CC')}$. On voit que pour $(CN) = 0$, c'est-à-dire quand la distance R_{ON} est très grande comparée à $R_{CC'}$, $\varepsilon = C + 4(CC')$. Dans ce cas, on a une molécule C_2 de deux atomes de carbone, qui possède exactement l'énergie donnée par la formule relative à deux atomes isolés. Si l'on donne des valeurs plus grandes à z ,

c'est-à-dire, si l'on rapproche les atomes N, la courbe de ε monte d'abord, c'est-à-dire, les atomes N sont repoussés par la molécule saturée C_2 . Plus loin, la courbe descend, et les atomes N sont attirés, formant alors une molécule C_2N_2 . Les 4 traits de valence entre $C \equiv C$ sont séparés et sont employés à lier les atomes N de manière à constituer peu à peu une molécule $N \equiv C - C \equiv N$. Donc, pour former la molécule C_2N_2 à partir d'une molécule C_2 et deux atomes N, il est nécessaire de dépenser d'abord une *certaine énergie*.

On appelle cette énergie, en chimie, l'énergie d'*activation*. La fonction propre de l'état de la molécule est une combinaison linéaire des états purs de valence (45)

$$\Psi = c_1\varphi_1 + c_2\varphi_2 + c_3\varphi_3 + c_4\varphi_4$$

où les coefficients c_1, c_4 dépendent de $z = \frac{(CC')}{(CN)}$.

Dans le domaine, où la courbe descend, formant une molécule $N \equiv C - C' \equiv N'$ on peut montrer, que $c_1 \sim 1$, c'est-à-dire, que la fonction d'onde est *presque identique* à celle de l'état pur de valence φ_1 , qui correspond à la *formule chimique*.

Si nous considérons l'autre fig. 2 nous voyons que pour $\frac{(CC')}{(CN)} = 0$ l'énergie $C = C + 6(CN)$, qui est l'énergie de deux molécules non saturées $-C \equiv N$. Comme ce groupe $-CN$ possède une valence libre, on pourrait s'attendre à une attraction mutuelle des deux groupes $-CN$. Mais elle ne se produit pas. Nous voyons que la courbe monte également. C'est un phénomène très intéressant parce qu'il faut d'abord dépenser aussi une certaine énergie d'activation pour former une molécule C_1N_2 à partir de deux groupes CN. On a trouvé depuis peu de temps, par l'expérience, une énergie analogue d'activation dans la formation de C_2H_6 à partir de deux groupes CH_3 .

Après une petite montée d'activation, la courbe de ε descend et forme une molécule C_2N_2 . Dans la figure, le domaine, où la molécule est *stable* est encadré (la courbe pointillée se rapporte au chapitre V.).

V

Théorie semi-classique

1. — La théorie générale, que nous avons développée jusqu'ici, conduit, dans les cas simples, à des résultats qui sont généralement d'accord avec l'expérience. Mais l'idée de valence a trouvé ses applications les plus fécondes en chimie organique, c'est-à-dire dans le cas des molécules très grandes. Dans ce domaine règnent des lois très simples : par exemple l'énergie d'une telle molécule se compose additivement des contributions de chaque trait de valence. Malheureusement, il est très difficile d'appliquer la théorie à ces cas, parce que le nombre des états purs de valence commence à devenir extraordinairement grand, et il faudrait résoudre une équation algébrique de degré très élevé.

Pour appliquer la théorie à ces cas compliqués, il est peut-être permis de lui faire subir une simplification radicale, qui permet de traiter ces molécules en toute généralité, mais en renonçant à une description très détaillée.

Nous avons obtenu l'opérateur

$$t_{ab}\varphi = -p_{ab}\varphi + \sum_{X,Y} \varphi \frac{[AY][BX]}{[AX][BY]} \cdot p_{ax}p_{by}.$$

Introduisons, au lieu des fonctions de spins φ , d'autres fonctions, qui sont normalisées d'une autre manière.

$$(47) \quad \chi_{p_{ab} \dots} = - \frac{\varphi_{p_{ab} \dots}}{\sqrt{p_{ab}! p_{ax}! p_{by}! \dots}}.$$

Nous obtenons, en remarquant que

$$\begin{aligned} \varphi \frac{[AY][BX]}{[AX][BY]} &= \varphi(p_{ax} - 1, p_{by} - 1, p_{ay} + 1, p_{bx} + 1) \\ (48) \quad t_{ab}\chi_{p_{ab} \dots} &= -p_{ab}\chi_{p_{ab} \dots} \\ &+ \sum_{X,Y} \sqrt{p_{ax}p_{by}(p_{ay} + 1)(p_{bx} + 1)} \chi_{(p_{ax} - 1, p_{bx} + 1, p_{ay} + 1, p_{by} - 1)}. \end{aligned}$$

On peut définir alors deux opérateurs ∂_{ax}^+ et ∂_{ax}^- tels que ∂_{ax}^+ agissant sur χ élève son argument p_{ax} d'une unité et que ∂_{ax}^- abaisse cet argument d'une unité

$$(49) \quad \begin{cases} \partial_{ax}^+ \chi(p_{ax} \dots) = \chi(p_{ax} + 1) \\ \partial_{ax}^- \chi(p_{ax} \dots) = \chi(p_{ax} - 1). \end{cases}$$

Si enfin p_{ax} représente l'opérateur qui multiplie une fonction par son argument, nous pouvons définir deux autres opérateurs :

$$\gamma_{ax} \chi = \sqrt{p_{ax}} \partial_{ax}^+ \chi = \sqrt{p_{ax} + 1} \chi(p_{ax} + 1)$$

et

$$(50) \quad \gamma_{ax}^* \chi = \partial_{ax}^- \sqrt{p_{ax}} \chi = \sqrt{p_{ax}} \chi(p_{ax} - 1).$$

On voit que l'opérateur t_{ab} de notre théorie est égal à

$$(51) \quad t_{ab} = -p_{ab} + \sum_{x, y} \gamma_{ay} \gamma_{bx} \gamma_{ax}^* \gamma_{by}^*.$$

Considérons encore les opérateurs γ_{ax} , γ_{ax}^* . Si nous formons les deux produits

$$\begin{aligned} \gamma_{ax} \gamma_{ax}^* &= \sqrt{p_{ax}} \partial_{ax}^+ \partial_{ax}^- \sqrt{p_{ax}} = p_{ax} \\ \gamma_{ax}^* \gamma_{ax} &= \partial_{ax}^- \sqrt{p_{ax}} \sqrt{p_{ax}} \partial_{ax}^+ = p_{ax} + 1, \end{aligned}$$

nous obtenons

$$(52) \quad \gamma_{ax}^* \gamma_{ax} - \gamma_{ax} \gamma_{ax}^* = 1.$$

On voit que les deux opérateurs $\gamma_{ax}^* \frac{h}{2\pi i}$, et γ_{ax} sont *canoniquement conjugués*. Nous avons écrit notre opérateur t_{ab} de la chimie, en fonction de variables canoniques. Nous connaissons déjà des opérateurs de ce genre, par exemple dans la théorie de la lumière. DIRAC a introduit des opérateurs analogues pour obtenir les quanta de la lumière. Le nombre des quanta de la fréquence ν , n_ν , est aussi égal au produit

$$b_\nu^* b_\nu = n_\nu$$

de deux opérateurs canoniquement conjugués.

C'est une conséquence de la théorie générale des quanta, que

$$n_\nu = b_\nu^* b_\nu$$

doit être un *nombre entier*. Nous voyons que, dans notre théorie de la chimie, $p_{ax} = \gamma_{ax}^* \gamma_{ax}$ est aussi un nombre entier. Le nombre des traits de valence entre deux atomes correspond en un certain sens au nombre des quanta de lumière de la fréquence ν dans la théorie de DIRAC.

La fonction d'énergie H s'exprime en fonction de variables canoniques par

$$(53) \quad H = C + \sum_{a, b} (AB) t_{ab}$$

$$t_{ab} = -\gamma_{ab} \gamma_{ab}^* + \sum_{x, y} \gamma_{ay} \gamma_{ax} \gamma_{ax}^* \gamma_{by}^*$$

ou, si nous séparons dans la somme, les membres pour lesquels $x = y$, nous obtenons

$$(54) \quad t_{ab} = -p_{ab} + \sum_x p_{ax} p_{bx} + \sum_{x \neq y} \gamma_{ay} \gamma_{bx} \gamma_{ax}^* \gamma_{by}^*$$

Dans cette fonction de HAMILTON H , l'intégrale C et les intégrales d'échange interviennent encore comme paramètres. Déterminer l'énergie d'interaction revient à trouver les valeurs propres de cette fonction de HAMILTON H , exprimée en fonction des variables canoniques, ce qui est un problème quantique ordinaire.

2. — Nous passerons maintenant à une théorie *semi-classique*, considérant H comme une fonction de HAMILTON classique et les γ_{ax} , γ_{ax}^* comme des variables canoniques classiques, semblable aux q et p . Nous obtiendrons ainsi une mécanique classique pour les γ_{ax} , qui seront des fonctions du temps t . La manière dont γ_{ax} dépend du temps est décrite par les équations classiques de HAMILTON

$$(55) \quad \frac{\partial H}{\partial q} = -\dot{p}, \quad \frac{\partial H}{\partial p} = +\dot{q},$$

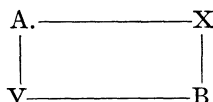
soit, dans notre cas, où $\gamma_{ax}^* \frac{h}{2\pi i}$ correspond à p et γ_{ax} à q ;

$$(56) \quad \frac{\partial H}{\partial \gamma_{ax}} = -\frac{h}{2\pi i} \dot{\gamma}_{ax}^*, \quad \frac{\partial H}{\partial \gamma_{ax}^*} = \frac{h}{2\pi i} \dot{\gamma}_{ax}$$

Dans cette théorie classique, les nombres p_{ab} ne sont pas nécessairement entiers, pour le moment.

Avant de chercher des solutions de ces équations, il faut examiner rapidement les conditions dans lesquelles il est permis de passer à cette théorie classique. Cette théorie représenterait une bonne approximation si le nombre des traits de valence p_{ab} était grand. De même, la théorie classique de la lumière constitue une bonne approximation, quand le nombre des quanta ν est très grand. Mais en chimie p_{ab} est égal à 1, 2 ou tout au plus 3. Pourtant la mécanique classique nous fournit bien une approximation de la mécanique quantique, si l'on impose à un certain nombre de quantités, par exemple au nombre des quanta de lumière des conditions de *quantification*. L'ancienne théorie de BOHR est fondée sur de telles hypothèses. Nous pouvons espérer que nous obtiendrons une certaine approximation de la réalité, même pour des petits nombres de quanta p_{ab} , si nous imposons plus tard, à ces nombres p_{ab} , etc., la condition d'être entiers.

3. — On peut trouver facilement des solutions très générales de nos équations classiques. Supposons qu'à un certain instant $t = 0$, nous ayons affaire à certaines valeurs des γ_{ax} et γ_{ax}^* . Ce sont nos conditions initiales. Nous admettons en outre qu'à l'instant $t = 0$ il ne se trouve pas quatre atomes A, B, X, Y, qui soient tous joints par des traits de valence suivant le schéma :



Supposons qu'un des quatre nombres p_{ax} , p_{ay} , p_{bx} , p_{by} soit égal à zéro. Dans ce cas le produit $\gamma_{ay}/\gamma_{bx} \gamma_{ax}^*/\gamma_{by}^*$ est nul parce que les γ_{ax} sont proportionnels à $\sqrt{p_{ax}}$, etc. Alors

$$(57) \quad \begin{cases} \frac{h}{2\pi i} \dot{\gamma}_{ax}^* = \frac{\partial H}{\partial \gamma_{ax}} = \sum_{a,b} (AB) \frac{\partial t_{ab}}{\partial \gamma_{ax}} = (AX) \gamma_{ax}^* - \sum_b (AB) p_{bx} \gamma_{ax}^* \\ \frac{h}{2\pi i} \dot{\gamma}_{ax} = - \frac{\partial H}{\partial \gamma_{ax}^*} = - (AX) \gamma_{ax} + \sum_b (AB) p_{bx} \gamma_{ax}, \end{cases}$$

et

$$\dot{p}_{ax} = \gamma_{ax} \dot{\gamma}_{ax}^* + \gamma_{ax}^* \dot{\gamma}_{ax} = 0.$$

Et nous voyons que $p_{ax} = \text{const.}$, pour $t = 0$, c'est-à-dire, dans ce cas, tous les nombres p des traits de valence sont constants au temps $t = 0$. On obtient une solution générale, si l'on pose

$$(58) \quad p_{ab} = \text{const.}$$

à tout instant.

On voit que p_{ab} ne change pas au cours du temps. On peut exiger alors, que ces nombres p_{ab} soient des nombres entiers. Cela correspond en un certain sens à l'ancienne théorie de BOHR.

Il est très facile de trouver l'énergie du système. Cette énergie est égale simplement à la fonction H , donc à

$$(59) \quad \varepsilon = C + \sum_{a, b} p_{ab}(AB) - \sum_{a, b, x} (AB)p_{ax}p_{bx}.$$

Nous allons discuter maintenant cette formule.

On voit que l'énergie se compose de différentes parties. Prenons d'abord les termes $C + \sum_{a, b} (AB)p_{ab}$. C est l'énergie de COULOMB de sys-

tème tout entier. Les termes $(AB)p_{ab}$ donnent une contribution qui se compose *additivement* des contributions de chaque trait de valence. La contribution de chaque trait de valence est l'énergie (AB) , c'est-à-dire *l'énergie de l'intégrale d'échange*. Cette énergie se présente avec un signe positif, c'est-à-dire que atomes A et B qui sont liés avec des traits de valence s'attirent.

La troisième partie $-\sum_x (AB)p_{ax}p_{bx}$ n'est différente de zéro que dans le cas où les deux atomes A et B sont liés avec un troisième atome X. Les traits de valence entre A et B effectuent alors une *répulsion* entre A et B égale à $-(AB)p_{ax}p_{bx}$. En effet la chimie organique nous apprend que toutes les molécules ayant un cycle de traits de valence sont moins stables que les autres molécules.

Abstraction faite de ce résultat, nous voyons que notre théorie semi-classique correspond à peu près aux images que l'on s'est faites de la structure des molécules en chimie organique.

Il est intéressant de savoir comment cette théorie se comporte vis-à-vis de la théorie exacte. On peut comparer les courbes de l'énergie que nous avons calculées au chapitre IV pour la molécule C_2N_2 ,

W. HEITLER

avec les courbes, obtenues à l'aide de notre théorie semi-classique et qui sont dessinées en pointillé sur la figure 2.

On voit que justement dans le domaine où la molécule est stable, les courbes ne sont pas très différentes. Mais d'autre part, pour les valeurs très petites ou très grandes de z , les courbes diffèrent beaucoup. Les phénomènes d'activation s'estompent à ce degré d'approximation.

Conférences faites à l'Institut Henri-Poincaré en mars 1933.

Manuscrit reçu le 14 Avril 1933.

— — — — —