



Applications de la Mécanique Quantique

De l'atome au solide

(FIP1 – L6)

Claude ASLANGUL
Université Pierre et Marie Curie
(Paris 6)
aslangul@lptmc.jussieu.fr

Préambule

Ce cours vise à présenter un tour d'horizon des premières applications de la Mécanique Quantique en Physique des basses énergies, de l'ordre de la dizaine d'eV au plus. Ses trois parties correspondent au découpage naturel qui résulte de ce choix.

La première partie (6 à 7 cours) traite de la physique de l'atome, en repartant du cas le plus simple, l'atome d'hydrogène (ch. 1) ; les buts principaux sont d'une part d'argumenter physiquement sur la nécessité du spin, d'autre part d'introduire les notions élémentaires permettant, dans un traitement semi-classique, de jeter les bases de la spectroscopie atomique. Le chapitre 2 pose le problème de l'indiscernabilité des particules identiques, énonce sa résolution et examine les premières conséquences du postulat de symétrisation. Le chapitre 3 est consacré à l'exposé élémentaire de la description des atomes à plusieurs électrons, en insistant sur les aspects physiques, et notamment en donnant l'explication de l'existence du magnétisme atomique, laquelle constitue un exemple peu cité des succès de la Mécanique Quantique. Le chapitre 4 termine la partie consacrée à la physique atomique ; il se veut une illustration des postulats quantiques en discutant quelques expériences assez récentes ayant fourni, de façon parfois spectaculaire et très inattendue, de nouvelles preuves de l'extraordinaire robustesse de la théorie quantique et ayant confirmé avec éclat ses prévisions, y compris dans ce qu'elles ont de plus "paradoxal" pour le sens commun (effet Zénon, sauts quantiques, cryptographie quantique).

La deuxième partie (3 cours environ) est consacré à la physique moléculaire. Le chapitre 5 expose l'approximation de Born - Oppenheimer qui est la base de la description quantique des molécules. Le chapitre 6 jette les bases physiques de la notion de liaison chimique, cependant que le chapitre 7 est consacré à la présentation élémentaire des principes de la spectroscopie moléculaire.

La troisième et dernière partie (6 à 7 cours) est consacrée à quelques aspects de la physique de la matière condensée ordinaire. Ce domaine est trop vaste pour ne pas exiger d'emblée des choix draconiens ; en particulier, la physique des verres (et des liquides) ne sera pas abordée, et l'objectif est de montrer comment la Mécanique quantique permet de comprendre les propriétés des cristaux, idéalement décrits dans l'hypothèse du solide parfaitement ordonné selon un réseau sans défauts. Une discussion semi-quantitative de la cohésion des cristaux est proposée dans le chapitre 8, qui permet de justifier la classification élémentaire des différents types de solides. Le chapitre 9 est consacré aux conséquences de l'ordre spatial parfait sur les propriétés électroniques ; le théorème de Bloch y est donné, ainsi que ses premières applications (approximation des électrons presque libres et approximation des liaisons fortes). Dans un cas comme dans l'autre, l'accent est à nouveau mis sur les vertus explicatives de la Mécanique quantique, seule susceptible de fournir un cadre cohérent à une distinction aussi primordiale que la distinction isolant/conducteur. Dans le chapitre 10, on abandonne l'hypothèse du réseau rigide en exposant le traitement élémentaire des vibrations de réseau, et les premières conséquences de l'existence des phonons sur la chaleur spécifique et les spectres de diffraction. Enfin, le dernier chapitre (ch. 11) est une simple introduction au transport dans les métaux, afin de montrer notamment comment le modèle classique de Drude trouve naturellement son prolongement dans le calcul quantique semi-classique de la conductivité électrique.

Le volume et le contenu de ces notes de cours dépassent – et de loin – ce qu'il est raisonnable et possible de traiter en un cours semestriel. Cette disproportion est le résultat d'un choix délibéré :

tenter de faire le tour d'horizon d'une question donnée – sans toutefois prétendre l'épuiser – en posant les idées principales mais sans omettre des compléments conceptuels et sans négliger les détails de calcul. En outre, et c'est tout particulièrement vrai pour le chapitre 1 qui est le point d'articulation avec le cours d'Édouard BREZIN, les recouvrements apparents ne seront pas repris, et ne figurent explicitement que pour rappeler les idées principales, fixer les notations et fournir des compléments utiles pour la compréhension. Enfin, certains points ne seront pas du tout abordés oralement et doivent être considérés comme des prolongements naturels des questions en cours de discussion, laissés à la libre appréciation du lecteur.

Ces notes n'ont aucun caractère exhaustif et ne doivent surtout pas être considérées comme un document de référence exclusif, bien au contraire : il est souhaitable (et souhaité !) que les développements qui ne seront pas mentionnés en amphi suscitent la curiosité d'en savoir davantage en allant consulter les nombreux livres traitant des questions abordées. Quelques références sont données à l'issue de chaque chapitre : elles sont seulement indicatives et reflètent essentiellement des goûts personnels.

Table des Matières

1	Atome d'hydrogène et compléments	1
1.1	Hamiltonien électrostatique	1
1.2	États propres liés	4
1.2.1	Comportements de la fonction radiale	7
1.2.2	Résolution de l'équation radiale	9
1.3	Symétrie dynamique du potentiel Coulombien	17
1.4	Interaction spin-orbite et structure fine	19
1.5	Équation de Dirac et limite de Pauli	23
1.5.1	Construction de l'équation de Dirac	23
1.5.2	États stationnaires d'un électron libre	29
1.6	Effet d'un champ statique externe : effet Stark et effet Zeeman	35
1.7	Interaction avec un champ électromagnétique (description semi-classique)	39
1.8	États de diffusion	42
1.8.1	Nature des états de diffusion	42
1.8.2	Méthode des déphasages	44
1.8.3	Le cas du potentiel Coulombien	47
2	Particules identiques	49
2.1	Indiscernabilité des particules identiques en Mécanique Quantique	49
2.2	Le postulat de symétrisation	52
2.3	Permutations. Opérateurs de symétrisation et d'antisymétrisation	56
2.4	Etats d'un système de particules indépendantes : différence fondamentale entre bosons et fermions	61
2.5	Introduction à la Seconde Quantification	66

3	Atomes à plusieurs électrons	73
3.1	Modèle à électrons indépendants	73
3.2	Termes spectraux. Multiplets	75
3.3	Théorème du Viriel	79
3.4	Au-delà de l'approximation à électrons indépendants	80
3.5	L'atome d'hélium	84
4	Illustration des postulats de la Mécanique Quantique	91
4.1	L'effet Zenon quantique	91
4.2	Sauts quantiques	102
4.3	Cryptographie quantique	110
4.3.1	Principes de la détection d'une écoute indésirable	111
4.3.2	Exemples	112
4.3.3	Communication de la clé secrète entre Alice et Bob	115
5	L'approximation de Born et Oppenheimer	121
5.1	Ordres de grandeur	121
5.2	Approximation de Born et Oppenheimer	123
5.2.1	Étape 1 : mouvement des électrons dans le champ des noyaux fixes	124
5.2.2	Étape 2 : inclusion du mouvement des noyaux	124
5.2.3	Approximation adiabatique	125
5.3	Discussion	126
5.3.1	À propos du mouvement des électrons	126
5.3.2	À propos du mouvement des noyaux	128
6	Structure électronique des molécules. Nature physique de la liaison chimique	131
6.1	L'ion moléculaire H_2^+	131
6.1.1	Solution exacte	132
6.1.2	Méthodologie pour une description approchée	135
6.2	Nature physique de la liaison chimique	139

7	Mouvement des noyaux.	
	Spectres de vibration et de rotation	141
7.1	Molécules diatomiques	142
7.1.1	Fonctions propres de vibration - rotation	142
7.1.2	Spectres de vibration et de rotation	146
7.2	Effet Raman	148
7.3	Molécules polyatomiques	149
7.3.1	Coordonnées normales	149
7.3.2	Rotation des molécules polyatomiques	152
8	Matière condensée ordonnée	157
8.1	Classification des solides ordonnés	157
8.2	Énergie de cohésion des solides ordonnés	158
8.2.1	Généralités	158
8.2.2	Cohésion des solides moléculaires	159
8.2.3	Cohésion des réseaux ioniques	163
8.2.4	Cohésion des solides covalents	165
8.2.5	Cohésion des métaux	166
8.3	Structures cristallines	169
8.3.1	Réseau de Bravais	169
8.3.2	Structure physique d'un réseau de Bravais	172
8.4	Réseau réciproque	173
8.4.1	Définition et propriétés du réseau réciproque	173
8.4.2	Exemples de réseaux réciproques	176
8.4.3	Plans réticulaires	178
8.5	Diffraction par un réseau	179
8.5.1	Diffusion des rayons X et des neutrons	179
8.5.2	Conditions de von Laue et de Bragg	183
8.5.3	Facteurs de structure (géométrique et atomique)	186

9	Electrons dans un cristal	191
9.1	Préliminaires	191
9.2	Théorème de Bloch	192
9.3	Premières conséquences du théorème de Bloch	198
9.3.1	Première zone de Brillouin	198
9.3.2	Équation pour la fonction $u_{\vec{k}}$	199
9.3.3	Impulsion de l'électron dans le cristal	200
9.3.4	Vitesse d'un électron de Bloch	201
9.3.5	Symétrie de la fonction $\varepsilon(\vec{k})$	202
9.3.6	Surface de Fermi	203
9.4	Applications choisies du théorème de Bloch	204
9.4.1	L'approximation des électrons presque libres	204
9.4.2	L'approximation des liaisons fortes	214
9.5	Conducteurs, semi-conducteurs, isolants	219
10	Vibrations d'un solide ordonné	225
10.1	L'approximation harmonique	225
10.2	Modes normaux d'un réseau unidimensionnel	227
10.3	Modes normaux d'un réseau unidimensionnel à deux atomes par maille	233
10.4	Modes normaux d'un réseau tridimensionnel	239
10.4.1	Réseau sans base	239
10.4.2	Réseau avec base	240
10.5	Quantification des vibrations de réseau	240
10.6	Contribution des phonons à la chaleur spécifique	241
10.7	Manifestation des phonons dans les spectres de diffraction	243
11	Notions de transport dans les solides	249
11.1	Généralités	249
11.2	Le modèle semi-classique	250
11.3	Conséquences du modèle semi-classique	252
11.3.1	Mouvement dans un champ électrique constant dans le temps	253
11.3.2	Mouvement dans un champ magnétique constant	254
11.4	Conductivité statique d'un métal	256

Partie I

Physique atomique

Chapitre 1

Atome d'hydrogène et compléments

1.1 Hamiltonien électrostatique

En première approximation, l'électron de l'atome d'hydrogène est soumis au seul champ électrostatique du noyau, à symétrie sphérique et dérivant du potentiel¹ $U(r) = \frac{|e|}{(4\pi\epsilon_0 r)}$. Une fois effectuée la réduction du problème à deux corps, le Hamiltonien décrivant le mouvement relatif est celui d'une particule de masse réduite μ et d'énergie potentielle $V(r) = eU(r)$:

$$H = \frac{\vec{p}^2}{2\mu} + V(r) \equiv \frac{\vec{p}^2}{2\mu} - \frac{e^2}{r} \quad (e^2 = \frac{e^2}{4\pi\epsilon_0}) . \quad (1.1)$$

En représentation- q , $\vec{p} = -i\hbar\vec{\nabla}$, de sorte que le terme cinétique T fait intervenir l'opérateur Laplacien Δ . Compte tenu de la symétrie, le choix naturel consiste à adopter les coordonnées sphériques² (r, θ, ϕ) . L'expression correspondante de T peut s'obtenir en effectuant le changement de variables, facile mais laborieux, dans le Laplacien. Le résultat est le suivant :

$$T \equiv \frac{\vec{p}^2}{2\mu} = -\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{\vec{L}^2}{2\mu r^2} , \quad (1.2)$$

où³ $\vec{L} = \vec{r} \times \vec{p}$. Le terme $\frac{\vec{L}^2}{2\mu r^2}$ est appelé terme centrifuge : il apparaît dans l'équation fondamentale de la dynamique pour la variable radiale r et, reporté au second membre, ressort comme une force d'inertie complémentaire ; comme il s'agit de rotation, c'est un terme centrifuge. L'expression (1.2) de l'énergie cinétique n'est pas surprenante et peut d'ailleurs être retrouvée par des moyens moins lourds. En effet, classiquement, on peut écrire :

$$\vec{L}^2 = (\vec{r} \times \vec{p})^2 = r^2 p^2 \sin^2 \Theta = r^2 p^2 (1 - \cos^2 \Theta) \equiv r^2 \vec{p}^2 - (\vec{r} \cdot \vec{p})^2 , \quad (1.3)$$

Θ étant l'angle entre \vec{r} et \vec{p} . Il en résulte :

$$\vec{p}^2 = \left(\frac{\vec{r} \cdot \vec{p}}{r} \right)^2 + \frac{\vec{L}^2}{r^2} . \quad (1.4)$$

Cette expression classique doit être convenablement symétrisée afin d'engendrer *in fine* un opérateur hermitique après la substitution $\vec{p} \rightarrow -i\hbar\vec{\nabla}$. Il n'y a pas d'ambiguïté pour le rapport $\frac{\vec{L}^2}{r^2}$ car les deux opérateurs commutent

¹On désignera partout par e la charge de l'électron ($e = -1.6 \times 10^{-19}$ C).

² $0 \leq r < +\infty$, $0 \leq \theta < \pi$, $0 \leq \phi < 2\pi$.

³On vérifie facilement, à partir de $[q_u, p_v] = i\hbar\delta_{uv}$, que $\vec{r} \times \vec{p} = -\vec{p} \times \vec{r}$; il est donc inutile de symétriser l'expression classique avant de passer aux opérateurs.

entre eux. La seule difficulté réside dans le terme p_r^2 :

$$p_r^2 \equiv \left(\frac{\vec{r} \cdot \vec{p}}{r} \right)^2 \iff p_r = \frac{\vec{r} \cdot \vec{p}}{r} , \quad (1.5)$$

dont la symétrisation “naturelle” est :

$$p_r = \frac{1}{2} \left(\frac{\vec{r}}{r} \cdot \vec{p} + \vec{p} \cdot \frac{\vec{r}}{r} \right) . \quad (1.6)$$

En vertu de (1.7), on peut écrire⁴ :

$$\vec{p} \cdot \frac{\vec{r}}{r} = [\vec{p}, \frac{\vec{r}}{r}] + \frac{\vec{r}}{r} \cdot \vec{p} = -i\hbar \left(\vec{\nabla} \cdot \frac{\vec{r}}{r} \right) + \frac{\vec{r}}{r} \cdot \vec{p} . \quad (1.8)$$

L'expression symétrisée (1.6) conduit alors à :

$$p_r = \frac{\vec{r}}{r} \cdot \vec{p} - \frac{i\hbar}{2} \left(\vec{\nabla} \cdot \frac{\vec{r}}{r} \right) = -i\hbar \left[\frac{\vec{r}}{r} \cdot \vec{\nabla} + \frac{1}{2} \left(\vec{\nabla} \cdot \frac{\vec{r}}{r} \right) \right] . \quad (1.9)$$

Le premier terme entre crochets contient la composante du gradient le long du rayon-vecteur : c'est simplement $\frac{\partial}{\partial r}$. Le deuxième terme se calcule facilement à partir de sa définition :

$$\left(\vec{\nabla} \cdot \frac{\vec{r}}{r} \right) = \frac{\partial}{\partial x} \frac{x}{r} + \frac{\partial}{\partial y} \frac{y}{r} + \frac{\partial}{\partial z} \frac{z}{r} . \quad (1.10)$$

Le terme en x est :

$$\frac{\partial}{\partial x} \frac{x}{r} = \frac{1}{r} - \frac{x^2}{r^3} . \quad (1.11)$$

En additionnant les trois termes en x , y et z , il vient donc :

$$\left(\vec{\nabla} \cdot \frac{\vec{r}}{r} \right) = \frac{3}{r} - \frac{x^2 + y^2 + z^2}{r^3} = \frac{2}{r} , \quad (1.12)$$

d'où, selon (1.9) :

$$p_r = -i\hbar \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \equiv -i\hbar \frac{1}{r} \frac{\partial}{\partial r} r . \quad (1.13)$$

On remarque au passage que p_r n'est pas égal à $-i\hbar \frac{\partial}{\partial r}$ en conséquence du fait que r n'est pas une coordonnée rectangulaire. Avec l'expression (1.13), on trouve toujours $[r, p_r] = i\hbar$, mais cette relation n'a rien d'évident *a priori* puisque précisément les coordonnées ne sont pas cartésiennes⁵.

Il reste à calculer le carré de p_r ; en prenant garde à l'ordre des opérateurs dans l'élévation au carré du binôme :

$$\left(\frac{\partial}{\partial r} + \frac{1}{r} \right)^2 \bullet = \left(\frac{\partial}{\partial r} \frac{\partial}{\partial r} + \frac{\partial}{\partial r} \frac{1}{r} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \right) \bullet , \quad (1.14)$$

on trouve⁶ :

$$p_r^2 = -\hbar^2 \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) \equiv -\hbar^2 \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) . \quad (1.16)$$

⁴Pour toute fonction vectorielle $\vec{f}(\vec{r})$ on a :

$$[\vec{p}, \vec{f}(\vec{r})] = -i\hbar \left(\vec{\nabla} \cdot \vec{f}(\vec{r}) \right) \mathbf{1} ; \quad (1.7)$$

les parenthèses sont là pour signifier que l'opérateur gradient, au second membre, n'agit que sur $\vec{f}(\vec{r})$, pas sur une fonction sous-entendue à droite.

⁵toute expression du genre $\varpi(r) \equiv -i\hbar \left[\frac{\partial}{\partial r} + \phi(r) \right]$ conduit d'ailleurs elle aussi à la relation de commutation $[r, \varpi(r)] = i\hbar$.

⁶D'une façon générale, dans \mathbb{R}^d , p_r^2 est donné par :

$$p_r^2 = -\hbar^2 \left(\frac{\partial^2}{\partial r^2} + \frac{d-1}{r} \frac{\partial}{\partial r} + \frac{(d-1)(d-3)}{4r^2} \right) . \quad (1.15)$$

En ajoutant $\frac{\vec{L}^2}{2\mu r^2}$, l'expression (1.2) est retrouvée.

En définitive, le Hamiltonien du problème central dans \mathbb{R}^3 s'écrit :

$$H = \frac{p_r^2}{2\mu} + \frac{\vec{L}^2}{2\mu r^2} + V(r) \equiv -\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{\vec{L}^2}{2\mu r^2} + V(r) \equiv T + V, \quad (1.17)$$

où chaque opérateur différentiel agit sur *tout* ce qui se trouve à sa droite. T est l'opérateur représentant l'énergie cinétique de la particule.

Pour la référence ultérieure, il est utile de préciser les conditions à satisfaire pour que cet opérateur soit hermitique. Le carré du moment cinétique commute avec r , donc le second terme de (1.17) est visiblement hermitique ; $V(r)$ l'est aussi. La seule question porte sur le premier terme, qui sera hermitique si p_r l'est. Il s'agit d'un opérateur différentiel, il faut donc que toutes les fonctions sur lequel il agit se comportent convenablement aux bornes $r = 0$ et $r = +\infty$ (la condition d'hermiticité se vérifie par des intégrations par parties et il faut bien que les termes tout intégrés soient nuls).

Soit $\psi_1(r)$ et $\psi_2(r)$ deux fonctions quelconques ; la condition d'hermiticité de p_r s'écrit par définition :

$$\langle \psi_1 | p_r \psi_2 \rangle = \langle p_r \psi_1 | \psi_2 \rangle \iff \int_0^{+\infty} r^2 dr \psi_1^*(r) [p_r \psi_2(r)] = \int_0^{+\infty} r^2 dr [p_r \psi_1(r)]^* \psi_2(r) \quad (1.18)$$

soit, explicitement compte tenu de (1.13) :

$$-i\hbar \int_0^{+\infty} r^2 dr \psi_1^*(r) \left[\left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \psi_2(r) \right] = +i\hbar \int_0^{+\infty} r^2 dr \left[\left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \psi_1^*(r) \right] \psi_2(r). \quad (1.19)$$

Le premier membre est :

$$-i\hbar \int_0^{+\infty} r^2 dr \psi_1^*(r) [\psi_2'(r) + r^{-1}\psi_2(r)]. \quad (1.20)$$

Le second membre se transforme par intégration par parties ; on trouve :

$$+i\hbar [r^2 \psi_1^*(r) \psi_2(r)]_0^{+\infty} - i\hbar \int_0^{+\infty} dr \psi_1^*(r) [r \psi_2(r) + r^2 \psi_2'(r)]. \quad (1.21)$$

En comparant (1.20) et (1.21), on voit que la condition d'hermiticité (1.19) s'écrit :

$$[r^2 \psi_1^*(r) \psi_2(r)]_0^{+\infty} = 0. \quad (1.22)$$

En particulier, pour tout élément diagonal, il faut :

$$[r^2 |\psi(r)|^2]_0^{+\infty} = 0. \quad (1.23)$$

Pour que ces conditions soient satisfaites, il est *suffisant* que toutes les fonctions de l'espace vectoriel satisfassent les égalités suivantes :

$$\lim_{r \rightarrow 0} [r |\psi(r)|] = 0, \quad \lim_{r \rightarrow +\infty} [r |\psi(r)|] = 0. \quad (1.24)$$

Ces conditions suffisantes, notamment celle à l'infini, sont en général satisfaites pour les états liés⁷, bien que la condition de normalisabilité, à elle seule, ne suffise pas à assurer l'hermiticité de p_r . En effet, si ψ se comporte comme $\frac{1}{r}$ à l'origine, $r^2 |\psi|^2$ est sommable en zéro et pourtant dans ce cas $r\psi(r)$ tend vers une constante *finie* à l'origine, en violation de (1.24)

⁷Pour un état lié, donc normalisable, il est *nécessaire* que $\lim_{r \rightarrow +\infty} [r^2 |\psi(r)|^2] = 0$.

Remarque

L'hermiticité de p_r assure celle de H . Toutefois, p_r n'est pas une observable au sens strict⁸ : ses fonctions propres n'appartiennent pas à l'espace de fonctions qui sont à la fois de carré sommable et telles que $[r\psi(r)]_{r=0} = 0$. En effet, d'après (1.13) l'équation propre pour p_r est :

$$-i\hbar \frac{1}{r} \frac{\partial}{\partial r} [rf(r)] = \lambda f(r) . \quad (1.25)$$

En multipliant membre à membre par r , on voit que les solutions sont :

$$f(r) = \frac{C^{\text{ste}}}{r} e^{\frac{i}{\hbar}\lambda r} . \quad (1.26)$$

Ces fonctions ne sont pas de carré sommable et ne satisfont pas la condition suffisante assurant que p_r est hermitique. On ne peut donc faire jouer les postulats de la Mécanique Quantique à propos de p_r , qui n'est pas une observable puisque ses états propres sont en-dehors de l'espace sur lequel cet opérateur est hermitique⁹. De toute façon, c'est l'hermiticité de p_r^2 qui est nécessaire (et suffisante) ; il n'est d'ailleurs pas évident que l'hermiticité de p_r^2 exige celle de p_r – étant entendu que si un opérateur est hermitique, son carré l'est aussi.

1.2 États propres liés

Comme $\lim_{r \rightarrow +\infty} V(r) = 0$, les états liés ont une énergie négative. Classiquement, la particule reste confinée entre deux valeurs r_- et r_+ (la trajectoire est alors plus précisément une ellipse) ; quantiquement, cet état lié est représenté par une fonction d'onde qui décroît essentiellement comme une exponentielle quand r tend vers l'infini.

La résolution du problème aux valeurs et fonctions propres consiste donc à trouver les fonctions $\psi(\vec{r})$ satisfaisant l'équation aux dérivées partielles résultant de (1.1) et (1.2) :

$$H \psi(\vec{r}) = E \psi(\vec{r}) , \quad (1.27)$$

avec la condition $E < 0$. Comme on le verra, cette condition d'apparence banale contient en germe la quantification spontanée de l'énergie des états liés.

H possède la symétrie sphérique et de ce fait commute avec n'importe laquelle des composantes du moment cinétique orbital :

$$[H, \vec{L}] = 0 , \quad (1.28)$$

Techniquement, ceci résulte d'une part du fait que \vec{L} et T commutent, puisque l'ordre des deux opérations : rotation, dérivation par rapport à r , est indifférent (géométriquement, l'indifférence à l'ordre est évidente). D'autre part, \vec{L} commute avec $V(r)$: quand on fait tourner le système, par définition d'un champ central, l'énergie potentielle ne change pas. L'équation (1.28) assure que l'on peut trouver des états propres communs à (H, \vec{L}^2, L_z) . Si l'on note $\psi(r, \theta, \phi)$ ces derniers, on voit que, compte tenu de la forme¹⁰ de H , tous les états propres de ce dernier opérateur peuvent s'obtenir comme des combinaisons linéaires¹¹ de fonctions à variables séparées :

$$\psi(r, \theta, \phi) = R(r) Y_{lm}(\theta, \phi) , \quad (1.29)$$

⁸voir [1], p. 293.

⁹Cette propriété de p_r n'est pas si particulière que cela : p_x n'est pas non plus de ce point de vue une observable puisque les ondes planes ne sont pas normalisables. Il n'empêche que la condition d'hermiticité est tout autant requise pour p_x que pour p_r .

¹⁰ H est une combinaison linéaire de deux opérateurs – l'un radial, l'autre angulaire – où les coefficients multiplicatifs dépendent au plus des variables autres que celles concernées par l'opérateur.

¹¹En raison de la symétrie de rotation autour notamment de Oz , tous les états propres du type (1.29), de même l et différant par leur m , ont la même énergie. Toute combinaison linéaire de ces états est donc encore propre, avec la même énergie.

où les Y_{lm} sont les harmoniques sphériques introduites antérieurement, lors de l'exposé de la théorie du moment cinétique (orbital), telles que $\vec{L}^2 Y_{lm} = \hbar^2 l(l+1) Y_{lm}$, $L_z Y_{lm} = m\hbar Y_{lm}$. En effet, reportons une telle forme dans $H\psi \equiv [T + V]\psi = E\psi$; il vient :

$$\left[-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{\vec{L}^2}{2\mu r^2} + V(r) \right] R(r) Y_{lm}(\theta, \phi) = E R(r) Y_{lm}(\theta, \phi) . \quad (1.30)$$

En faisant agir le carré du moment cinétique \vec{L}^2 sur ses états propres Y_{lm} , et après simplification¹², il reste l'équation dite radiale pour la seule fonction $R(r)$:

$$-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) R(r) + \frac{\hbar^2 l(l+1)}{2\mu r^2} R(r) + V(r) R(r) = E R(r) . \quad (1.31)$$

Il est souvent avantageux d'introduire une autre fonction inconnue $u(r) = rR(r)$. On voit sans peine à partir de (1.31) que $u(r)$ satisfait l'équation suivante :

$$-\frac{\hbar^2}{2\mu} \frac{d^2 u}{dr^2} + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] u(r) \equiv -\frac{\hbar^2}{2\mu} \frac{d^2 u}{dr^2} + V_{\text{eff}}(r) u(r) = E u(r) . \quad (1.32)$$

Cette équation, nettement plus simple que (1.31), présente de plus l'avantage d'être formellement un problème à une dimension (mais réduite à \mathbb{R}_+) avec un potentiel effectif incorporant le terme centrifuge (ce dernier est nul pour un moment cinétique nul¹³, ce qui n'est pas surprenant). Ainsi, le problème pour $u(r)$ est exactement le même que celui d'une particule à une dimension, d'énergie potentielle $V_{\text{eff}}(x)$ et en outre confinée sur le demi-axe réel positif par une barrière infranchissable en $x = 0$. Tout ce que l'on sait à propos du mouvement à une dimension (réalité des fonctions d'onde, dégénérescence, *etc.*) peut ainsi être utilisé, tant que la fonction $u(r)$ est seule pertinente.

Lorsque l'énergie E est positive¹⁴, le mouvement classique à la Kepler (hyperbole – parabole si $E = 0$) n'est pas borné ; le mouvement quantique correspondant est décrit par une fonction d'onde qui oscille à l'infini et représente un état non-lié. Dans le cas contraire, $E < 0$, la particule classique reste confinée entre deux valeurs r_- et r_+ (ellipse, ou cercle quand $r_- = r_+$) : l'équivalent quantique est un état lié dont la fonction d'onde décroît essentiellement comme une exponentielle quand r tend vers l'infini. Toutefois, les conditions aux limites pour la fonction radiale sont différentes de celles rencontrées dans les problèmes à une dimension, et peuvent dépendre du potentiel $V(r)$ considéré. Elles seront précisées au coup par coup, dans chaque cas particulier étudié dans la suite. En tout état de cause, la condition d'hermiticité de p_r – qui assure celle de H – s'écrit :

$$\lim_{r \rightarrow 0} [rR(r)] = 0 \iff \lim_{r \rightarrow 0} [u(r)] = 0 , \quad \lim_{r \rightarrow \infty} [rR(r)] = 0 \iff \lim_{r \rightarrow \infty} [u(r)] = 0 . \quad (1.33)$$

La sous-section 1.2.1 donne une analyse détaillée du comportement de la fonction radiale en $r = 0$, $r = +\infty$ et au voisinage d'un saut de potentiel.

Notons enfin que les Y_{lm} étant toujours conventionnellement normalisées par rapport aux variables angulaires, la normalisation de la fonction d'onde d'un état lié impose :

$$\int_0^{+\infty} r^2 dr R^2(r) = 1 \iff \int_0^{+\infty} dr u^2(r) = 1 . \quad (1.34)$$

L'analyse de l'équation de conservation :

$$\frac{\partial \rho}{\partial t} + \text{div } \vec{j} = 0 , \quad \rho = \Psi^* \Psi , \quad \vec{j} = \frac{\hbar}{2i\mu} [\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^*] \quad (1.35)$$

¹²L'action explicite de \vec{L}^2 fait apparaître le simple produit $\hbar^2 l(l+1) Y_{lm}$, de sorte que l'harmonique sphérique se trouve en facteur dans l'équation complète et peut être omise dans la suite.

¹³Les états $l = 0, 1, 2, 3, \dots$ sont traditionnellement notés s, p, d, f, ...

¹⁴On suppose que $\lim_{r \rightarrow +\infty} V(r) = 0$. Plus généralement – sauf cas exceptionnel –, le seuil en énergie séparant états liés et non-liés est la valeur de cette limite ; si elle est infinie (par exemple : oscillateur harmonique à trois dimensions), tous les états sont liés. On connaît quelques cas exotiques où il existe des états liés dont l'énergie est noyée dans le *continuum* d'états de diffusion.

permet de réaliser que la fonction radiale R est essentiellement réelle. En effet, quand Ψ est un état stationnaire (donc de la forme $e^{\frac{i}{\hbar}Et} RY_{lm}$), la densité ρ est constante en temps et il reste $\operatorname{div} \vec{j} = 0$. Comme les angles sont découplés de r et sans interaction entre eux, les trois composantes j_r , j_θ et j_ϕ sont séparément des constantes dans l'espace ; la valeur de ces constantes dépend de la nature de l'état, lié ou non.

Pour un état lié, seule la composante j_ϕ peut être différente de zéro (elle ne l'est pas forcément). La raison à ceci est d'ordre purement topologique : l'angle ϕ est la seule coordonnée qui se "boucle" dans le système, permettant des courants permanents non nuls. Au contraire, l'angle θ (qui varie entre 0 et π) n'autorise pas de circulation permanente, et il en va de même pour la coordonnée radiale r , qui va de 0 à $+\infty$. En particulier, la composante radiale j_r , nulle pour un état lié, a pour expression :

$$j_r = \frac{\hbar}{2i\mu} |Y_{lm}|^2 \left[R^* \frac{dR}{dr} - R \frac{dR^*}{dr} \right] \quad (1.36)$$

et il en résulte que :

$$R^* \frac{dR}{dr} = R \frac{dR^*}{dr} \iff \ln R = \ln R^* + C^{\text{ste}} \iff R^* \propto R . \quad (1.37)$$

R et R^* sont deux fonctions proportionnelles et représentent donc le même état physique. La partie radiale d'un état propre lié peut en conséquence toujours être prise réelle.

La réalité de R peut se voir d'une autre façon, en invoquant l'absence de dégénérescence et la symétrie par renversement du temps¹⁵. La fonction $u(r)$ introduite ci-dessus satisfait de fait l'équation à une dimension sur \mathbb{R}_+ (1.32) ; partant de là, on peut établir un théorème du Wronskien, comme on le fait pour les problèmes à une dimension sur \mathbb{R} . On trouve ainsi :

$$[W(u_1, u_2)]_a^b \equiv [u_1 u_2' - u_2 u_1']_a^b = (E_1 - E_2) \int_a^b dr u_1(r) u_2(r) . \quad (1.38)$$

En particulier, si u_1 et u_2 sont associées à la même valeur propre E ($E_1 = E_2 = E$), la variation du Wronskien entre a et b est nulle :

$$[W(u_1, u_2)]_a^b = 0 , \quad (1.39)$$

et puisque les deux valeurs a et b sont quelconques, on en déduit :

$$W(u_1, u_2) = C^{\text{ste}} . \quad (1.40)$$

Ceci permet d'établir que les états discrets (*i.e.* normalisables) $u(r)$ sont non-dégénérés. En effet, pour de tels états, les fonctions u_i et u_i' sont nécessairement nulles à l'infini, le Wronskien est donc nul à l'infini. Comme il prend la même valeur partout, il est nul partout. Il vient ainsi, $\forall r$:

$$u_1(r) u_2'(r) - u_2(r) u_1'(r) = 0 \iff u_2(r) \propto u_1(r) . \quad (1.41)$$

Ainsi, deux fonctions u_1 et u_2 associées à la même valeur propre E sont proportionnelles : à chaque E correspond de fait une seule fonction propre u ¹⁶. Compte tenu de la symétrie par renversement du temps, u et u^* correspondent à une et une seule valeur propre ; deux telles fonctions diffèrent donc au plus par une phase globale, dénuée de sens physique. En définitive, la fonction radiale R peut bien toujours être prise réelle.

¹⁵Cette même symétrie permet aussi de se convaincre que Y_{lm} et Y_{l-m} ont la même énergie – indépendamment de la symétrie de rotation autour de Oz . L'application d'un champ magnétique brise l'invariance par renversement du temps (les ampériens changent de sens).

¹⁶ H reste bien sûr en général dégénéré : n et l étant fixés, tous les états propres distincts de L_z ont la même énergie en l'absence de direction privilégiée. En outre, le phénomène de dégénérescence "accidentelle" (E ne dépendant en fait que de n) reste toujours possible. Pour le champ Coulombien, la dégénérescence "accidentelle" donne la même énergie à toutes les fonctions radiales de même n : R_{nl} et $R_{n'l'}$ ont la même énergie. Il n'y a pas de théorème du Wronskien pour un tel couple : la démonstration ci-dessus suppose que les deux fonctions u_1 et u_2 sont propres du même Hamiltonien ; or le Hamiltonien central (1.32) dépend de l par le terme centrifuge.

1.2.1 Comportements de la fonction radiale

La discussion précédente montre que la fonction d'onde du champ central peut être recherchée sous la forme ($R(r) = \frac{u(r)}{r}$) :

$$\psi(r, \theta, \phi) = \frac{u(r)}{r} Y_{lm}(\theta, \phi) , \quad (1.42)$$

où la seule fonction encore inconnue $u(r)$ satisfait l'équation différentielle :

$$-\frac{\hbar^2}{2\mu} \frac{d^2 u}{dr^2} + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] u(r) = E u(r) . \quad (1.43)$$

Comme r ne change pas dans l'inversion d'espace, la fonction ψ a la même parité que Y_{lm} , soit $(-1)^l$. Les états propres sont donc pairs (resp. impairs) si l est un entier pair (resp. impair).

Tout état lié doit être normalisable ; les Y_{lm} étant supposées normalisées à part, il convient seulement d'assurer que $u(r)$ est une fonction de module carré sommable selon (1.34), ce qui élimine une grande partie des solutions mathématiques de l'équation différentielle (1.43). En outre, cette équation est vraie partout sauf évidemment en $r = 0$; la singularité du point $r = 0$ doit donc faire l'objet d'une prescription spéciale, qui a déjà été établie en arguant du fait que p_r est hermitique (voir (1.33)), ce qui assure que le Hamiltonien l'est aussi.

La discussion générale des conditions que doit satisfaire toute solution $u(r)$ physiquement acceptable est sans intérêt ; pour les potentiels possédant à la fois des états liés et des états non-liés¹⁷, l'immense majorité des cas pratiques est couverte quand on fait les hypothèses additionnelles suivantes sur la fonction $V(r)$:

$$V(r) \rightarrow 0 \quad \text{si} \quad r \rightarrow +\infty , \quad (1.44)$$

$$V(r) \simeq A r^\alpha \quad \text{si} \quad r \rightarrow 0 , \quad (1.45)$$

où $\alpha \geq -1$. La discussion couvre donc des potentiels divergents à l'origine et en particulier le cas du champ coulombien.

Examinons d'abord précisément ce qui se passe à l'origine, lorsque l'hypothèse (1.45) est satisfaite. Partant de l'équation pour u , on l'intègre dans un voisinage à droite :

$$\int_0^{\delta r} dr \left[-\frac{\hbar^2}{2\mu} \frac{d^2 u}{dr^2} + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] u(r) \right] = E \int_0^{\delta r} dr u(r) . \quad (1.46)$$

Le second membre sera nul à la limite $\delta r \rightarrow 0$ et on peut d'ores et déjà l'annuler :

$$-\frac{\hbar^2}{2\mu} \left[\frac{du}{dr} \right]_0^{\delta r} + \int_0^{\delta r} dr \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + A r^\alpha \right] u(r) = 0 . \quad (1.47)$$

Faisons maintenant l'hypothèse que u se comporte comme r^β dans le voisinage de l'origine (comme u doit tendre vers zéro, ceci impose $\beta > 0$). Il en résulte :

$$-\frac{\hbar^2}{2\mu} [\beta r^{\beta-1}]_0^{\delta r} + \int_0^{\delta r} dr \left[\frac{\hbar^2 l(l+1)}{2\mu} r^{\beta-2} + A r^{\alpha+\beta} \right] = 0 . \quad (1.48)$$

À condition que β soit différent de 1, il vient, après intégration et regroupement des termes :

$$-\frac{\hbar^2}{2\mu} \left[\left(\beta - \frac{l(l+1)}{\beta-1} \right) r^{\beta-1} \right]_0^{\delta r} + \left[\frac{A}{\alpha+\beta+1} r^{\alpha+\beta+1} \right]_0^{\delta r} = 0 . \quad (1.49)$$

Comme $\alpha+1 \geq 0$ et $\beta > 0$, le second terme tend vers zéro dans tous les cas : il faut donc $\beta > 1$, auquel cas R est en $r^{\beta-1}$ et tend vers zéro. Dans le cas où $\beta = 1$, $u \sim Cr$, l'intégration dans (1.49) fournit un terme logarithmique ; en effet, (1.48) donne alors :

$$\int_0^{\delta r} dr \left[\frac{\hbar^2}{2\mu} l(l+1) r^{-1} + A r^{\alpha+1} \right] = 0 \iff \left[\frac{\hbar^2}{2\mu} l(l+1) \ln r + \frac{A}{\alpha+2} r^{\alpha+2} \right]_0^{\delta r} = 0 . \quad (1.50)$$

¹⁷L'existence d'états non-liés est assurée par le fait que $V(r)$ a une limite finie quand $r \rightarrow +\infty$. Ce n'est pas le cas pour l'oscillateur harmonique à trois dimensions qui, comme son homologue sur \mathbb{R} , ne possède que des états liés.

Il est toujours impossible de satisfaire ceci, à cause du terme divergeant logarithmiquement, sauf si $l = 0$. Donc, seuls les états s peuvent se comporter comme Cr (pour u) et comme une constante (pour R).

L'analyse précédente repose sur une intégration locale et est donc sûre ; elle montre bien que u doit tendre vers zéro à l'origine – mais ne permet pas de trouver l'exposant β . Pour l'obtenir, il suffit de manipuler l'équation différentielle elle-même, avec toujours l'hypothèse $u \propto r^\beta$. Il vient :

$$-\frac{\hbar^2}{2\mu} \beta(\beta - 1) r^{\beta-2} + \frac{\hbar^2 l(l+1)}{2\mu} r^{\beta-2} + A r^{\alpha+\beta} = E r^\beta ; \quad (1.51)$$

β étant positif et $\alpha \geq -1$, on a $\alpha + \beta \geq -1$; l'annulation des termes les plus divergents donne :

$$\beta(\beta - 1) = l(l + 1) \iff \beta = l + 1 . \quad (1.52)$$

En définitive, en présence d'un potentiel $V(r) \propto r^\alpha$, ($\alpha \geq -1$), u se comporte comme r^{l+1} à l'origine ; la fonction radiale $R(r)$ varie donc comme r^l ; seuls les états s ont une densité de probabilité de présence non nulle en $r = 0$ ($R(r) \propto r^0$) :

$$V(r) \propto r^\alpha \quad (r \sim 0, \alpha \geq -1) \implies R(r) \propto r^l \iff u(r) \propto r^{l+1} \quad (r \sim 0) . \quad (1.53)$$

Discutons maintenant brièvement le comportement de $R(r)$ à l'infini, en se cantonnant aux états liés ($E < 0$ compte tenu de (1.44)). On voit tout de suite que u ne peut avoir un comportement en pure loi-puissance ($u \propto r^{-\lambda}$, $\lambda > 0$) à l'infini, car il serait impossible de satisfaire l'équation (1.32) avec les termes dominants. En revanche, ceci devient possible si u contient un facteur exponentiel ; il en résulte :

$$R(r) \propto r^\lambda e^{-kr} \quad (r \rightarrow +\infty, k = \sqrt{\frac{-2\mu E}{\hbar^2}}) \quad (1.54)$$

où la valeur précise de l'exposant – positive ou négative, peu importe – reste à trouver ; le point important à retenir est que R a alors un comportement essentiellement exponentiel.

Examinons enfin ce qui se passe lorsque $V(r)$ présente un saut fini pour une certaine valeur r_0 . L'intégration de l'équation pour u de part et d'autre de cette singularité fournit :

$$-\frac{\hbar^2}{2\mu} [u'(r_0 + 0) - u'(r_0 - 0)] + \int_{r_0-0}^{r_0+0} dr \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] u(r) = E \int_{r_0-0}^{r_0+0} dr u(r) . \quad (1.55)$$

Comme $|u|^2$ est une densité de probabilité, $|u|^2$ doit être partout localement sommable. En admettant (au pire) pour u une divergence comme $(r - r_0)^{-\lambda}$, il faut que $\lambda < \frac{1}{2}$ et alors l'intégrale du second membre donne zéro. Dans l'intégrale au premier membre, le terme en $\frac{l(l+1)}{r^2}$ est constant près de r_0 et $u(r)$ est sommable. Quant au terme en $V(r)u(r)$, il est donc en $(r - r_0)^{-\lambda} \theta(r - r_0)$, avec $\lambda < \frac{1}{2}$: son intégrale est nulle. Il reste finalement :

$$u'(r_0 + 0) - u'(r_0 - 0) = 0 ; \quad (1.56)$$

cette égalité exprime la continuité de la dérivée et, par voie de conséquence, celle de $u(r)$. Tout comme la fonction d'onde en dimension 1, la fonction radiale est continue et à dérivée continue même en présence d'un saut *fini* de potentiel.

Remarque

Les problèmes à trois dimensions sont nettement plus complexes que ceux à une dimension – notamment, le simple puits carré introduit des fonctions spéciales. En outre, tout ce que l'on sait des propriétés générales du mouvement à une dimension ne se généralise pas forcément à trois dimensions. Par exemple, alors que tout état lié à une dimension¹⁸ est non-dégénéré, ce n'est visiblement pas le cas dans \mathbb{R}^3 (pour un potentiel à symétrie sphérique, il y a déjà la dégénérescence liée à la symétrie sphérique). Autre exemple : le courant de probabilité est nul dans tout état stationnaire à une dimension ; ce n'est plus vrai à trois dimensions : il peut exister des courants permanents (stationnaires) qui “tournent en rond”¹⁹, ce qui n'est pas possible dans \mathbb{R} pour des raisons purement topologiques.

¹⁸sur *tout* \mathbb{R} ; sur un intervalle fini ou semi-infini, certaines propriétés de dégénérescence sont modifiées, notamment la double dégénérescence de tout état non-lié.

¹⁹Dans un état de l'atome d'hydrogène où L_z n'est pas nul, il y a une “boucle de courant” perpendiculaire à Oz .

1.2.2 Résolution de l'équation radiale

Comme toujours, les conditions requises par le sens physique attribué à la fonction d'onde provoquent la quantification spontanée de l'énergie des états liés. Elles ont été vues à l'œuvre à propos des harmoniques sphériques : c'est le caractère monovalué de la fonction d'onde (purement spatiale) qui engendre la quantification de la composante L_z du moment cinétique (et l'obligation pour le nombre quantique l d'être entier²⁰ – ce qui a d'ailleurs comme conséquence que $2l + 1$ est un nombre *impair*, un détail (!) qui a son importance). De même, la condition de normalisabilité de la fonction radiale $R(r)$ ne sera satisfaite que si l'énergie prend ses valeurs dans un ensemble discret $\{E_n\}$, isomorphe à \mathbb{N} pour le potentiel hydrogénoïde. Il sera en outre un peu surprenant de constater, en bout de course, que la formule donnant l'énergie est exactement la même que celle issue de la théorie de Bohr : si on doit s'attendre à ce que les deux expressions de l'énergie coïncident dans la limite des grands nombres quantiques, différant l'une de l'autre par des corrections tendant vers zéro quand $n \rightarrow +\infty$, rien n'assure d'avance qu'elles soient identiques $\forall n$. C'est pourtant ce que l'on trouve : les deux expressions de l'énergie coïncident quel que soit n .

L'équation à résoudre pour la fonction $u(r)$ est²¹ :

$$-\frac{\hbar^2}{2\mu} \frac{d^2 u}{dr^2} + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{e'^2}{r} \right] u(r) = E u(r) . \quad (1.57)$$

Pour les états liés, l'énergie E est négative ; en effet, quand r tend vers l'infini, le terme potentiel est négligeable et il reste :

$$-\frac{\hbar^2}{2\mu} \frac{d^2 u}{dr^2} \simeq E u(r) . \quad (1.58)$$

Pour que la solution u soit normalisable, il est nécessaire qu'elle tende vers zéro à l'infini : si E était positif, u oscillerait sans cesse ; au contraire, avec $E < 0$, on devine un comportement essentiellement exponentiel, conformément à ce qui a été montré dans la sous-section 1.2.1.

Pour trouver les bonnes solutions physiques de cette équation, on utilise la méthode polynômiale comme cela a été fait à propos de l'oscillateur harmonique à une dimension²². Comme d'habitude, afin de simplifier le plus tôt possible le problème mathématique sur la base de considérations physiques, posons-nous la question du comportement précis à l'infini des solutions ; ceci permet d'écartier d'emblée celles qui ne sont certainement pas acceptables. Quand on reporte dans (1.57) un comportement du type :

$$u(r) \simeq r^\lambda e^{-kr} \quad (r \rightarrow +\infty) , \quad (1.59)$$

et ne retenant que les termes dominants²³, on obtient :

$$-\frac{\hbar^2 k^2}{2\mu} r^\lambda e^{-kr} \simeq E r^\lambda e^{-kr} . \quad (1.60)$$

E étant négatif, la forme simplifiée (1.59) est donc solution si $k = \pm \sqrt{\frac{-2\mu E}{\hbar^2}}$; les seules solutions physiques se comportent donc comme suit à l'infini :

$$u(r) \simeq r^\lambda e^{-kr} , \quad k = \sqrt{\frac{-2\mu E}{\hbar^2}} > 0 \quad (r \rightarrow +\infty) , \quad (1.61)$$

l'exposant λ étant pour l'instant indéterminé (il peut visiblement être positif ou négatif).

Par ailleurs, on sait d'après (1.53) que le comportement de u à l'origine est :

$$u(r) \simeq r^{l+1} , \quad (r \simeq 0) . \quad (1.62)$$

²⁰Pour un moment cinétique de *spin*, ce nombre peut être entier ou demi-entier.

²¹Pour un ion hydrogénoïde de charge nucléaire Z , il suffit de remplacer partout e'^2 par Ze'^2 .

²²voir le cours d'Édouard Brézin.

²³Le terme dominant venant de la dérivée seconde est $r^\lambda e^{-kr}$. Les termes en r^{-1} et r^{-2} venant du potentiel effectif dans (1.57) sont négligeables.

Toutes ces considérations suggèrent des changements de variable et de fonction inconnue. Outre le changement évident $\rho = kr$ introduisant une variable radiale adimensionnée, il est en effet commode d'exhiber les comportements asymptotiques maintenant déterminés en introduisant une nouvelle fonction $w(\rho)$:

$$u(r) = \rho^{l+1} e^{-\rho} w(\rho) , \quad \rho = kr . \quad (1.63)$$

La fonction $w(\rho)$ ne doit pas croître plus vite que e^ρ quand $r \rightarrow +\infty$ puisque la condition de normalisabilité impose que $u(r)$ tende vers zéro à l'infini. En outre, la fonction $w(\rho)$ doit être finie à l'origine (notamment : pas de puissances négatives de r !) puisque le comportement de u , connu d'après (1.53), a été mis en exergue. La substitution dans l'équation complète (1.57) donne :

$$\rho \frac{d^2 w}{d\rho^2} + 2(l+1-\rho) \frac{dw}{d\rho} + (\lambda - 2l - 2) w = 0 \quad (1.64)$$

avec :

$$\lambda = \sqrt{\frac{2\mu c'^4}{-\hbar^2 E}} . \quad (1.65)$$

On pose maintenant pour $w(\rho)$ un développement en série entière ne contenant que des termes de puissance positive ou nulle :

$$w(\rho) = \sum_{p=0}^{+\infty} c_p \rho^p . \quad (1.66)$$

Notons que c_0 est différent de zéro puisque r^{l+1} est déjà précisément le comportement de la fonction u près de $r = 0$: la série doit donc se comporter comme une constante dans ce voisinage. Le report dans l'équation (1.64) fournit la récurrence suivante²⁴ :

$$(p+1)[p+2(l+1)] c_{p+1} + (\lambda - 2l - 2 - 2p) c_p = 0 , \quad p \in \mathbb{N} , \quad c_0 \neq 0 . \quad (1.67)$$

Ceci permet bien d'exprimer tous les coefficients en fonction du premier c_0 . La relation (1.67) définit donc complètement l'ensemble des solutions mathématiques qui sont *régulières* à l'origine, mais cette dernière condition n'est pas suffisante pour élire toutes ces solutions comme physiquement acceptables : on va en effet montrer que la série (1.66) se comporte comme $e^{+2\rho}$ à l'infini ; elle ne peut conduire à une solution convenable que si tous ses coefficients nuls à partir d'un certain rang, auquel cas la série se réduit alors à un *polynôme* ; quel que soit le degré de ce polynôme, l'exponentielle décroissante dans (1.63) finit par l'emporter et u , tendant vers zéro à l'infini, possède le bon comportement.

La divergence exponentielle de la série (1.66) se met en évidence comme suit. À l et E fixés, pour p grand, la récurrence (1.67) s'écrit à peu près :

$$p^2 c_{p+1} - 2p c_p \simeq 0 \quad \Longleftrightarrow \quad \frac{c_{p+1}}{c_p} \simeq \frac{2}{p} . \quad (1.68)$$

Il existe une série bien connue dont le rapport de deux termes consécutifs est $\frac{2}{p}$, c'est $e^{2\rho}$:

$$e^{2\rho} = \sum_{p=0}^{+\infty} \frac{1}{p!} (2\rho)^p \equiv \sum_{p=0}^{+\infty} \gamma_p \rho^p , \quad \frac{\gamma_{p+1}}{\gamma_p} = \frac{2}{p} . \quad (1.69)$$

Pour ρ tendant vers l'infini, ce sont les grandes valeurs de p qui comptent ; il en résulte que, tant que w est donnée par la série entière, w se comporte à peu près comme $e^{2\rho}$ quand $r \rightarrow +\infty$. Une autre façon d'établir ce fait consiste à poser que pour r grand, on a $w \sim \rho^q e^{\alpha\rho}$. Le report dans (1.64) donne, après simplification par $\rho^{q-1} e^{\alpha\rho}$:

$$q(q-1) + 2q\alpha\rho + \alpha^2\rho^2 + 2(l+1-\rho)(q+\alpha\rho) + (\lambda - 2l - 2)\rho \sim 0 . \quad (1.70)$$

²⁴On peut étendre les valeurs de p aux valeurs entières négatives à condition de garder en tête que $c_{-1} = 0$. Par ailleurs, la présence du facteur $(p+1)$ pour c_{p+1} assure le découplage des $c_{p \geq 0}$ et des éventuels $c_{p < 0}$; ce fait traduit l'existence de deux classes de solutions linéairement indépendantes – comme pour toute équation du second ordre. L'une des classes contient toutes les solutions régulières à l'origine, celles que l'on retient ici. L'autre classe contient toutes les solutions divergentes en $\rho = 0$, mises à l'écart par le choix (1.63) qui résulte des conditions physiques imposées aux solutions à retenir exclusivement.

Seuls les termes les plus divergents doivent être considérés²⁵, soit le terme en ρ^2 ; son coefficient, égal à $(\alpha^2 - 2\alpha)$, doit être nul, ce qui fournit bien la seule valeur acceptable²⁶ $\alpha = 2$, confirmant l'analyse fondée sur la comparaison des séries.

Au total, suivant (1.63), on trouve que $u(r)$ se comporte comme :

$$u(r) \simeq \rho^{l+1} e^{-\rho} e^{+2\rho} = \rho^{l+1} e^{+\rho} \quad (1.71)$$

Ainsi, si on garde la série (1.66) telle quelle, la fonction $u(r)$ *diverge exponentiellement* quand r tend vers l'infini. La seule et unique façon de soigner cette divergence inacceptable physiquement est d'annuler tous les coefficients c_p à partir d'un certain rang ; pour ceci, une seule possibilité : il faut (et il suffit) que la quantité $\lambda - 2l - 2 - 2p$ s'annule pour un certain entier p_0 ; dans ces conditions, $c_{p_0+1} = c_{p_0+2} = \dots = 0$, la série se réduit à un polynôme de degré p_0 , et $u(r)$ se comporte comme suit à l'infini :

$$u(r) \simeq \rho^{l+1+p_0} e^{-\rho} . \quad (1.72)$$

p_0 est visiblement un entier positif ou nul ; quelle que soit la valeur finie de p_0 , l'exponentielle dans (1.72) finit toujours par l'emporter. Les seules bonnes solutions sont donc celles pour lesquelles :

$$\lambda - 2l - 2 = 2p_0 , \quad p_0 \in \mathbb{N}. \quad (1.73)$$

p_0 et l sont des entiers positifs ou nuls, donc les valeurs possibles de λ sont les entiers pairs strictement positifs :

$$\lambda = 2n , \quad n \in \mathbb{N}^* ; \quad (1.74)$$

compte tenu de la définition de λ , (1.65), ceci donne :

$$E \in \{E_n\} : \quad E_n = - \frac{\mu e'^4}{2n^2 \hbar^2} \quad (n \in \mathbb{N}^*). \quad (1.75)$$

Ainsi, une fois de plus, la condition de normalisabilité des fonctions d'onde des états liés vient faire le tri parmi toutes les solutions mathématiques. Tout comme la condition de Bohr - Wilson - Sommerfeld $\oint p dq = \text{entier} \times h$ le fait parmi toutes les solutions "classiques", la condition de normalisabilité passe au peigne fin toutes les solutions mathématiques pour en extraire un ensemble infini dénombrable, produisant *de facto* la quantification de l'énergie.

Les fonctions d'onde R , qui dépendent évidemment de l'entier l figurant explicitement dans l'équation différentielle, seront également repérées par l'indice n associé à la valeur de l'énergie ; on notera ainsi $R_{nl}(r)$ la fonction radiale qui, multipliée par Y_{lm} , constitue l'état propre de (H, \vec{L}^2, L_z) pour l'atome d'hydrogène :

$$\left(\frac{\vec{p}^2}{2\mu} - \frac{e'^2}{r} \right) \psi_{nlm}(r, \theta, \phi) = E_n \psi_{nlm}(r, \theta, \phi) , \quad \psi_{nlm}(r, \theta, \phi) = R_{nl}(r) Y_{lm}(\theta, \phi) . \quad (1.76)$$

L'expression (1.75) de l'énergie est strictement identique à celle de la théorie de Bohr, dont on sait qu'elle reproduit²⁷ la formule de Balmer pour les termes spectraux de l'hydrogène, compte tenu de la correction de masse réduite. Comme annoncé, cette coïncidence $\forall n$ a quelque chose de surprenant²⁸.

²⁵ Les autres devraient être comparés avec des termes omis en posant $w \sim \rho^q e^{\alpha\rho}$.

²⁶ L'équation $\alpha^2 - 2\alpha = 0$ a aussi la solution $\alpha = 0$, auquel cas la fonction $w(\rho)$ se comporte à l'infini comme un polynôme de degré q , indéterminé à ce stade. Compte tenu de la conclusion qui va suivre, retenir cette solution n'enrichit donc pas l'ensemble des solutions acceptables.

²⁷ On se souvient que Bohr a construit sa théorie pour qu'il en soit ainsi.

²⁸ Les auteurs qui la relèvent se bornent à la constater, ajoutant le plus souvent qu'il s'agit d'une "coïncidence mathématique" : on n'est guère avancé ! Par ailleurs, on peut dès à présent noter une propriété remarquable de l'expression de l'énergie (1.75) : elle ne dépend pas du nombre quantique l associé au module du moment cinétique orbital. Cette propriété est spécifique du potentiel Coulombien pur et donne lieu à ce qui est appelé souvent – mais incorrectement – dégénérescence accidentelle. En réalité, cette dégénérescence résulte d'une symétrie remarquable du potentiel en $\frac{1}{r}$, qui donne lieu à la conservation dynamique du vecteur de Lenz - Runge (voir sous-section 1.3). Tout écart à la loi en $\frac{1}{r}$ – par exemple un terme en $\frac{1}{r^2}$ déstabilisant (parfois utilisé pour simuler un écrantage aux courtes distances) –, supprime cette propriété et l'énergie devient alors une fonction explicite de n et de l .

A contrario, l'analyse des états non-liés n'introduit aucune condition de normalisabilité. Ces états représentent la diffusion par le potentiel de Coulomb de particules émises par une source située à l'infini. Il y a bien sûr des conditions aux limites à satisfaire, qui prendront en compte notamment la valeur du courant de particules émis par la source et la localisation de celle-ci. Elles ne feront que généraliser ce qui a été vu à une dimension d'espace : dans ce dernier cas, on sait bien que c'est la position de la source et ses caractéristiques²⁹, qui permettent de déterminer toutes les constantes d'intégration apparaissant dans la résolution de l'équation aux valeurs propres³⁰. En l'absence d'une condition de normalisabilité, l'énergie n'est pas quantifiée.

Le spectre du Hamiltonien coulombien pur se compose donc de deux parties :

1. une partie discrète bornée inférieurement³¹ du côté négatif par $E_{n=1}$, constituée de niveaux de plus en plus denses quand n augmente, en nombre infini dénombrable, avec un point d'accumulation en $E = 0$. Cette propriété est caractéristique³² des potentiels qui tendent vers zéro moins vite que $\frac{1}{r^2}$ ($r^2V(r) \rightarrow +\infty$ si $r \rightarrow +\infty$) ; pour un potentiel allant plus rapidement vers zéro que $\frac{1}{r^2}$ ($r^2V(r) \rightarrow 0$ si $r \rightarrow +\infty$), le nombre d'états liés est fini (et parfois nul d'ailleurs si le potentiel n'est pas assez attracteur et/ou pas assez étendu dans l'espace) ; le puits carré sphérique appartient à cette catégorie.
2. une partie continue (dense) s'étendant de 0 à $+\infty$, correspondant aux états non-liés (états stationnaires de diffusion).

L'énergie du fondamental est $E_1 \simeq -13.6$ eV. Par ailleurs, le rayon de la première orbite de Bohr (dans l'approximation du noyau infiniment massif) est³³ :

$$a_0 = \frac{\hbar^2}{me'^2} \simeq 0.53 \text{ \AA} \quad (1.77)$$

ce qui permet de réécrire l'expression de l'énergie :

$$E_n = -\frac{\mu}{m_e} \frac{e'^2}{2n^2 a_0} . \quad (1.78)$$

Une autre expression est encore utile, qui fait apparaître la constante de structure fine, α :

$$E_n = -\frac{\alpha^2}{2n^2} \mu c^2 \quad \alpha = \frac{e'^2}{\hbar c} = \frac{1}{137.007 \dots} . \quad (1.79)$$

Comme l'entier p_0 , où la série (1.66) est tronquée, est positif ou nul, les seules valeurs possibles de l pour n fixé satisfont³⁴ (voir (1.73) et (1.74)) :

$$2n - 2l - 2 \geq 0 \quad \iff \quad l \leq n - 1 . \quad (1.80)$$

On sait déjà que l est positif ou nul ; au total, à n fixé, c'est-à-dire à énergie fixée, les seules valeurs possibles de l sont :

$$l = 0, 1, 2, \dots, n - 1 . \quad (1.81)$$

²⁹ Par exemple le courant qu'elle émet.

³⁰ Ainsi, pour un potentiel localisé $V(x)$ – nul en-dehors d'un intervalle fini $[a, b]$ –, la solution pour $x > b$ est de la forme $Ae^{ikx} + Be^{-ikx}$, $k > 0$. Si la source est localisée en $x = -\infty$, la constante B est en fait nulle. Le module de A sera finalement trouvé, *via* les coefficients de réflexion et de transmission, par calage sur le courant de la source, $j_s : \frac{\hbar k}{m} |A|^2 = j_s$.

³¹ Le fait que l'énergie E soit bornée inférieurement est déjà en soi un succès majeur de la théorie. En effet, un argument simple de stabilité maximale, développé dans un cadre strictement classique s'appuyant sur le théorème du Viriel montre que l'état d'énergie la plus basse est celui où l'électron est ... sur le noyau ! Comme alors l'énergie est infinie négative, c'est bien vrai que l'on ne peut pas faire "mieux". L'atome classique n'a vraiment aucune chance d'exister : outre cette instabilité purement mécanique, il est aussi foncièrement instable électrodynamiquement parlant (l'électron, particule chargée accélérée, rayonne, perd peu à peu (!?) son énergie et se précipite sur le noyau en un temps de l'ordre de 10^{-8} s).

³² [1], p. 354

³³ m_e désigne la masse de l'électron, $m_e \simeq 9 \times 10^{-31}$ kg.

³⁴ Les états $l = 0$ correspondent aux orbites circulaires de la théorie de Bohr.

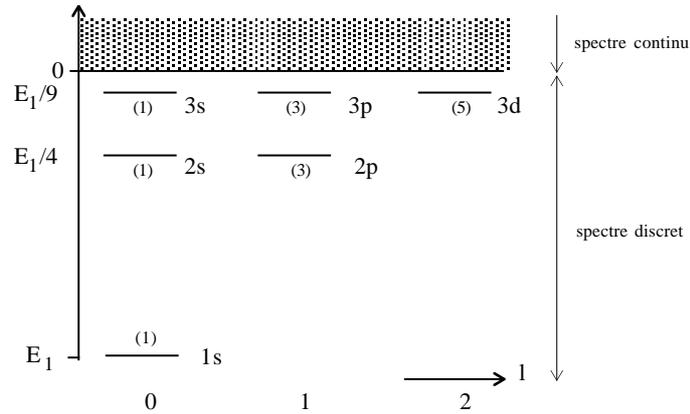


Figure 1.1: Spectre d'énergie de l'atome d'hydrogène et notation spectroscopique.

L'énergie des états liés de l'atome d'hydrogène ne dépend pas de l ; comme à l fixé il y a $2l + 1$ valeurs possibles de m , la dégénérescence totale³⁵ est :

$$g_n = \sum_{l=0}^{n-1} (2l + 1) = 2 \frac{(n-1)n}{2} + n = n^2 . \quad (1.82)$$

L'indépendance de l'énergie par rapport à l a été brièvement mentionnée plus haut et n'était pas prévisible au vu des seuls arguments de symétrie de rotation ; c'est pourquoi on la qualifie parfois de dégénérescence "accidentelle"³⁶. Les différents états de l'électron dans l'atome d'hydrogène sont donc entièrement caractérisés par la donnée des trois nombres quantiques n , l et m , tant que l'on maintient le spin à l'écart. Pour des raisons historiques, tenant à la classification "zoologique" des états construite en une période où même la théorie de Bohr n'était pas faite, l'usage est de représenter ce triplet de nombres, ou plus précisément les deux premiers par un symbole appelé "notation spectroscopique". On le construit en mettant en tête la valeur de n , puis on accole une lettre en correspondance avec la valeur de l :

$$l = 0 \leftrightarrow s \quad l = 1 \leftrightarrow p \quad l = 2 \leftrightarrow d \quad l = 3 \leftrightarrow f \quad l = 4 \leftrightarrow g \quad \dots \quad (1.83)$$

s est pour "sharp", p pour "pure", d pour "diffuse", *etc.*, autant de qualificatifs attribués aux raies spectrales à l'aube de la spectroscopie atomique. La notation spectroscopique est ainsi : 1s (couche K) ; 2s, 2p (couche L) ; 3s, 3p, 3d (couche M), *etc.*

À propos des polynômes de Laguerre

Les fonctions radiales ainsi obtenues sont construites avec les *polynômes* satisfaisant (1.64) où λ est remplacé par sa valeur $\lambda = 2n$, compte tenu de (1.74) :

$$\rho \frac{d^2 w}{d\rho^2} + 2(l + 1 - \rho) \frac{dw}{d\rho} + (2n - 2l - 2) w = 0 . \quad (1.84)$$

Cette équation est un cas particulier de l'équation dite de Laplace ; écrite pour une fonction $L(z)$, cette équation est :

$$z \frac{d^2 L}{dz^2} + (\beta - z) \frac{dL}{dz} - \alpha L = 0 \quad (1.85)$$

³⁵Compte non-tenu du spin ! La prise en compte du spin $\frac{1}{2}$ de l'électron ajoute un facteur 2 : la dégénérescence totale est alors égale à $2n^2$, d'où les nombres d'occupation maximaux des couches atomiques (2, 8, 18, ...).

³⁶La qualification "accidentelle" est source de contresens. Il existe en fait une symétrie subtile, comme expliqué plus bas (sous-section 1.3) ; cette symétrie, dans l'espace des impulsions, remarquée par Fock en 1935, a été étudiée notamment par McIntosh [2] et par Bander & Itzykson [3]. En outre, cette dégénérescence particulière est liée au fait qu'il existe un autre jeu de coordonnées (les coordonnées paraboliques) pour lequel les variables spatiales se séparent aussi – ces coordonnées paraboliques sont d'ailleurs particulièrement bien adaptées au traitement exact (non-perturbatif) de l'effet Stark.

où α et β sont des constantes complexes quelconques. Il suffit en effet de poser :

$$w(\rho) = L(z = 2\rho) \quad (1.86)$$

pour que l'équation (1.84) devienne :

$$z \frac{d^2 L}{dz^2} + [2(l+1) - z] \frac{dL}{dz} + (n-l-1)L = 0, \quad (1.87)$$

qui, par comparaison avec (1.84), donne l'identification $\alpha = -n + l + 1$, $\beta = 2l + 2$. Les polynômes satisfaisant cette équation particulière de Laplace sont appelés polynômes associés de Laguerre, notés L_{n-l-1}^{2l+1} , aux propriétés bien répertoriées.

Les $L_p^0(z) \equiv L_p$ satisfont par définition l'équation différentielle homogène :

$$z \frac{d^2 L_p}{dz^2} + (1-z) \frac{dL_p}{dz} + p L_p = 0, \quad (1.88)$$

et, à ce stade, sont donc tous définis à un facteur près ; on achève de les définir en posant conventionnellement :

$$L_p(z=0) = p!. \quad (1.89)$$

Les L_p sont appelés polynômes de Laguerre. $L_p(z)$ est un polynôme de degré p , dont on peut montrer que tous les zéros sont réels.

Par ailleurs, il est facile de voir que la $q^{\text{ème}}$ dérivée, $L^{(q)}$, de la solution de cette équation obéit à :

$$z \frac{d^2 L^{(q)}}{dz^2} + (q+1-z) \frac{dL^{(q)}}{dz} + (p-q)L^{(q)} = 0 \quad (1.90)$$

qui est bien l'équation différentielle (1.87) obtenue plus haut avec l'identification $q = 2l + 1$, $p - q = n - l - 1$. Ceci conduit aux polynômes *associés* de Laguerre, précisément définis comme :

$$L_p^q(z) = (-1)^q \frac{d^q}{dz^q} L_{p+q}^0(z). \quad (1.91)$$

$L_p^q(z)$ est la $q^{\text{ème}}$ dérivée d'un polynôme de degré $p+q$, c'est donc encore un polynôme de degré p , tout comme $L_p(z)$.

Ceci étant précisé, les L_p peuvent s'écrire de diverses façons ; par exemple, on peut partir de :

$$L_p(z) = \frac{p!}{2i\pi} \oint ds \frac{e^{-sz}}{s^{p+1}} (1+s)^p \quad (1.92)$$

où l'intégrale est prise le long d'un petit contour entourant une fois l'origine du plan de s dans le sens positif³⁷. De l'expression intégrale (1.92), on déduit facilement :

$$L_p(z) = (-1)^p \frac{p!}{2i\pi} e^z \frac{d^p}{dz^p} \oint ds \frac{e^{-(s+1)z}}{s^{p+1}} = (-1)^p \frac{p!}{2i\pi} e^z \frac{d^p}{dz^p} \oint ds \frac{e^{-sz}}{s^{p+1}} e^{-z} \quad (1.94)$$

En développant maintenant l'exponentielle e^{-sz} sous l'intégrale, et en appliquant le théorème des résidus, il vient simplement :

$$L_p(z) = e^z \frac{d^p}{dz^p} [z^p e^{-z}]. \quad (1.95)$$

³⁷L'expression (1.92) "sort" naturellement quand on résout l'équation (1.88) par la méthode de Laplace, qui fonctionne efficacement pour toute équation différentielle dont les coefficients sont au plus des fonctions linéaires de la variable, réduisant d'une unité l'ordre de l'équation différentielle. Cette méthode consiste à poser la fonction inconnue sous la forme :

$$L_p(z) = \oint_C ds e^{-sz} f(s); \quad (1.93)$$

C est un contour à préciser ultérieurement et f devient la fonction inconnue. La terminologie est évidente : l'intégrale ressemble manifestement à une transformation de Laplace ordinaire (pour plus de détails, voir [4], appendice a).

L'exponentielle en préfacteur assure bien que L_p est un simple polynôme, de degré p , dont le coefficient du facteur de plus haut degré est d'ailleurs égal à $(-1)^p$. Le calcul de la dérivée $p^{\text{ème}}$ donne :

$$L_p(z) = \sum_{r=0}^p \frac{(-1)^r p!^2}{(p-r)! r!^2} z^r . \quad (1.96)$$

La fonction génératrice $F(t, z)$ des L_p est par définition :

$$F(t, z) = \sum_{p=0}^{+\infty} L_p(z) \frac{t^p}{p!} . \quad (1.97)$$

$F(t, z)$ est donc la fonction dont les coefficients de la série de Taylor autour de $t = 0$ sont³⁸ les $L_p(z)$:

$$L_p(z) = \left(\frac{\partial^p}{\partial t^p} F(t, z) \right)_{t=0} . \quad (1.99)$$

$F(t, z)$ s'obtient en utilisant l'expression (1.96) d'où :

$$F(t, z) = \sum_{r=0}^{+\infty} \frac{(-zt)^r}{r!} \sum_{p=0}^{+\infty} C_{p+r}^p t^p . \quad (1.100)$$

La sommation interne s'écrit :

$$\sum_{p=0}^{+\infty} C_{p+r}^p t^p = \sum_{p=0}^{+\infty} \frac{t^p (p+r)!}{p! r!} = \sum_{p=0}^{+\infty} (p+r)(p+r-1) \dots (r+2)(r+1) \frac{t^p}{p!} . \quad (1.101)$$

Ceci ressemble à une série de Taylor ; de fait, en observant que :

$$\frac{d^p}{dt^p} \frac{1}{(1-t)^{r+1}} = \frac{(r+1)(r+2) \dots (r+p)}{(1-t)^{r+1+p}} \implies (r+1)(r+2) \dots (r+p) = \left(\frac{d^p}{dt^p} \frac{1}{(1-t)^{r+1}} \right)_{t=0} , \quad (1.102)$$

on voit que la somme interne n'est autre que la série de Taylor de $(1-t)^{-(r+1)}$ autour de l'origine. D'où, reportant dans (1.100) :

$$F(t, z) = \sum_{r=0}^{+\infty} \frac{(-zt)^r}{r!} \frac{1}{(1-t)^{r+1}} = \frac{1}{1-t} e^{-\frac{tz}{1-t}} . \quad (1.103)$$

La série de Taylor (1.97) ne converge que si $|t| < 1$, mais sa somme peut être prolongée analytiquement et devient ainsi une fonction analytique dans le plan complexe ouvert privé du point $t = 1$, où elle a une singularité essentielle, tant que $z \neq 0$. Sur l'expression de $F(t, z)$, ou sur (1.96), on vérifie bien la valeur conventionnelle $L_p(0) = p!$:

$$F(t, z=0) = \sum_{r=0}^{+\infty} L_p(0) \frac{t^p}{p!} = \frac{1}{1-t} = \sum_{p=0}^{+\infty} t^p \quad (|t| < 1) . \quad (1.104)$$

Comme déjà mentionné, le polynôme associé, L_p^q , en tant que dérivée $q^{\text{ème}}$ d'un polynôme de degré $p+q$, est un polynôme de degré p . En raison de la définition (1.91) des L_p^q , leur fonction génératrice est $F_q(t, z)$:

$$F_q(t, z) \equiv \sum_{p=0}^{+\infty} L_p^q(z) \frac{t^p}{(p+q)!} = \sum_{p=0}^{+\infty} (-1)^q \frac{d^q}{dz^q} L_{p+q}(z) \frac{t^p}{(p+q)!} = (-t)^{-q} \frac{d^q}{dz^q} \sum_{p=q}^{+\infty} L_p(z) \frac{t^p}{p!} . \quad (1.105)$$

³⁸De ce fait, par la formule de Cauchy, on a :

$$L_p(z) = \frac{p!}{2i\pi} \oint \frac{F(\xi, z)}{\xi^{p+1}} d\xi , \quad (1.98)$$

où le contour est une petite boucle autour de l'origine. En injectant l'expression (1.103) de $F(t, z)$ et en posant $\frac{\xi}{1-\xi} = s$, on retombe bien sur l'expression intégrale (1.92) de $L_p(z)$.

La somme n'est pas exactement $F(t, z)$, mais comme les termes manquants sont des monômes en z de degré variant de 0 à $q - 1$, leur $q^{\text{ème}}$ dérivée est nulle, et on peut de fait commencer la sommation à $p = 0$. Au total :

$$F_q(t, z) = (-t)^{-q} \frac{d^q}{dz^q} \sum_{p=0}^{+\infty} L_p(z) \frac{t^p}{p!} = (-t)^{-q} \frac{\partial^q}{\partial z^q} F(t, z) = \frac{1}{(1-t)^{q+1}} e^{-\frac{tz}{1-t}} . \quad (1.106)$$

Faisant $z = 0$, on en déduit :

$$F_q(t, z = 0) = \sum_{p=0}^{+\infty} L_p^q(0) \frac{t^p}{(p+q)!} = \frac{1}{(1-t)^{q+1}} ; \quad (1.107)$$

reprenant alors le développement de Taylor de $(1-t)^{-(q+1)}$ (voir (1.102)), on voit que :

$$L_p^q(z = 0) = \frac{[p+q]!^2}{p!q!} . \quad (1.108)$$

En définitive, les fonctions radiales sont les $R_{nl}(r)$:

$$R_{nl}(r) = C_{nl} r^l L_{n-l-1}^{2l+1}(2k_n r) e^{-k_n r} \quad (1.109)$$

où C_{nl} est la constante de normalisation radiale déduite de :

$$\int_0^{+\infty} r^2 dr R_{nl}^2(r) = 1 , \quad (1.110)$$

la partie angulaire étant elle aussi normalisée à part. Les k_n sont les nombres d'onde déduits de la définition de k (voir (1.61)) et de l'expression de E_n donnée en (1.75) :

$$k_n = \frac{\mu}{m_e} \frac{1}{na_0} . \quad (1.111)$$

L_{n-l-1}^{2l+1} est un polynôme de degré $n-l-1$. R_{nl} contient toute la dépendance radiale. L'expression normalisée explicite des R_{nl} est :

$$R_{nl}(r) = \left[(2k_n)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3} \right]^{\frac{1}{2}} (2k_n r)^l L_{n-l-1}^{2l+1}(2k_n r) e^{-k_n r} . \quad (1.112)$$

Le cas où l prend sa valeur maximale $n-1$ est simple à analyser puisqu'alors le polynôme associé de Laguerre est de degré zéro (c'est donc une constante) ; $u(r)$ varie alors comme $r^{l+1=n}$ (et $R(r)$ est en r^{n-1}) ; dans ces conditions, la fonction radiale est :

$$R_{nl=n-1}(r) = C_{nn-1} r^{n-1} e^{-k_n r} . \quad (1.113)$$

Pour ces états $l = n-1$, la densité de probabilité radiale est proportionnelle à $r^{2n} e^{-2k_n r}$, le facteur additionnel r^2 provenant de l'élément de volume radial $r^2 dr$. Cette fonction présente un maximum pour $r = r_{\max}$ donné par :

$$r_{\max} = \frac{n}{k_n} = \frac{m_e}{\mu} n^2 a_0 \simeq n^2 a_0 . \quad (1.114)$$

Cette distance au noyau où la densité de probabilité est maximale est donc égale à l'orbite de l'électron dans le $n^{\text{ème}}$ état prévu par la théorie de Bohr. Par ailleurs, comme u se comporte comme r^{l+1} près de $r = 0$, R_{nl} varie comme r^l : seuls les états s ($l = 0$) ont une densité de probabilité non nulle à l'origine.

Dans les fonctions d'onde complètes :

$$\psi_{nlm}(r, \theta, \phi) = R_{nl}(r) Y_{lm}(\theta, \phi) , \quad (1.115)$$

les Y_{lm} portent toute la dépendance angulaire ; les états $l = 0$ sont à symétrie sphérique, comme on le sait, les états $l = 1, m = 0, \pm 1$ ont au contraire une symétrie axiale en module (leur module est invariant par

rotation autour de Oz). Il ne faut pas être choqué d'obtenir des états propres ayant une symétrie inférieure à celle du Hamiltonien : après tout, le choix de l'axe Oz , axe de quantification du moment cinétique, est totalement arbitraire et n'importe quel autre choix conviendrait tout autant. D'ailleurs, toute combinaison linéaire impliquant les trois harmoniques sphériques $Y_{1m=0,\pm 1}$ est tout autant acceptable – puisqu'aucune direction n'est privilégiée, ce qui se traduit par la dégénérescence de l'énergie vis-à-vis de m , quelle que soit la forme précise du potentiel central $V(r)$; une telle combinaison ne possède pas la symétrie axiale simple de l'une des Y_{1m} – la même remarque vaut bien sûr pour un l quelconque.

Pour terminer, calculons le courant \vec{j} pour un état stationnaire Ψ construit avec l'état propre ψ_{nlm} . On se souvient que Y_{lm} contient le facteur $e^{im\phi}$ et qu'il s'agit de la seule contribution complexe à la fonction d'onde ; par ailleurs, le gradient a pour composantes sphériques :

$$\vec{\nabla} = \left(\frac{\partial}{\partial r}, \frac{1}{r} \frac{\partial}{\partial \theta}, \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \right) . \quad (1.116)$$

Comme l'expression (1.35) du courant contient la différence entre une quantité complexe et sa conjuguée, seuls comptent les termes imaginaires purs – tous les autres se compensent deux à deux ; de plus, seule la dérivation en ϕ donne une contribution complexe. Au total, j_r et j_θ sont nuls et le vecteur \vec{j} a pour seule composante non nulle j_ϕ , indiquant un courant tournant autour de Oz ; un calcul simple donne :

$$j_\phi = \frac{m\hbar}{\mu r \sin \theta} |\psi_{nlm}|^2 . \quad (1.117)$$

Le fluide de probabilité tourne donc autour de Oz ; ce n'est pas une rotation rigide (“en bloc”) – caractérisée par le fait que le courant j_ϕ est proportionnel à la distance à l'axe de rotation, soit $r \sin \theta$. Il est facile de calculer le moment cinétique global \vec{K} lié à la rotation du fluide ; visiblement dirigé le long de Oz , il se réduit à sa seule composante K_z . On a :

$$d\vec{K} = \mu \vec{r} \times \vec{j} d^3r \quad (1.118)$$

et :

$$K_z = \mu \int_{\mathbb{R}^3} d^3r j_\phi \sin \theta = \mu \int_{\mathbb{R}^3} d^3r \frac{m\hbar}{\mu} |\psi_{nlm}|^2 = m\hbar . \quad (1.119)$$

Autrement dit, le moment cinétique de rotation du courant de probabilité n'est autre que la composante L_z du moment cinétique orbital. Cet écoulement permanent est visiblement rendu possible parce que le système considéré est plongé dans \mathbb{R}^3 . Il est nul dans un état sphérique (états s) et, à l quelconque, pour l'état ayant $m = 0$ ($\langle L_z \rangle = 0$).

1.3 Symétrie dynamique du potentiel Coulombien

On sait que symétrie et dégénérescence sont étroitement liées, une évidence d'un point de vue physique : lors d'une transformation qui n'affecte pas le système (puisque'il s'agit d'une symétrie), son énergie ne saurait changer. Dans le cas de l'atome d'hydrogène, on pouvait notamment s'attendre, en l'absence de toute direction privilégiée, à ce que l'énergie ne dépende pas du nombre quantique magnétique m , puisque l'axe de quantification Oz est totalement arbitraire. En revanche, les mêmes arguments ne permettaient pas de prévoir que la valeur de l'énergie serait indépendante du module du moment cinétique représentée par le nombre quantique l . Cette dégénérescence “additionnelle” se présente à chaque fois que l'équation aux valeurs et fonctions propres peut être résolue dans différents systèmes de coordonnées et doit pouvoir se relier à une symétrie qui n'est pas forcément de nature géométrique. Ces symétries sont appelées symétries dynamiques car elles résultent d'une forme très particulière de l'énergie potentielle, disparaissant au moindre écart par rapport à cette loi. Ici, il s'agit du potentiel Coulombien pur, en $\frac{1}{r}$; tout terme additif (effet d'écran à longue portée, répulsion à courte distance, etc.) supprime cette dégénérescence “accidentelle” (elle ne fut pas tout de suite reconnue comme émanant d'une symétrie relativement subtile).

Classiquement, on sait que le potentiel Coulombien donne lieu à une constante du mouvement remarquable, le vecteur dit de Lenz-Runge \vec{R} dont la définition est :

$$\vec{R} = \frac{\vec{p}}{\mu} \times \vec{L} - e'^2 \frac{\vec{r}}{r} . \quad (1.120)$$

Il n'est pas difficile, en effet, de montrer que les crochets de Poisson de n'importe laquelle des composantes de ce vecteur avec H sont nuls. L'opérateur quantique correspondant se forme en symétrisant comme d'habitude :

$$\vec{R} = \frac{1}{2} \left(\frac{\vec{p}}{\mu} \times \vec{L} - \vec{L} \times \frac{\vec{p}}{\mu} \right) - e'^2 \frac{\vec{r}}{r} . \quad (1.121)$$

\vec{R} et \vec{L} ne commutent pas entre eux. Les commutateurs correspondants s'expriment commodément à l'aide du symbole ϵ_{uvw} tel que :

$$[L_u, L_v] = i\hbar \epsilon_{uvw} L_w , \quad (1.122)$$

et on trouve :

$$[L_u, R_v] = i\hbar \epsilon_{uvw} R_w \quad [R_u, R_v] = -2 \frac{i\hbar}{\mu} \epsilon_{uvw} L_w H . \quad (1.123)$$

$$[L_u^2, R_v] = i\hbar \epsilon_{uvw} (L_u R_w + R_w L_u) \quad \vec{L} \cdot \vec{R} = \vec{R} \cdot \vec{L} = 0 . \quad (1.124)$$

En outre, bien sûr :

$$[H, R_v] = 0 . \quad (1.125)$$

Ceci étant, on peut voir immédiatement que l'existence de cette constante du mouvement propre au champ Coulombien conduit à une dégénérescence de l'énergie vis-à-vis de l . En effet, par (1.125), on a :

$$\langle n l m | [H, R_z] | n l' m \rangle = 0 . \quad (1.126)$$

Appelons provisoirement E_{nl} l'énergie d'un état propre sphérique de H ; (1.126) donne :

$$(E_{nl} - E_{nl'}) \langle n l m | R_z | n l' m \rangle = 0 . \quad (1.127)$$

Comme R_z et \vec{L}^2 ne commutent pas, on peut affirmer que sur la base propre de \vec{L}^2 où cet opérateur est représenté par une matrice diagonale, la matrice de R_z n'est pas diagonale ; ses éléments non-diagonaux sont donc, en général, non nuls. Il résulte alors nécessairement de (1.127) que :

$$E_{nl} = E_{nl'} , \quad (1.128)$$

ce qui établit l'indépendance de l'énergie vis-à-vis de l . En définitive, la dégénérescence particulière vient du fait qu'il existe une constante du mouvement, \vec{R} , qui ne commute pas avec \vec{L}^2 , L_z . Deux ECOG sont ainsi disponibles : (H, \vec{L}^2, L_z) ou (H, \vec{R}^2, R_z) , correspondant aux deux jeux de coordonnées (sphériques ou paraboliques) pour lesquelles les coordonnées (variables) se séparent, *i.e.* permettent de factoriser les états propres.

Les trois composantes de \vec{L} sont les générateurs infinitésimaux des rotations dans \mathbb{R}^3 et forment une algèbre fermée ; on peut donc se poser la question de l'identification d'une algèbre "plus grande", incluant les composantes de \vec{R} , associée à un groupe de transformations connues³⁹. Dans ce but, commençons par redéfinir un autre vecteur \vec{R}' par :

$$\vec{R}' = \sqrt{\frac{-\mu}{2E}} \vec{R} , \quad (1.129)$$

puis renumérotions les différentes grandeurs dynamiques comme suit :

$$\vec{r} = (r_1, r_2, r_3) , \quad \vec{p} = (p_1, p_2, p_3) , \quad (1.130)$$

$$\vec{L} = (L_x, L_y, L_z) \equiv (L_{23}, L_{31}, L_{12}) . \quad (1.131)$$

Dans ces notations, on a les relations compactes :

$$L_{ij} = r_i p_j - r_j p_i , \quad [r_i, p_j] = i\hbar \delta_{ij} , \quad (1.132)$$

avec $i, j = 1, 2, 3$. Prolongeons maintenant ces deux dernières équations en "inventant" une quatrième dimension (coordonnée) r_4 et le moment correspondant p_4 , et posons :

$$R'_x = L_{14} , \quad R'_y = L_{24} , \quad R'_z = L_{34} . \quad (1.133)$$

³⁹Voir [5], p. 234.

Il est alors facile de vérifier que les 6 quantités L_{ij} satisfont :

$$[L_{ij}, L_{kl}] = i\hbar (L_{li}\delta_{jk} + L_{kj}\delta_{il} + L_{ik}\delta_{jl} + L_{jl}\delta_{ik}) , \quad (1.134)$$

montrant que les L_{ij} forment une algèbre fermée. Cette équation constitue aussi une écriture compacte, pour ces 6 générateurs, des relations de commutation des composantes du moment cinétique et du vecteur de Lenz-Runge. On peut alors concevoir que, tout comme L_x , L_y et L_z sont des générateurs des rotations dans \mathbb{R}^3 , groupe désigné par $O(3)$, les 6 générateurs ainsi définis constituent ceux d'un groupe d'opérations dans un espace plus vaste ; il s'agit en fait du groupe des rotations dans \mathbb{R}^4 , noté $O(4)$. Bien sûr, la symétrie correspondante n'est pas de nature géométrique puisque r_4 et p_4 n'ont pas le sens physique d'une coordonnée et d'un moment conjugué. Il s'agit cependant bel et bien d'une symétrie au sens général, symétrie induisant la dégénérescence sur le nombre l , d'où la qualification un peu abusive, mais usuelle, de dégénérescence *accidentelle*.

1.4 Interaction spin-orbite et structure fine

Jusqu'à présent, l'électron a été considéré comme ponctuel et sans aucune structure ou degré de liberté interne. Dans cette optique, en représentation- q pour fixer les idées, tous les degrés de liberté quantiques sont exprimables à l'aide des trois coordonnées d'espace, x , y et z par exemple. Toute l'information sur l'état du système à l'instant t est alors réputée entièrement contenue dans la connaissance de la fonction d'onde $\Psi(x, y, z ; t)$, satisfaisant l'équation de Schrödinger $i\hbar\frac{\partial}{\partial t}\Psi = H\Psi$, un état initial étant supposé donné.

Une telle description est insuffisante, comme on va le voir. Cette insuffisance provient des preuves expérimentales démontrant qu'une particule telle que l'électron possède un moment magnétique *propre*, indépendamment de tout mouvement de rotation dans l'espace autour d'un centre. L'existence de ce moment magnétique entraîne à son tour l'existence d'un moment cinétique propre, ou intrinsèque, qui a été baptisé *spin* car on croyait au début que ce degré de liberté était lié à une rotation de la particule sur elle-même⁴⁰. Ce degré de liberté est "interne" – bien que l'électron continue à être considéré comme une particule ponctuelle ; c'est, au même titre que la charge ou la masse, un attribut intrinsèque, donné une fois pour toutes. Il s'avère impossible de donner du spin une image classique ; se représenter l'électron comme une petite bille de rayon non-nul qui tourne sur elle-même conduit à des absurdités⁴¹ – il reste cependant que le spin d'une particule est son moment cinétique dans le référentiel où elle est au repos ([6], p. II. 4).

L'hypothèse du spin de l'électron a été formulée par Uhlenbeck et Goudsmit en 1925 pour rendre compte pragmatiquement des spectres des atomes complexes ; alliée au Principe de Pauli, elle est le fondement de la classification périodique des éléments de Mendeleiev. Bien au-delà de ces vertus explicatives, cette hypothèse a des implications et des conséquences considérables : le spin permet de distinguer deux classes de particules (les fermions et les bosons) ; c'est cette distinction qui permet, finalement, de comprendre ni plus ni moins que la stabilité de la matière !

Les preuves de l'existence d'un moment magnétique propre de l'électron sont multiples. L'expérience la plus importante – sur ce point comme pour d'autres – est certainement l'expérience de Stern et Gerlach. Dans sa version historique, elle impliquait des atomes d'argent dans l'état fondamental dont la symétrie sphérique exclut l'existence d'un moment cinétique de type orbital non-nul en moyenne et, par voie de conséquence, celle d'un moment magnétique permanent. Un moment cinétique orbital *fini*, à lui seul, aurait pu expliquer l'existence d'un moment magnétique engendré par une boucle de courant, mais, de toute façon, le nombre de taches observées sur l'écran est incompatible avec un moment cinétique orbital : pour l'atome d'Ag, on observe deux taches (nombre *pair*) alors qu'un moment cinétique orbital conduit à un nombre impair de valeurs possibles pour la projection le long d'un axe donné. L'apparition d'un nombre pair de taches exclut un moment magnétique d'origine orbital, lié classiquement au moment cinétique par le facteur gyromagnétique ordinaire.

Une autre preuve est fournie par l'effet Zeeman anormal. L'effet Zeeman désigne généralement l'éclatement des raies spectrales d'un atome quand on le soumet à un (petit) champ magnétique ; en présence du

⁴⁰On parle du spin de la Terre pour évoquer sa révolution diurne.

⁴¹Par exemple, on trouve qu'un point situé à la périphérie de l'électron a une vitesse linéaire très supérieure à c !!!

champ, chaque raie devient en général multiple et est composée de plusieurs raies très voisines (*splitting*), dont l'écart en fréquence par rapport à la raie unique observée en l'absence de champ est, en champ faible, simplement proportionnelle au module du champ appliqué. Les premières observations faisaient état d'une raie en champ nul se décomposant en trois raies très proches, plus généralement en un nombre impair de raies, dont une explication a même pu être fournie dans le cadre de l'Ancienne Théorie des Quanta. En revanche, des observations ultérieures ont révélé la possibilité d'un éclatement en un nombre pair de raies, phénomène incompréhensible tant que seuls des moments cinétiques orbitaux sont en jeu.

Enfin, l'existence de la structure fine apporte une autre preuve de l'existence du moment magnétique propre de l'électron. On sait qu'un moment magnétique $\vec{\mu}$ qui se déplace à la vitesse \vec{v} dans un champ électrique $\vec{\mathcal{E}}$ se couple, par son mouvement⁴², à ce champ électrique ; l'interaction est donnée par $\mu\vec{v} \times \vec{\mathcal{E}}$. Si le champ est à symétrie sphérique, il est forcément de la forme $\vec{\mathcal{E}} = f(r) \vec{r}$; on trouve alors la combinaison $\vec{r} \times \vec{v}$, proportionnelle au moment cinétique orbital \vec{L} , d'où finalement dans le Hamiltonien un terme du genre :

$$V_{\text{magn}} = a(r) \vec{L} \cdot \vec{\mu} . \quad (1.135)$$

Ce résultat s'obtient comme suit à partir d'arguments classiques. Dans le repère de l'électron, celui-ci voit tourner le noyau (une particule chargée !) autour de lui à la vitesse \vec{v}_N , produisant le champ magnétique $\vec{\mathcal{B}}(\vec{r}') = \delta(\vec{r}') \rho(\vec{r}') \vec{v}_N \times (\vec{r} - \vec{r}')$ puisque le noyau est pris ponctuel et situé à l'origine) :

$$\vec{\mathcal{B}}(\vec{r}) = \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} d^3r' \frac{\rho(\vec{r}') \vec{v}_N \times (\vec{r} - \vec{r}')}{\|\vec{r} - \vec{r}'\|^3} = \frac{\mu_0}{4\pi} \frac{Z|e| \vec{v}_N \times \vec{r}}{r^3} . \quad (1.136)$$

Dans le repère du noyau, la vitesse \vec{v} de l'électron est égale à $-\vec{v}_N$. Le champ magnétique est donc aussi :

$$\vec{\mathcal{B}}(\vec{r}) = -\frac{\mu_0}{4\pi} \frac{Z|e| \vec{v} \times \vec{r}}{r^3} \equiv \frac{\mu_0}{4\pi} \frac{1}{m} \frac{Z|e|}{r^3} \vec{L} , \quad (1.137)$$

où $\vec{L} = \vec{r} \times m\vec{v}$ est le moment cinétique orbital de l'électron en rotation autour du noyau. En désignant par $U(r) = \frac{Z|e|}{4\pi\epsilon_0 r}$ le potentiel électrostatique du noyau, (1.137) s'écrit⁴³ :

$$\vec{\mathcal{B}}(\vec{r}) = -\frac{1}{mc^2} \frac{1}{r} \frac{dU}{dr} \vec{L} . \quad (1.138)$$

L'électron se couple à ce champ magnétique par l'intermédiaire de son moment magnétique supposé ; si l'on croit l'argument classique qui vient d'être suivi, le terme d'interaction est :

$$-\vec{\mu} \cdot \vec{\mathcal{B}} = +\frac{1}{mc^2} \frac{1}{r} \frac{dU}{dr} \vec{L} \cdot \vec{\mu} \equiv a(r) \vec{L} \cdot \vec{\mu} \equiv V_{\text{magn}} . \quad (1.139)$$

En réalité, en raison d'un effet subtil appelé précession de Thomas, l'expression correcte de $a(r)$ contient un facteur $\frac{1}{2}$ additionnel⁴⁴ :

$$a(r) = +\frac{1}{2mc^2} \frac{1}{r} \frac{dU}{dr} . \quad (1.140)$$

L'origine physique de V_{magn} justifie que l'on donne à ce terme le nom d'interaction spin-orbite. Il produit des (petites) variations de l'énergie des différents états ; en effet, en remplaçant r par a_0 dans (1.139) et en admettant que $\vec{\mu}$ est lié à un moment cinétique par un facteur gyromagnétique du genre $\frac{|e|\hbar}{m}$ (à des facteurs près)⁴⁵, on trouve, en ordre de grandeur :

$$V_{\text{magn}} \sim \frac{1}{mc^2} \frac{1}{a_0^3} \frac{Z|e|}{4\pi\epsilon_0} \hbar \frac{|e|\hbar}{2m} \hbar \sim \left(\frac{\hbar}{mca_0} \right)^2 E_{n=1} \quad (1.141)$$

⁴²Tout comme une charge électrique q en mouvement dans un champ magnétique $\vec{\mathcal{B}}$ ressent une force $\vec{F} = q\vec{v} \times \vec{\mathcal{B}}$.

⁴³ $\epsilon_0 \mu_0 c^2 = 1$.

⁴⁴voir la Remarque à la fin de la sous-section 1.5.2, et notamment l'expression (1.244).

⁴⁵L'argument est visiblement d'inspiration classique ; il est cependant recevable sur un strict plan d'homogénéité et en ne jouant qu'avec les constantes fondamentales disponibles.

où $E_{n=1} \sim \frac{e'^2}{a_0}$ est l'ordre de grandeur de l'énergie fondamentale de l'atome d'hydrogène. Compte tenu de la définition de la première orbite de Bohr, $a_0 = \frac{\hbar^2}{me'^2}$, il vient :

$$V_{\text{magn}} \sim \alpha^2 E_{n=1} \quad (1.142)$$

où α est toujours la constante de structure fine, définie en (1.79). Les déplacements en énergie varient certainement d'un état à l'autre, mais ils restent petits vis-à-vis des énergies purement électrostatiques (voir (1.152)). En tout cas, combinés avec la règle de Bohr, ils donnent lieu à l'apparition d'une structure dans les raies spectrales : avec une résolution médiocre, on ne voit qu'une seule raie. En regardant mieux, on constate que chaque raie est le plus souvent constituée de plusieurs raies très voisines : c'est ce que l'on appelle la structure fine⁴⁶. Sommerfeld en a fourni une explication remarquable pour la raie d'émission H_α ($n = 3 \rightarrow n = 2$) de la série de Balmer en effectuant un calcul relativiste dans le cadre de l'Ancienne Théorie des Quanta. Ceci n'est pas tout à fait un hasard⁴⁷.

L'existence du moment magnétique propre de l'électron étant admise sous la pression des faits expérimentaux, un argument⁴⁸ permet de comprendre que, nécessairement, l'électron possède un moment cinétique *propre*. En effet, la présence d'un terme comme V_{magn} dans le Hamiltonien H entraîne que le vecteur \vec{L} n'est plus une constante du mouvement : comme les différentes composantes de \vec{L} ne commutent pas entre elles, le commutateur de l'une quelconque d'entre elles avec H n'est pas nul. Il en résulte, en l'état actuel des choses, que le moment cinétique d'un système isolé comme l'atome n'est pas une constante du mouvement !

La difficulté se résout d'elle-même si, par analogie avec le cas orbital, on fait l'hypothèse que le moment magnétique intrinsèque est inévitablement associé à un moment cinétique propre, noté \vec{S} , dont l'existence est postulée, et tel que $\vec{\mu} \propto \vec{S}$; dans ces conditions, l'interaction spin-orbite est de la forme :

$$V_{\text{magn}} = A(r) \vec{L} \cdot \vec{S} . \quad (1.143)$$

Ceci étant admis, il est naturel de définir le moment cinétique total, noté \vec{J} , par la combinaison :

$$\vec{J} = \vec{L} + \vec{S} \quad (1.144)$$

et c'est ce moment total dont on attend qu'il soit, lui, une constante du mouvement. En définitive, il est impossible d'admettre l'existence d'un moment magnétique propre de l'électron non lié à un moment cinétique, tout en voulant maintenir la conservation du moment cinétique d'un système isolé.

Montrons de fait que \vec{J} défini par (1.144) est une constante du mouvement quand le Hamiltonien est complété par un terme du genre V_{magn} . On a :

$$\vec{L} \cdot \vec{S} = \frac{1}{2} [(\vec{L} + \vec{S})^2 - \vec{L}^2 - \vec{S}^2] = \frac{1}{2} [\vec{J}^2 - \vec{L}^2 - \vec{S}^2] \quad (1.145)$$

d'où :

$$[\vec{J}, \vec{L} \cdot \vec{S}] = \frac{1}{2} [\vec{J}, \vec{J}^2 - \vec{L}^2 - \vec{S}^2] = -\frac{1}{2} [\vec{J}, \vec{L}^2 + \vec{S}^2] . \quad (1.146)$$

Les moments orbital et de spin commutent entre eux ; il reste :

$$[\vec{J}, \vec{L} \cdot \vec{S}] = -\frac{1}{2} [\vec{L}, \vec{L}^2] - \frac{1}{2} [\vec{S}, \vec{S}^2] = 0 . \quad (1.147)$$

\vec{J} est bien le moment cinétique qui, même en présence d'interaction spin-orbite, est une constante du mouvement – une obligation pour un système isolé.

Étant acquis que c'est bien \vec{J} qui est le moment cinétique de l'électron avec spin en orbite, \vec{L} et \vec{S} deviennent des opérateurs vectoriels : il est en effet facile de vérifier que le commutateur $[\vec{V}, \vec{J}]$, avec $\vec{V} = \vec{L}$ ou \vec{S} ,

⁴⁶La structure hyperfine (impliquant des effets encore plus petits) résulte d'un couplage du même type avec le spin nucléaire ; cette fois c'est le noyau qui, par son spin (donc son moment magnétique), se couple au champ magnétique créé par la rotation des électrons.

⁴⁷Comme on le verra plus loin, section 1.5, le spin est d'origine relativiste.

⁴⁸voir [7] p. 255.

est bien égal à ce qui a été énoncé dans la théorie du moment cinétique pour élire un opérateur à trois composantes en tant qu'opérateur vectoriel⁴⁹.

Avant de poursuivre, quelques rappels et définitions sont utiles, à propos du lien étroit existant entre moment cinétique et moment magnétique. On sait que le moment magnétique d'une boucle de courant de surface S parcourue par un courant d'intensité I est un vecteur normal au circuit, donné par la règle du tire-bouchon et de module IS . En prenant une vision classique, l'atome est un petit circuit constitué par un électron de vitesse v en orbite circulaire autour du noyau. Pour fabriquer un courant, on dispose d'une charge (celle de l'électron) et d'un temps, la période T de rotation de l'électron autour du noyau ; le courant atomique est donc $\frac{|e|}{T}$. Si R est le rayon de l'orbite, cette période vaut $\frac{2\pi R}{v}$ et la surface est πR^2 . Le module du moment magnétique vaut alors⁵⁰ :

$$\mu = \frac{|e|}{(2\pi R/v)} \pi R^2 = \frac{|e|}{2m} mvR \quad . \quad (1.148)$$

On reconnaît le module du moment cinétique orbital, de sorte que la relation vectorielle précise, compte tenu de la règle du tire-bouchon, est :

$$\vec{\mu} = \frac{e}{2m} \vec{L} \quad . \quad (1.149)$$

L'électron étant chargé négativement, les deux vecteurs sont dirigés en sens contraires. Pour une particule de masse M et de charge q , la relation est :

$$\vec{\mu} = \frac{q}{2M} \vec{L} \equiv \gamma \vec{L} \quad , \quad (1.150)$$

ce qui définit le rapport gyromagnétique γ . Comme l'unité fondamentale de moment cinétique est \hbar , il est d'usage, pour l'électron par exemple, d'écrire le moment magnétique résultant de son mouvement orbital sous la forme :

$$\vec{\mu} = \frac{e\hbar}{2m} \frac{\vec{L}}{\hbar} \quad , \quad (1.151)$$

écriture où apparaît le magnéton⁵¹ de Bohr, μ_B :

$$\mu_B = \frac{e\hbar}{2m} = -0.9274 \times 10^{-23} \text{ Joule/Tesla} = -5.789 \times 10^{-5} \text{ eV/Tesla} \quad . \quad (1.152)$$

Alors (1.151) devient :

$$\vec{\mu} = \mu_B \hbar^{-1} \vec{L} \quad . \quad (1.153)$$

On introduit aussi le magnéton nucléaire, μ_N , relatif au proton :

$$\mu_N = \frac{|e|\hbar}{2M_p} = 5.051 \times 10^{-27} \text{ Joule/Tesla} = 3.153 \times 10^{-8} \text{ eV/Tesla} \quad . \quad (1.154)$$

Tout ceci repose sur des notions classiques et se réfère par conséquent à un moment magnétique issu d'un moment cinétique orbital. Étant convaincu de l'existence d'un moment magnétique intrinsèque, on pose naturellement – sur de simples considérations dimensionnelles – des relations analogues entre le moment cinétique de spin et le moment magnétique révélé par l'expérience. Comme le moment magnétique lié au spin ne résulte pas d'une révolution géométrique, il peut apparaître – outre le facteur charge/masse – un facteur numérique que l'on espère d'ordre unité sinon le calibrage dimensionnel serait malhabile. Finalement, pour le moment magnétique de spin, on pose pour l'électron :

$$\vec{\mu}_S = g_e \frac{e}{2m} \vec{S} = g_e \mu_B \hbar^{-1} \vec{S} \equiv \gamma_e \vec{S} \quad . \quad (1.155)$$

Dans (1.155), g_e est donc un nombre pur ($g_e = 2.002319314\dots$) pris égal à 2 sauf cas exceptionnel⁵² ; γ_e est le facteur gyromagnétique de spin et est négatif, tout comme μ_B : le moment magnétique lié au spin et le spin sont

⁴⁹On a évidemment $[\vec{L}, \vec{S}] = 0$, de sorte que \vec{L} (resp. \vec{S}) n'est pas un opérateur vectoriel vis-à-vis de \vec{S} (resp. \vec{L}).

⁵⁰Comme μ est la notation traditionnelle pour un moment magnétique – mais c'est aussi celle pour la masse réduite – gare aux confusions ! Ici, m est la masse de l'électron et on raisonne avec un noyau infiniment massif. Pour le calcul précis conduisant à (1.148), voir les bons ouvrages.

⁵¹On adopte partout la convention par laquelle μ_B est *négatif* ; il semble que ce soit l'usage le plus répandu ([7], [8], *etc.*).

⁵²La théorie de Dirac donne $g_e = 2$ (voir section (1.5)). L'écart fini $g_e - 2$ résulte de corrections électrodynamiques ; son calcul est l'un des enjeux de l'Électrodynamique Quantique, dont le succès à cet égard est proprement fabuleux : les valeurs théoriques et expérimentales coïncident sur une douzaine de chiffres significatifs... On a [9] $g_{th} = 2.002319304402 \pm 6 \times 10^{-11}$, cependant que la valeur expérimentale connue est $g_{exp} = 2.002319304376 \pm 8 \times 10^{-12}$...

donc dirigés en sens contraires. Au total, γ_e est une caractéristique intrinsèque de l'électron. Pour le proton⁵³, de façon analogue :

$$\vec{\mu}_S = g_p \frac{|e|}{2M_p} \vec{S} = g_p \mu_N \hbar^{-1} \vec{S} \equiv \gamma_p \vec{S} \quad (1.156)$$

avec $g_p = 5.59$. Pour le neutron⁵⁴, $g_n = -3.82630$, conférant au neutron un moment magnétique égal à 1,91315 magnéton nucléaire (tout comme l'électron, le neutron a ses deux moments de spin et magnétique opposés). L'existence d'un moment magnétique pour le neutron (globalement *neutre*⁵⁵) peut sembler mystérieuse au vu de la relation classique (1.150) – qu'il vaut mieux dans ce cas lire à l'envers, partant de l'idée que c'est le spin qui est l'attribut intrinsèque fondamental (tout comme une charge, une masse).

C'est le spin, quand il est demi-entier, qui permet d'expliquer le nombre pair de raies Zeeman anormales et le nombre pair d'impacts sur l'écran dans l'expérience de Stern et Gerlach, étant entendu que l'on observe aussi des cas où ces nombres sont impairs, bien qu'il s'agisse toujours d'électrons au sein d'atomes. De fait, pour un atome complexe (et en ne considérant que les électrons) la parité du nombre de "signaux" résulte de l'addition de tous les moments cinétiques, orbitaux et de spin, de tous les électrons. Même si \vec{L} est non-pertinent pour le cas examiné (parce que l'atome est, du point de vue orbital, dans un état sphérique $\vec{L} = 0$), le spin à lui seul peut donner soit des nombres pairs, soit des nombres impairs : tout dépend de la *parité* du nombre d'électrons dans l'atome. La théorie générale du moment cinétique permet d'affirmer que les atomes à nombre pair (resp. impair) d'électrons ont un spin entier (resp. demi-entier)⁵⁶.

1.5 Équation de Dirac et limite de Pauli

Il s'agit de présenter – sous forme élémentaire – quelques arguments, mettant en évidence la nature relativiste du degré de liberté interne que constitue le spin. Dans un premier temps, seul le cas d'une particule *libre* sera considéré, un électron pour fixer les idées. Cette version première de la théorie est due à Dirac et s'appelle historiquement "Théorie de l'électron de Dirac". Par la suite, on examinera le cas de l'électron lié au sein de l'atome d'hydrogène ; ceci permettra, dans la limite faiblement relativiste, d'introduire le Hamiltonien de Pauli et également de fournir la justification détaillée de l'expression de l'interaction spin-orbite donnée en (1.140).

1.5.1 Construction de l'équation de Dirac

Il est raisonnable de croire que la construction de l'équation d'onde relativiste peut être entreprise comme celle de son homologue non-relativiste. Finalement, pour une particule libre non-relativiste, l'équation de Schrödinger :

$$i\hbar \frac{\partial}{\partial t} \bullet = \frac{1}{2m} (-i\hbar \vec{\nabla})^2 \bullet \quad (1.157)$$

s'obtient à partir de l'expression classique :

$$H = \frac{p^2}{2m} , \quad (1.158)$$

en y faisant la substitution :

$$H \longrightarrow i\hbar \frac{\partial}{\partial t} , \quad \vec{p} \longrightarrow -i\hbar \vec{\nabla} . \quad (1.159)$$

⁵³également de spin $S = \frac{1}{2}$, et de masse 938.258 MeV. Selon l'habitude, dire qu'une particule est de spin S c'est donner la valeur du nombre quantique permettant de calculer le carré du module de son moment cinétique intrinsèque. Pour "une particule de spin S ", celui-ci est égal à $S(S+1)\hbar^2$.

⁵⁴également de spin $S = \frac{1}{2}$, mais de masse très légèrement supérieure à celle du proton : $M_n = 939.553$ MeV. Le neutron (libre) est instable (durée de vie de l'ordre de la dizaine de minutes), alors que le proton est stable, pour autant qu'on sache.

⁵⁵La charge du neutron est expérimentalement inférieure à 2×10^{-22} charge élémentaire [10].

⁵⁶Cette affirmation laisse de côté le spin nucléaire. Pour l'hydrogène ordinaire ("léger"), le noyau a un spin $\frac{1}{2}$ (un proton) et, au total, le spin de l'atome vaut soit 0, soit 1 ; ceci permet, si l'on veut, de distinguer deux "espèces" d'hydrogène. Cette classification n'est utile que pour les problèmes où la structure hyperfine joue un rôle important, ou lorsque l'aspect boson *versus* fermion est pertinent.

De la même façon, pour le cas relativiste, on part du Hamiltonien classique (*i. e.* non quantique) pour une particule libre de masse m :

$$H^2 = p^2 c^2 + m^2 c^4 ; \quad (1.160)$$

En y faisant la substitution (1.159), on obtient :

$$-\hbar^2 \frac{\partial^2}{\partial t^2} = -\hbar^2 c^2 \Delta + m^2 c^4 \iff \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta + \frac{m^2 c^2}{\hbar^2} \right) \Psi = 0 . \quad (1.161)$$

$\frac{\hbar}{mc}$ qui apparaît dans cette équation est la seule longueur disponible avec la constante de Planck, la masse de l'électron et la vitesse de la lumière ; c'est la longueur d'onde Compton de l'électron et vaut environ 0.024 Å. L'équation (1.161) s'appelle équation de Klein - Gordon. Elle conduit à des difficultés de deux ordres :

1. C'est une équation du 2^{ème} ordre par rapport au temps, alors que jusqu'à présent on a admis que la connaissance de l'état à $t = 0$ suffit à déterminer l'état à tout instant ultérieur. La généralisation relativiste ne devrait pas exiger un changement de l'ordre de l'équation fondamentale, changement qui procède par tout ou rien (comme une symétrie, présente ou absente). On peut envisager des développements en $1/c$ pour retrouver la limite non-relativiste. Par continuité⁵⁷, on ne voit pas bien comment on pourrait passer de l'ordre 2 à l'ordre 1.

Toutefois, cette difficulté peut être contournée, au moins formellement. La résolution proposée ici est d'ailleurs éclairante et se révélera féconde pour régler d'autres difficultés ultérieures. En effet, une équation du second ordre peut en fait résulter de la contraction de deux équations plus "fondamentales", toutes deux d'ordre 1. Prenons l'exemple trivial d'un oscillateur (classique) à une dimension. L'équation dynamique est :

$$m\ddot{x} = -kx . \quad (1.162)$$

Elle est du second ordre, mais résulte des deux équations de Hamilton :

$$\dot{x} = \frac{p}{m} , \quad \dot{p} = -kx , \quad (1.163)$$

qui sont toutes deux du 1^{er} ordre. On peut d'ailleurs les écrire sous forme matricielle :

$$\frac{d}{dt} \begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{m} \\ -k & 0 \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} \quad (1.164)$$

qui constitue bien une équation du 1^{er} ordre, portant sur une description dynamique "complexifiée" (bien sûr, le nombre de conditions initiales n'a pas changé, il en faut toujours deux). De la même façon, on peut réécrire l'équation de Klein - Gordon en affirmant que l'état du système est complètement décrit par une fonction d'onde à *deux* composantes Ψ_a et Ψ_b . En effet, (1.161) s'écrit aussi :

$$\left(1 + \frac{i\hbar}{mc^2} \frac{\partial}{\partial t} \right) \left(1 - \frac{i\hbar}{mc^2} \frac{\partial}{\partial t} \right) \Psi = \frac{\hbar^2}{m^2 c^2} \Delta \Psi , \quad (1.165)$$

ce qui suggère d'introduire les deux combinaisons :

$$\Psi_a = \Psi + \frac{i\hbar}{mc^2} \frac{\partial \Psi}{\partial t} , \quad \Psi_b = \Psi - \frac{i\hbar}{mc^2} \frac{\partial \Psi}{\partial t} , \quad (1.166)$$

d'où résulte :

$$\Psi = \frac{1}{2} (\Psi_a + \Psi_b) , \quad \frac{\partial \Psi}{\partial t} = \frac{mc^2}{2i\hbar} (\Psi_a - \Psi_b) . \quad (1.167)$$

Dès lors, (1.165) peut s'écrire de deux façons :

$$\left(1 - \frac{i\hbar}{mc^2} \frac{\partial}{\partial t} \right) \Psi_a = \frac{\hbar^2}{2m^2 c^2} \Delta (\Psi_a + \Psi_b) , \quad \left(1 + \frac{i\hbar}{mc^2} \frac{\partial}{\partial t} \right) \Psi_b = \frac{\hbar^2}{2m^2 c^2} \Delta (\Psi_a + \Psi_b) \quad (1.168)$$

⁵⁷Bien sûr, la limite quantique non-relativiste pourrait être singulière, mais ce n'est pas le cas ; le "petit paramètre" de la Mécanique Quantique (\hbar !) apparaît dans le terme de plus haut degré dans l'équation de Schrödinger, faisant de la limite classique (non-quantique et non-relativiste) une limite singulière. Rien de tel ne se produit en ce qui concerne la limite non-relativiste.

soit, sous la forme matricielle :

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \Psi_a \\ \Psi_b \end{bmatrix} = \begin{bmatrix} mc^2 - \frac{\hbar^2}{2m} \Delta & -\frac{\hbar^2}{2m} \Delta \\ +\frac{\hbar^2}{2m} \Delta & -mc^2 + \frac{\hbar^2}{2m} \Delta \end{bmatrix} \begin{bmatrix} \Psi_a \\ \Psi_b \end{bmatrix} . \quad (1.169)$$

On retrouve ainsi formellement une équation du 1^{er} ordre par rapport au temps – il reste toutefois que deux conditions initiales sont toujours à prescrire. En ce sens, la difficulté n'est résolue que *formellement*.

2. Il existe en réalité une autre difficulté, beaucoup plus grave. L'équation (1.161) étant donnée, et si l'on maintient l'interprétation usuelle de Ψ en tant qu'amplitude de probabilité (pour avoir une probabilité fabriquée avec $|\Psi|^2$), il convient de vérifier la cohérence du tout en établissant une équation de conservation, qui doit *découler* de (1.161). En gardant comme courant l'expression symétrisée que l'on peut construire avec Ψ , Ψ^* et la vitesse $\frac{1}{m}\vec{p} = -\frac{i\hbar}{m}\vec{\nabla}$, on pose :

$$\vec{j} = \frac{\hbar}{2im} (\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^*) . \quad (1.170)$$

En ce qui concerne la densité ρ , que l'on attend quadratique en Ψ et Ψ^* , on constate qu'en fait $\Psi\Psi^*$ ne peut pas convenir puisque l'équation de conservation introduit une dérivée première en temps, et que l'équation de Klein - Gordon contient une dérivée *seconde*. La densité doit donc être fabriquée à partir de Ψ et de $\frac{\partial\Psi}{\partial t}$; une démarche par essai et erreur montre que si l'on pose :

$$\rho = \frac{i\hbar}{2mc^2} \left(\Psi^* \frac{\partial\Psi}{\partial t} - \Psi \frac{\partial\Psi^*}{\partial t} \right) , \quad (1.171)$$

alors l'équation de Klein - Gordon (1.161) prend la la forme souhaitée d'une équation de conservation :

$$\frac{\partial\rho}{\partial t} + \text{div} \vec{j} = 0 . \quad (1.172)$$

Le seul ennui est que la densité ρ définie en (1.171) ... n'a aucune raison d'être positive ! Les équations ne peuvent être retenues qu'au prix d'une réinterprétation, qui a été faite par Pauli et Weisskopf ([11], p. 764). Fondamentalement, en multipliant \vec{j} et ρ par une charge, ceux-ci représentent alors un courant et une densité de charge. Comme, de toute façon, le nombre de particules ne saurait être fixé au sein d'une théorie relativiste (possibilité de création de paires), \vec{j} et ρ peuvent être associés à la différence entre le nombre de charges positives et le nombre de charges négatives ([12], p. 195). La théorie devient alors une théorie à une charge (la charge totale), et perd son statut de théorie à une particule.

De toute façon, l'équation de Klein - Gordon n'est pas satisfaisante pour l'électron, en dépit de la présence, dans sa version (1.169), des deux combinaisons Ψ_a et Ψ_b (qui peuvent faire penser aux deux composantes d'un spineur de rang 2). En effet, si l'on examine la limite non-relativiste ([12], p. 199), on trouve que l'une des composantes est d'ordre 0 en v/c , alors que l'autre est d'ordre 1 : ces deux composantes sont donc loin de jouer un rôle symétrique, alors que, même dans une théorie pseudo-relativiste à la Pauli, les deux composantes du spineur doivent être du même ordre de grandeur. Au total, la limite non-relativiste produit finalement une fonction d'onde à *une* composante, et ne convient donc pas pour l'électron. Au mieux, l'équation de Klein - Gordon convient pour une particule de spin *nul*.

S'agissant de construire une équation relativiste, il faut en fait s'ancrer sur l'obligation de faire jouer des rôles symétriques aux coordonnées d'espace et au temps. En particulier, si l'on maintient l'idée d'une équation d'ordre 1 en temps, les coordonnées doivent également figurer par des dérivées du 1^{er} ordre. Quand on revient à la forme "linéarisée" de Klein - Gordon (1.169), on devine ce qu'il faut tenter de faire. Dans la matrice au second membre, apparaît le Laplacien (dérivées spatiales du 2^{ème} ordre) ; pour se ramener à des dérivées du 1^{er} ordre, on peut utiliser la même astuce que pour le temps dans le cas de l'oscillateur harmonique, et à nouveau "complexifier" la représentation. Ainsi apparaît la suggestion forte de doubler la dimension de la matrice 2×2 , ce qui revient à dire que chaque combinaison Ψ_a et Ψ_b est en fait elle-même un *doublet*, composantes de l'élément d'un espace vectoriel à deux dimensions, produisant finalement une fonction d'onde à *quatre* composantes ; en

retenant cette hypothèse de travail, celles-ci sont désormais notées Ψ_μ , $\mu = 1, 2, 3, 4$:

$$\Psi = \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{bmatrix} . \quad (1.173)$$

Dès lors, on posera que la densité de probabilité est :

$$\rho(\vec{r}, t) = \sum_{\mu=1}^4 |\Psi_\mu(\vec{r}, t)|^2 . \quad (1.174)$$

Ceci étant retenu, l'équation (du premier ordre en temps et espace) la plus simple que l'on puisse former est une forme linéaire de $i\hbar \frac{\partial}{\partial t}$ et de $\vec{p} = -i\hbar \vec{\nabla}$; par ailleurs, pour une vitesse nulle, l'énergie doit se réduire à un terme contenant l'énergie dite de repos mc^2 . Au total, on peut *a priori* poser la forme suivante :

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}, t) = (c \vec{\alpha} \cdot \vec{p} + \beta mc^2) \Psi(\vec{r}, t) \equiv H_D \Psi(\vec{r}, t) . \quad (1.175)$$

H_D est le Hamiltonien de Dirac ; $\vec{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$ est un ensemble de trois objets agissant sur des vecteurs à quatre composantes : ce sont des matrices 4×4 , tout comme β . $\vec{\alpha}$ et β sont sans dimension physique ; ces matrices mélangent entre elles les quatre composantes Ψ_μ . En présence d'un champ de forces statique associé à l'énergie potentielle $V(\vec{r})$, la généralisation de (1.175) est⁵⁸ :

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}, t) = [c \vec{\alpha} \cdot \vec{p} + V(\vec{r}) \mathbf{1} + \beta mc^2] \Psi(\vec{r}, t) . \quad (1.176)$$

L'équation (1.175) s'écrit aussi :

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \vec{\alpha} \cdot \vec{\nabla} + \frac{imc}{\hbar} \beta \right) \Psi(\vec{r}, t) = 0 , \quad (1.177)$$

soit, sous forme plus explicite :

$$\left[\frac{1}{c} \frac{\partial}{\partial t} + \left(\sum_{u=x,y,z} \alpha_u \frac{\partial}{\partial u} \right) + \frac{imc}{\hbar} \beta \right] \Psi(\vec{r}, t) = 0 . \quad (1.178)$$

Toute la question est maintenant de trouver les matrices $\vec{\alpha}$ et β . Il convient bien sûr de maintenir le lien avec l'expression classique (1.160), qui est vraie pour *toute* particule et qui produit l'équation de Klein - Gordon (1.161) par simple substitution des opérateurs aux grandeurs classiques⁵⁹. Une façon de maintenir le contact est de faire des manipulations sur l'opérateur entre crochets dans (1.178) pour lui donner l'allure de l'opérateur apparaissant dans (1.161). Au total, il s'agit de trouver une marche à suivre permettant d'écrire des relations entre les matrices $\vec{\alpha}$ et β cherchées et, finalement, de les déterminer.

Pour exhiber à partir de (1.178) un premier membre qui "ressemble" à Klein - Gordon, il faut fabriquer des dérivées secondes en temps et en espace (images opératoriennes de la relation classique (1.160), qui est quadratique en H et \vec{p}), sans pour autant faire apparaître de dérivées croisées temps - espace ; compte tenu de la forme $A + B$ de l'opérateur au premier membre de (1.178), un peu de réflexion montre qu'il faut introduire l'opérateur de la forme $A - B$, soit :

$$\frac{1}{c} \frac{\partial}{\partial t} - \left(\sum_{v=x,y,z} \alpha_v \frac{\partial}{\partial v} \right) - \frac{imc}{\hbar} \beta \quad (1.179)$$

⁵⁸Dans la suite, $\mathbf{1}$ note tantôt la matrice identité 4×4 , tantôt la matrice identité 2×2 ; le contexte permet de lever toute ambiguïté.

⁵⁹Pour cette raison, l'équation de Klein - Gordon apparaît comme un point de passage obligé entre la relation classique (*i. e.* non quantique) et l'équation cherchée

et le faire agir sur le premier membre de (1.178). On obtient ainsi :

$$\left[\frac{1}{c} \frac{\partial}{\partial t} - \left(\sum_v \alpha_v \frac{\partial}{\partial v} \right) - \frac{imc}{\hbar} \beta \right] \left[\frac{1}{c} \frac{\partial}{\partial t} + \left(\sum_u \alpha_u \frac{\partial}{\partial u} \right) + \frac{imc}{\hbar} \beta \right] \Psi(\vec{r}, t) = 0 . \quad (1.180)$$

Il suffit maintenant d'expliciter les termes au premier membre, et de jouer avec les indices muets de sommation pour obtenir⁶⁰ :

$$\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{1}{2} \sum_{u,v} (\alpha_u \alpha_v + \alpha_v \alpha_u) \frac{\partial^2}{\partial u \partial v} - \frac{imc}{\hbar} \sum_u (\alpha_u \beta + \beta \alpha_u) \frac{\partial}{\partial u} + \left(\frac{mc}{\hbar} \right)^2 \beta^2 \right] \Psi(\vec{r}, t) = 0 . \quad (1.181)$$

Alors, cette équation se réduit à celle de Klein - Gordon à condition que les relations suivantes soient vérifiées :

$$\alpha_u \alpha_v + \alpha_v \alpha_u = 2 \delta_{uv} \iff \alpha_u^2 = \mathbf{1} \quad (u = x, y, z) \quad \text{et} \quad \alpha_u \alpha_v + \alpha_v \alpha_u = 0 \quad (u \neq v) \quad (1.182)$$

et :

$$\beta^2 = \mathbf{1} , \quad \alpha_u \beta + \beta \alpha_u = 0 \quad (u = x, y, z) . \quad (1.183)$$

Les α_u et β sont bien des matrices : les relations ci-dessus ne sauraient être satisfaites par des nombres ordinaires⁶¹.

Les relations (1.182) et (1.183) étant admises, il n'est pas difficile de montrer que les matrices 4×4 cherchées peuvent s'écrire par blocs 2×2 comme suit⁶² :

$$\vec{\alpha} = \begin{bmatrix} \mathbf{0} & \vec{\sigma} \\ \vec{\sigma} & \mathbf{0} \end{bmatrix} , \quad \beta = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \quad (1.184)$$

où les $\vec{\sigma}$ sont les matrices de Pauli et où $\mathbf{1}$ est ici la matrice identité 2×2 , $\mathbf{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. L'ensemble des quatre matrices données par (1.184) est appelé "représentation-standard", étant entendu qu'il existe une infinité de couples $(\vec{\alpha}, \beta)$ vérifiant (1.182) et (1.183), déductibles de (1.184) par une transformation unitaire quelconque. En tout cas, dans cette représentation, H_D s'écrit :

$$H_D = \begin{bmatrix} mc^2 \mathbf{1} & c \vec{\sigma} \cdot \vec{p} \\ c \vec{\sigma} \cdot \vec{p} & -mc^2 \mathbf{1} \end{bmatrix} \equiv mc^2 \beta + c \vec{\alpha} \cdot \vec{p} . \quad (1.185)$$

Notons que le carré de H_D n'est autre que la matrice $(m^2 c^4 + \vec{p}^2 c^2) \mathbf{1}$, en résultat du fait que l'on a imposé aux quatre Ψ_μ de satisfaire (1.161).

On va maintenant établir le premier résultat spectaculaire de la théorie de Dirac, à savoir l'apparition naturelle du spin sans apport supplémentaire. Il s'agit maintenant de montrer que l'opérateur :

$$\vec{\Sigma} = \begin{bmatrix} \vec{\sigma} & \mathbf{0} \\ \mathbf{0} & \vec{\sigma} \end{bmatrix} \quad (1.186)$$

fournit le degré de liberté de spin (très précisément, le spin est égal à $\frac{\hbar}{2} \vec{\Sigma}$). La base de l'argument consiste à montrer que le vecteur :

$$\vec{J} = \vec{L} + \frac{\hbar}{2} \vec{\Sigma} , \quad (1.187)$$

qui se veut être un moment cinétique puisqu'il incorpore déjà le moment cinétique orbital, est effectivement une constante du mouvement lorsque le Hamiltonien est celui qui figure au second membre de (1.176), soit :

$$H_D = [c \vec{\alpha} \cdot \vec{p} + V(\vec{r}) \mathbf{1} + \beta mc^2] . \quad (1.188)$$

⁶⁰ *A priori* les matrices α_u ne commutent pas entre elles, alors que $\frac{\partial^2}{\partial u \partial v} = \frac{\partial^2}{\partial v \partial u}$.

⁶¹ On connaît cependant des nombres qui obéissent à ce type d'algèbre, ou qui, en tout cas anticommulent : ce sont les nombres (variables) de Grassmann, souvent notés ξ_i et tels que $\xi_i \xi_{i'} + \xi_{i'} \xi_i = 0$. Ces nombres, d'apparence exotique, apparaissent dans la formulation en intégrale fonctionnelle (à la Feynman) des systèmes de fermions.

⁶² $\mathbf{0}$ note un bloc matriciel 2×2 identiquement nul.

Le terme potentiel est proportionnel à la matrice identité : il commute donc avec les matrices $\vec{\alpha}$ et β . Étant à symétrie sphérique, il commute avec \vec{L} . Il en résulte :

$$[\vec{J}, H_D] = [\vec{J}, c\vec{\alpha}\cdot\vec{p} + \beta mc^2] . \quad (1.189)$$

Calculons d'abord le commutateur de \vec{L} avec la partie "cinétique" de H_D . Dans l'espace à 4 dimensions, le moment cinétique orbital est simplement :

$$\vec{L} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} (\vec{r} \times \vec{p}) , \quad (1.190)$$

puisque \vec{p} agit de la même façon sur chaque composante ($\vec{p}\Psi_\mu = -i\hbar\vec{\nabla}\Psi_\mu$). Il en résulte :

$$[\vec{L}, H_D] = [(\vec{r} \times \vec{p}) \mathbf{1}, c\vec{\alpha}\cdot\vec{p} + \beta mc^2] = [(\vec{r} \times \vec{p}) \mathbf{1}, c\vec{\alpha}\cdot\vec{p}] . \quad (1.191)$$

Par exemple, on a :

$$[L_z, H_D] = c[xp_y - yp_x, \sum_u \alpha_u p_u] = i\hbar c(\alpha_x p_y - \alpha_y p_x) \equiv i\hbar c(\vec{\alpha} \times \vec{p})_z , \quad (1.192)$$

d'où :

$$[\vec{L}, H_D] = i\hbar c(\vec{\alpha} \times \vec{p}) . \quad (1.193)$$

Par ailleurs⁶³ :

$$[\vec{\Sigma}, H_D] = \left[\begin{bmatrix} \vec{\sigma} & \mathbf{0} \\ \mathbf{0} & \vec{\sigma} \end{bmatrix}, \begin{bmatrix} mc^2 \mathbf{1} & c\vec{\sigma}\cdot\vec{p} \\ c\vec{\sigma}\cdot\vec{p} & -mc^2 \mathbf{1} \end{bmatrix} \right] = \begin{bmatrix} \mathbf{0} & c[\vec{\sigma}, \vec{\sigma}\cdot\vec{p}] \\ c[\vec{\sigma}, \vec{\sigma}\cdot\vec{p}] & \mathbf{0} \end{bmatrix} . \quad (1.194)$$

On a, par exemple :

$$[\sigma_x, \vec{\sigma}\cdot\vec{p}] = [\sigma_x, \sum_u \sigma_u p_u] = 2i(p_y \sigma_z - p_z \sigma_y) \equiv 2i(\vec{p} \times \vec{\sigma})_x , \quad (1.195)$$

d'où :

$$[\vec{\Sigma}, H_D] = 2ic \begin{bmatrix} \mathbf{0} & \vec{p} \times \vec{\sigma} \\ \vec{p} \times \vec{\sigma} & \mathbf{0} \end{bmatrix} \equiv 2ic \vec{p} \times \vec{\alpha} . \quad (1.196)$$

Prenant en compte (1.193) et (1.196), le commutateur de \vec{J} avec le Hamiltonien de Dirac est :

$$[\vec{J}, H_D] \equiv [\vec{L} + \frac{\hbar}{2} \vec{\Sigma}, H_D] = i\hbar c(\vec{\alpha} \times \vec{p} + \vec{p} \times \vec{\alpha}) = 0 . \quad (1.197)$$

Ce résultat permet bien d'interpréter $\frac{\hbar}{2} \vec{\Sigma}$ comme l'expression du moment cinétique intrinsèque de spin de l'électron : ajouté vectoriellement au moment cinétique orbital, le moment cinétique résultant \vec{J} est une constante du mouvement. La théorie de Dirac engendre le spin sous la forme :

$$\vec{S} = \frac{\hbar}{2} \vec{\Sigma} = \frac{\hbar}{2} \begin{bmatrix} \vec{\sigma} & \mathbf{0} \\ \mathbf{0} & \vec{\sigma} \end{bmatrix} , \quad \vec{J} = \vec{L} + \vec{S} . \quad (1.198)$$

La même conclusion est préservée en présence d'un champ $V(r)$ à symétrie sphérique (\vec{J} commute avec le Hamiltonien de (1.176)).

Ainsi, la théorie relativiste de Dirac contient en soi le spin électronique. En outre, comme annoncé plus haut, elle fournit la valeur $g_e = 2$, comme on le voit en prenant la limite non-relativiste pour une particule libre en présence d'un champ magnétique. Cette limite ([11], p. 807, et voir la Remarque plus loin) fournit le Hamiltonien :

$$H_{c \rightarrow \infty} = \frac{1}{2m} (\vec{p} - e\vec{A})^2 - \frac{e\hbar}{2m} \vec{\sigma}\cdot\vec{B} + eU . \quad (1.199)$$

⁶³Penser à utiliser la multiplication par bloc des matrices.

appelé Hamiltonien de Pauli. Le couplage magnétique est $-\frac{e}{m} \vec{S} \cdot \vec{B}$ fait ressortir le moment magnétique $\vec{\mu} = \frac{e}{m} \vec{S}$, soit :

$$\vec{\mu} = 2 \frac{e\hbar}{2m} \vec{S} \iff g_e = 2 . \quad (1.200)$$

Ceci n'est pas strictement correct ; l'écart $g_e - 2 = 0.002319314\dots$ résulte de corrections "radiatives" et ne peut apparaître que dans une théorie traitant matière et photons sur un pied d'égalité, les englobant dans un même cadre quantique (c'est l'Électrodynamique Quantique). Quoi qu'il en soit, la présence de $\vec{\sigma}$ dans (1.199) montre que le Hamiltonien de Pauli agit sur une fonction d'onde à *deux composantes* ; ainsi, la limite faiblement relativiste reproduit exactement le Hamiltonien Zeeman (1.261) – à $g_e - 2$ près, évidemment.

1.5.2 États stationnaires d'un électron libre

Pour un électron libre, il n'est pas difficile de trouver les états stationnaires en théorie de Dirac ; en posant :

$$\Psi(\vec{r}, t) = e^{\frac{i}{\hbar} Et} \psi(\vec{r}) , \quad (1.201)$$

(1.175) devient :

$$E \psi(\vec{r}) = H_D \psi(\vec{r}) \quad (1.202)$$

et constitue l'équation aux valeurs et fonctions propres de Dirac pour un électron libre. En l'absence de champ de forces, l'impulsion est une constante du mouvement (\vec{p} commute avec H_D), et on peut chercher des solutions sous la forme d'ondes planes, correspondant à une impulsion déterminée, c'est-à-dire fonctions propres de \vec{p} :

$$\psi(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} C(\vec{k}) . \quad (1.203)$$

Dit autrement, l'invariance par translation résultant de l'absence d'une force assure que les générateurs des translations (les trois composantes de \vec{p}) sont des constantes du mouvement. Dans (1.203), $\psi(\vec{r})$ et $C(\vec{k})$ sont des matrices-colonne à quatre éléments ; sous forme explicite :

$$\psi_\mu(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} C_\mu(\vec{k}) \quad (\mu = 1, 2, 3, 4) . \quad (1.204)$$

Bien évidemment, C_μ est indépendant de \vec{r} . En reportant dans (1.202) et en utilisant :

$$\vec{p} e^{i\vec{k} \cdot \vec{r}} = -i\hbar \vec{\nabla} e^{i\vec{k} \cdot \vec{r}} = \hbar \vec{k} e^{i\vec{k} \cdot \vec{r}} \quad (1.205)$$

($\hbar \vec{k}$ est donc la valeur propre de \vec{p}), on obtient :

$$(c \hbar \vec{k} \cdot \vec{\alpha} + \beta mc^2) C(\vec{k}) = E C(\vec{k}) , \quad (1.206)$$

soit, écrivant par blocs 2×2 la matrice 4×4 de H_D :

$$\begin{bmatrix} mc^2 \mathbf{1} & c \hbar \vec{k} \cdot \vec{\sigma} \\ c \hbar \vec{k} \cdot \vec{\sigma} & -mc^2 \mathbf{1} \end{bmatrix} \begin{bmatrix} C_1(\vec{k}) \\ C_2(\vec{k}) \\ C_3(\vec{k}) \\ C_4(\vec{k}) \end{bmatrix} = E \begin{bmatrix} C_1(\vec{k}) \\ C_2(\vec{k}) \\ C_3(\vec{k}) \\ C_4(\vec{k}) \end{bmatrix} \iff H_D C(\vec{k}) = E C(\vec{k}) . \quad (1.207)$$

Pour trouver commodément l'énergie E , on commence par appliquer deux fois H_D à C :

$$H_D^2 C(\vec{k}) = E^2 C(\vec{k}) ; \quad (1.208)$$

en calculant (toujours par blocs) le carré de la matrice figurant dans (1.207), on trouve facilement la forme matricielle⁶⁴ de (1.208) :

$$\begin{bmatrix} (mc^2)^2 \mathbf{1} + \hbar^2 c^2 (\vec{k} \cdot \vec{\sigma})(\vec{k} \cdot \vec{\sigma}) & \mathbf{0} \\ \mathbf{0} & (mc^2)^2 \mathbf{1} + \hbar^2 c^2 (\vec{k} \cdot \vec{\sigma})(\vec{k} \cdot \vec{\sigma}) \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = E^2 \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} . \quad (1.209)$$

⁶⁴Dans la suite, la dépendance en \vec{k} des C_μ est sous-entendue pour alléger les notations.

Compte tenu de $(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) = (\vec{A} \cdot \vec{B}) \mathbf{1} + i \vec{\sigma} \cdot (\vec{A} \times \vec{B})$, $(\vec{k} \cdot \vec{\sigma})(\vec{k} \cdot \vec{\sigma})$ est égal à $k^2 \mathbf{1}$. Il en résulte que le carré de l'énergie est :

$$E^2 = (mc^2)^2 + \hbar^2 k^2 c^2 . \quad (1.210)$$

L'énergie est donc donnée (au signe près) dès que $\hbar \vec{k}$ est fixé, une autre conséquence du fait que, la particule étant libre, son impulsion $\vec{p} = \hbar \vec{k}$ est une constante du mouvement. En désignant par p le module de l'impulsion, on retrouve alors la relation :

$$E^2 = (mc^2)^2 + (pc)^2 , \quad (1.211)$$

d'où les deux formes possibles de la relation de dispersion :

$$E = \pm \sqrt{(mc^2)^2 + (pc)^2} \equiv \pm E_p , \quad (1.212)$$

E_p désignant la racine carrée positive. Ainsi, deux classes de solutions apparaissent, les unes d'énergie positive, supérieure ou égale à $E_{p=0} = mc^2$, les autres d'énergie négative, bornées supérieurement par $-E_{p=0} = -mc^2$. Les deux classes de solutions sont donc séparées⁶⁵ l'une de l'autres par un *gap* égal à $2mc^2$; l'interprétation de ce fait sera donnée plus loin.

Il s'avère commode de paramétriser comme suit :

$$p = mc \sinh \phi , \quad E = \pm mc^2 \cosh \phi , \quad (1.213)$$

et alors la limite non relativiste $p \ll mc$ correspond à $\phi \rightarrow 0$.

Comme l'espace est de dimension 4, chaque valeur de l'énergie (à \vec{k} donné et pour une classe) est dégénérée deux fois. Il existe donc une autre observable qui, jointe à H_D , constitue un ECOG ; ses vecteurs propres seront trouvés en formant les bonnes combinaisons linéaires d'états de *même* énergie.

Pour simplifier, on peut choisir l'axe Oz le long de la valeur de l'impulsion $\hbar \vec{k}$ fixée. Dès lors, (1.207) prend la forme explicite plus simple :

$$\begin{bmatrix} mc^2 \mathbf{1} & pc \sigma_z \\ pc \sigma_z & -mc^2 \mathbf{1} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = E \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} . \quad (1.214)$$

Comme chaque bloc (diagonal ou non-diagonal) est une matrice diagonale, ce système 4×4 se dissocie en deux systèmes 2×2 découplés :

$$\begin{cases} (mc^2 - E) C_1 + pc C_3 = 0 \\ pc C_1 - (mc^2 + E) C_3 = 0 \end{cases} , \quad \begin{cases} (mc^2 - E) C_2 - pc C_4 = 0 \\ -pc C_2 - (mc^2 + E) C_4 = 0 \end{cases} . \quad (1.215)$$

Ces deux systèmes ont la *même* équation caractéristique en E , dont les racines sont données par les expressions (1.212) ; cette identité n'est autre que la dégénérescence annoncée plus haut. Les valeurs propres étant déterminées, il reste à trouver les vecteurs propres. De toute évidence, $\vec{\Sigma} \cdot \frac{\vec{p}}{p} = \Sigma_z$ commute avec H_D , puisque (voir 1.194) :

$$[\Sigma_z, H_D] = \left[\begin{bmatrix} \sigma_z & \mathbf{0} \\ \mathbf{0} & \sigma_z \end{bmatrix}, \begin{bmatrix} mc^2 \mathbf{1} & c p \sigma_z \\ c p \sigma_z & -mc^2 \mathbf{1} \end{bmatrix} \right] = \begin{bmatrix} \mathbf{0} & c [\sigma_z, p \sigma_z] \\ c [\sigma_z, p \sigma_z] & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} . \quad (1.216)$$

σ_z est ici l'observable qui participe à l'ECOC, et représente l'angle entre le spin \vec{S} et la direction de propagation. L'équation propre pour Σ_z s'écrit :

$$\begin{bmatrix} \sigma_z & \mathbf{0} \\ \mathbf{0} & \sigma_z \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = \sigma \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} \quad (\sigma = \pm 1) . \quad (1.217)$$

⁶⁵ Pour une particule de masse nulle (le photon par exemple), on a simplement $E = \pm \hbar kc$ et le *gap* est nul.

Cette équation montre qu'une combinaison linéaire arbitraire des composantes 1 et 3 correspond à la valeur propre $\sigma = +1$ de Σ_z , qu'une combinaison arbitraire des composantes 2 et 4 correspond à la valeur propre $\sigma = -1$. Les combinaisons linéaires qui sont *aussi* propres de H_D s'obtiennent en diagonalisant les deux systèmes (1.215), celui de gauche correspondant à $\Sigma_z = +1$, celui de droite à $\Sigma_z = -1$. Au total, en désignant par ε le signe de l'énergie ($E = \varepsilon E_p$, voir (1.212)), les vecteurs propres seront donc étiquetés, à impulsion $\hbar\vec{k}$ donnée, par le couple de nombres quantiques (ε, σ) et notés $\Psi_{\varepsilon, \sigma}$.

Pour les solutions à énergie positive, un calcul facile donne :

$$\Psi_{+, +1} = \frac{1}{\sqrt{\cosh \phi}} \begin{bmatrix} \cosh \frac{\phi}{2} \\ 0 \\ \sinh \frac{\phi}{2} \\ 0 \end{bmatrix}, \quad \Psi_{+, -1} = \frac{1}{\sqrt{\cosh \phi}} \begin{bmatrix} 0 \\ \cosh \frac{\phi}{2} \\ 0 \\ -\sinh \frac{\phi}{2} \end{bmatrix}. \quad (1.218)$$

De même, pour les énergies négatives, on trouve :

$$\Psi_{-, +1} = \frac{1}{\sqrt{\cosh \phi}} \begin{bmatrix} -\sinh \frac{\phi}{2} \\ 0 \\ \cosh \frac{\phi}{2} \\ 0 \end{bmatrix}, \quad \Psi_{-, -1} = \frac{1}{\sqrt{\cosh \phi}} \begin{bmatrix} 0 \\ \sinh \frac{\phi}{2} \\ 0 \\ \cosh \frac{\phi}{2} \end{bmatrix}. \quad (1.219)$$

Comme il se doit, les deux états $|\Psi_{\varepsilon, \pm 1}\rangle$ et $|\Psi_{-\varepsilon, \pm 1}\rangle$ sont orthogonaux. Pour chaque énergie, l'observable Σ_z peut prendre les deux valeurs ± 1 ; la valeur propre $+1$ représente un état où la composante S_z du spin est dans le même sens que l'impulsion $\hbar\vec{k}$, -1 étant associé à l'état où le spin est "en arrière" vis-à-vis de l'impulsion. Au total, on imagine une rotation autour de la direction de propagation qui est tantôt dans un sens, tantôt dans l'autre : c'est pourquoi on parle d'états à *hélicité* positive et négative, respectivement.

L'identification de $\vec{\Sigma}$ avec le spin (au facteur $\frac{\hbar}{2}$ près) est d'autant plus convaincante que dans la limite non-relativiste $pc \ll mc^2$ (i.e. $\phi \ll 1$) et pour les solutions à énergie positive ($\varepsilon = +$), il vient :

$$\Psi_{+, +1} \simeq \begin{bmatrix} 1 \\ 0 \\ \frac{p}{2mc} \\ 0 \end{bmatrix}, \quad \Psi_{+, -1} \simeq \begin{bmatrix} 0 \\ 1 \\ 0 \\ \frac{-p}{2mc} \end{bmatrix} \quad (p \ll mc). \quad (1.220)$$

Dans cette limite, chaque état propre de (H_D, Σ_z) a une composante d'ordre zéro en $\frac{p}{mc}$ et une autre du premier ordre. Au total, à la limite stricte $c = +\infty$, la dimension de l'espace vectoriel passe de quatre à deux⁶⁶ :

$$\Psi_{+, +1} \rightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \Psi_{+, -1} \rightarrow \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (c \rightarrow +\infty). \quad (1.221)$$

On désigne par "petites" composantes les composantes qui tendent vers zéro dans cette limite, les autres étant naturellement appelées "grandes" composantes. Dans la limite non-relativiste, on obtient une particule d'énergie positive décrite par une fonction d'onde à deux composantes, lesquelles se distinguent par la valeur propre de Σ_z , c'est-à-dire de S_z .

Le même phénomène se produit pour les solutions à énergie négative :

$$\Psi_{-, +1} \simeq \begin{bmatrix} \frac{-p}{2mc} \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \Psi_{-, -1} \simeq \begin{bmatrix} 0 \\ \frac{p}{2mc} \\ 0 \\ 1 \end{bmatrix} \quad (p \ll mc). \quad (1.222)$$

⁶⁶On se souvient que ce n'est pas le cas pour l'équation de Klein - Gordon, ce qui rend celle-ci inapte à décrire l'électron.

Maintenant, ce sont les 3^{me} et 4^{me} composantes qui survivent à la limite $c \rightarrow +\infty$:

$$\Psi_{-,+1} \rightarrow \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \Psi_{-,-1} \rightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (c = +\infty). \quad (1.223)$$

L'énergie négative pose la question de la signification de ces états.

L'existence en toute généralité de solutions à énergie négative est à première vue problématique puisque l'on s'attend en toute hypothèse à trouver que l'énergie de l'électron est bornée inférieurement par mc^2 . L'interprétation proposée par Dirac est schématiquement la suivante et entérine l'idée suivant laquelle, dans toute théorie relativiste, le nombre de particules ne saurait être fixé à cause de la possibilité de création de paires. Il apparaît en bout de course que la théorie "à un électron" formulée ci-dessus se doit d'être complétée afin de prendre réellement tout son sens.

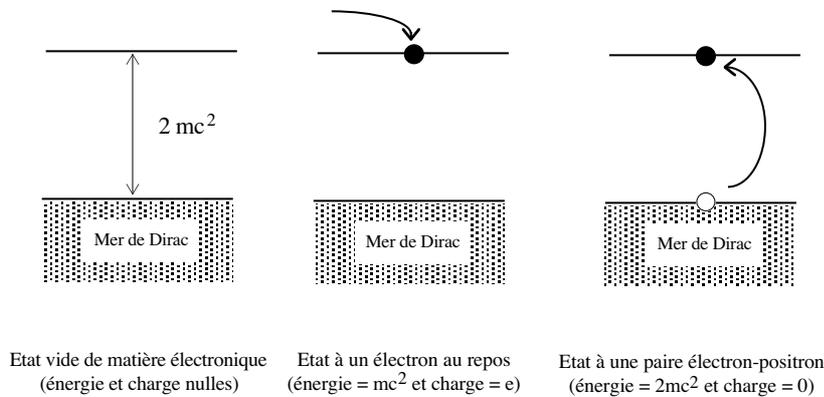


Figure 1.2: Illustration schématique de l'interprétation de Dirac.

Dirac postule l'existence d'un nombre (infini !) de particules occupant tous les états d'énergie négative⁶⁷ ; cette "bande" d'énergie remplie, appelée mer de Dirac, culmine en énergie à $-mc^2$ et est séparée par un *gap* égal à $2mc^2$ de l'état de plus basse énergie positive. Le système composé de la mer de Dirac remplie et d'aucune particule dans la bande positive est *défini* comme constituant le vide de matière électronique ; son énergie est conventionnellement prise égale à zéro⁶⁸, tout comme sa charge (ce qui est nettement moins intuitif).

Un état à un électron est formé en plaçant une particule supplémentaire dans un état d'énergie positive, ce qui coûte une énergie au moins égale à $mc^2 \leq mc^2 [1 - (v/c)^2]^{-1/2}$ (il faut créer un électron) ; en disposant N particules dans N états d'énergie positive, on fabrique un système à N électrons au sens usuel. D'une façon générale, les particules situées dans la bande d'énergie positive sont interprétées comme étant les électrons "ordinaires", la mer étant pleine par ailleurs (voir fig. 1.2).

Ce schéma a permis à Dirac de prévoir, en 1928, l'existence du positron, antiparticule de l'électron (même masse, même spin, charge opposée), découvert quelques années plus tard (en 1933) par C. D. Anderson. Le scénario initialement imaginé par Dirac est le suivant. L'état vide a – conventionnellement – une énergie et une charge nulles. Pour l'exciter – sans ajouter de particule de l'extérieur –, il faut lui fournir une énergie au moins égale à celle du *gap*, prenant une particule d'énergie inférieure à $-mc^2$ et en la hissant dans la bande d'énergie positive⁶⁹ : ceci coûte donc une énergie au moins égale à $2mc^2$. Après cette opération, on se retrouve

⁶⁷Ceci suppose que chaque état à une particule situé dans cette bande peut être occupé par un nombre fini de particules, en pratique une seule compte tenu du spin (Principe de Pauli).

⁶⁸Seules les *différences d'énergie* ont un sens physique.

⁶⁹Cette image d'une infinité d'états remplis comme constituant l'état fondamental est très féconde et réapparaît un peu partout dans la théorie du Problème à N -corps, dans des contextes très différents. Par exemple, l'état fondamental d'un isolant ou d'un semi-conducteur intrinsèque de bande interdite E_g correspond à un remplissage total (par des électrons) de la bande de valence – l'équivalent de la mer négative de Dirac. Les états excités se forment en prenant un électron de cette bande pour le hisser, en lui fournissant au moins l'énergie E_g , dans la bande de conduction. De la sorte, on crée une lacune positive dans la bande de valence et une particule chargée négativement dans la bande de conduction : c'est une paire électron-trou.

avec un électron d'une part, avec une lacune dans la mer de Dirac d'autre part. En vertu de la conservation de la charge, il faut attribuer au "trou" de la mer de Dirac une charge positive, exactement opposée à celle de l'électron : c'est un positron de Dirac, l'antiparticule de l'électron. La création d'une paire électron - positron coûte au moins deux fois l'énergie de repos $2mc^2$, soit environ 1.2 MeV.

Remarque

Il est possible de donner une autre justification du fait que l'opérateur $\vec{\Sigma}$ (1.186) apporte le degré de spin, suivant la relation précise :

$$\vec{S} = \frac{\hbar}{2} \vec{\Sigma} . \quad (1.224)$$

Pour cela, introduisons un champ électromagnétique $\vec{\mathcal{E}}, \vec{\mathcal{B}}$ dérivant du potentiel (\vec{A}, ϕ) et repartons du Hamiltonien de Dirac en y faisant la substitution habituelle :

$$E \longrightarrow E - e\phi \quad \vec{p} \longrightarrow \vec{p} - e\vec{A} . \quad (1.225)$$

L'équation de Dirac (1.175) prend alors la forme :

$$[E - e\phi - c\vec{\alpha}(\vec{p} - e\vec{A}) - \beta mc^2] \Psi = 0 , \quad (1.226)$$

où E et \vec{p} sont les opérateurs :

$$E = i\hbar \frac{\partial}{\partial t} \quad \vec{p} = -i\hbar \vec{\nabla} . \quad (1.227)$$

Multiplions maintenant (1.226) à gauche par $[E - e\phi + c\vec{\alpha}(\vec{p} - e\vec{A}) - \beta mc^2]$. Il vient :

$$\left[(E - e\phi)^2 - c^2 [\vec{\alpha}(\vec{p} - e\vec{A})]^2 - m^2 c^4 \right] \Psi = \left[c(E - e\phi)\vec{\alpha}(\vec{p} - e\vec{A}) - c\vec{\alpha}(\vec{p} - e\vec{A})(E - e\phi) \right] \Psi . \quad (1.228)$$

Compte tenu de l'expression des matrices $\vec{\alpha}$, il est facile d'établir la relation :

$$(\vec{V} \cdot \vec{\alpha})(\vec{W} \cdot \vec{\alpha}) = \vec{V} \cdot \vec{W} + i\vec{\Sigma} \cdot (\vec{V} \times \vec{W}) . \quad (1.229)$$

De la sorte, et prenant en compte la relation $\vec{\mathcal{B}} = \vec{\nabla} \times \vec{A}$, on voit que le premier membre de (1.228) est égal à :

$$\left[(E - e\phi)^2 - c^2 (\vec{p} - e\vec{A})^2 - m^2 c^4 + e\hbar c^2 \vec{\Sigma} \cdot \vec{\mathcal{B}} \right] \Psi . \quad (1.230)$$

Par ailleurs, des manipulations simples sur le second membre faisant usage de (1.227) et également de $\vec{\mathcal{E}} = -\frac{\partial \vec{A}}{\partial t} - \vec{\nabla} \phi$ le transforment en :

$$+ie\hbar c \vec{\alpha} \cdot \vec{\mathcal{E}} \Psi . \quad (1.231)$$

Finalement, l'équation de Dirac en présence du champ s'écrit :

$$\left[(E - e\phi)^2 - m^2 c^4 - c^2 (\vec{p} - e\vec{A})^2 + e\hbar c^2 \vec{\Sigma} \cdot \vec{\mathcal{B}} - ie\hbar c \vec{\alpha} \cdot \vec{\mathcal{E}} \right] \Psi = 0 . \quad (1.232)$$

Jusqu'à présent, aucune approximation n'a été faite. Examinons maintenant la limite faiblement relativiste en posant :

$$E = E_{\text{fr}} + mc^2 , \quad (1.233)$$

avec l'hypothèse supplémentaire $E_{\text{fr}}, e\phi \ll mc^2$. Les deux premiers termes de (1.232) se simplifient en :

$$(E - e\phi)^2 - m^2 c^4 \simeq 2mc^2 (E_{\text{fr}} - e\phi) , \quad (1.234)$$

ce qui permet de réécrire (1.232) comme suit :

$$E_{\text{fr}} \Psi = \left[\frac{1}{2m} (\vec{p} - e\vec{A})^2 + e\phi - \frac{e\hbar}{2m} \vec{\Sigma} \cdot \vec{\mathcal{B}} + i \frac{e\hbar}{2mc} \vec{\alpha} \cdot \vec{\mathcal{E}} \right] \Psi . \quad (1.235)$$

Ψ est toujours un vecteur à 4 composantes. Comme il s'agit maintenant de la limite faiblement relativiste et compte tenu du développement en $\frac{v}{c}$ déjà effectué, il suffit de considérer les deux premières composantes

(les “grandes” composantes). En désignant par Ψ_g la projection de Ψ dans le sous-espace 2×2 de ces grandes composantes, et puisque $\vec{\alpha}$ mélange les grandes et les petites composantes (celles-ci étant d'ordre $\frac{v}{c}$ par rapport aux premières), l'équation (1.235) devient à cet ordre :

$$E_{\text{fr}} \Psi_g = \left[\frac{1}{2m} (\vec{p} - e\vec{A})^2 + e\phi - \frac{e\hbar}{2m} \vec{\sigma} \cdot \vec{\mathcal{B}} \right] \Psi_g , \quad (1.236)$$

où $\vec{\sigma}$ est maintenant l'ensemble des trois matrices de Pauli (le terme $i(e\hbar)/(2mc)\vec{\alpha} \cdot \vec{\mathcal{E}}$ ainsi négligé produit l'interaction spin-orbite, voir ci-dessous). Les deux premiers termes du second membre sont familiers ; le troisième est un couplage du genre $-\vec{\mu} \cdot \vec{\mathcal{B}}$ et permet bien d'identifier le moment magnétique lié au spin en posant :

$$\vec{S} = \frac{\hbar}{2} \vec{\sigma} , \quad (1.237)$$

justifiant l'interprétation de $\vec{\Sigma}$. À nouveau, par comparaison avec (1.155), on observe que la théorie de Dirac fournit exactement $g_e = 2$. Comme déjà mentionné, l'écart $g_e - 2$ résulte de corrections radiatives, non-incluses par la théorie de Dirac.

Le Hamiltonien au second membre de (1.236) s'appelle le Hamiltonien de Pauli :

$$H_{\text{Pauli}} = \frac{1}{2m} (\vec{p} - e\vec{A})^2 + e\phi - \frac{e}{m} \vec{S} \cdot \vec{\mathcal{B}} , \quad (1.238)$$

compte tenu de (1.237). Ce résultat justifie la procédure pragmatique utilisée en théorie franchement non-relativiste, qui consiste à prendre le Hamiltonien ordinaire pour une particule chargée dans un champ électromagnétique et à rajouter *à la main* le couplage magnétique $-\vec{\mu} \cdot \vec{\mathcal{B}}$ résultant du spin.

Enfin, montrons comment obtenir l'expression correcte de l'opérateur spin-orbite V_{magn} , (1.140). En repartant du Hamiltonien de Dirac, en l'absence de champ extérieur mais pour une particule liée par le potentiel à symétrie sphérique $V(r)$, on a :

$$E \Psi = [c \vec{\alpha} \cdot \vec{p} + \beta mc^2 + V(\vec{r})] \Psi . \quad (1.239)$$

En désignant par Ψ_g et Ψ_p les couples de grandes et petites composantes, et en introduisant à nouveau E_{fr} comme en (1.233), l'équation (1.239) s'explique en deux équations couplées :

$$E_{\text{fr}} \Psi_g = c \vec{\sigma} \cdot \vec{p} \Psi_p + (\beta mc^2 + V) \Psi_g , \quad (1.240)$$

$$E_{\text{fr}} \Psi_p = c \vec{\sigma} \cdot \vec{p} \Psi_g + (-\beta mc^2 + V) \Psi_p . \quad (1.241)$$

Maintenant, par élimination de Ψ_p , on obtient :

$$(E_{\text{fr}} + 2mc^2) \Psi_g = \frac{1}{2m} \vec{\sigma} \cdot \vec{p} \frac{1}{1 + \frac{E_{\text{fr}} - V}{2mc^2}} \vec{\sigma} \cdot \vec{p} \Psi_g . \quad (1.242)$$

Le développement de la fraction au second membre et quelques manipulations algébriques⁷⁰ conduisent à :

$$E_{\text{fr}} \Psi_g = \left(\frac{\vec{p}^2}{2m} - \frac{p^4}{8m^3 c^2} + V - \frac{\hbar^2}{4m^2 c^2} \frac{dV}{dr} \frac{\partial}{\partial r} + \frac{1}{2m^2 c^2} \frac{1}{r} \frac{dV}{dr} \vec{L} \cdot \vec{S} \right) \Psi_g , \quad (1.244)$$

où $\vec{L} = \vec{r} \times \vec{p}$ et $\vec{S} = \frac{\hbar}{2} \vec{\sigma}$. Les deuxième et quatrième termes sont des corrections relativistes aux énergies cinétique et potentielle, d'ordre α^2 ; le terme en p^4 est une simple “correction de masse”, l'autre étant sans équivalent classique. Le dernier terme donne, comme souhaité, l'expression convenable de V_{magn} donnée en (1.140).

⁷⁰Utiliser notamment :

$$(\vec{\sigma} \cdot \vec{\nabla} V) (\vec{\sigma} \cdot \vec{p}) = (\vec{\nabla} V) \vec{p} + i \vec{\sigma} \cdot [\vec{\nabla} V \times \vec{p}] , \quad (1.243)$$

et la sphéricité du potentiel V ($\vec{\nabla} V = r^{-1} (dV/dr) \vec{r}$, etc)

1.6 Effet d'un champ statique externe : effet Stark et effet Zeeman

Il s'agit ici de décrire quantitativement les modifications du spectre de l'atome d'hydrogène quand il est soumis à des champs externes statiques (on revient au traitement non-relativiste). Comme on le verra par les ordres de grandeurs, sauf conditions extrêmes, la théorie des perturbations est ici l'outil pertinent⁷¹. L'effet Stark est l'éclatement (*splitting*) des sous-niveaux atomiques par application d'un champ électrique et l'effet Zeeman est le phénomène analogue provoqué par un champ magnétique.

Effet Stark

Pour l'effet Stark, il suffit ici d'examiner la situation la plus simple où seules les interactions électrostatiques sont considérées (on laisse ici de côté les couplages magnétiques du genre spin-orbite et donc toutes les complications liées à la structure fine). Dans ces conditions, le Hamiltonien H_0 de l'atome d'hydrogène en l'absence de champ appliqué est :

$$H_0 = \frac{\vec{p}^2}{2m} - \frac{e'^2}{r} . \quad (1.245)$$

Les états non-perturbés (sans spin) sont notés $|n, l, m\rangle$, avec la signification habituelle pour les nombres n, l, m . L'énergie à l'ordre zéro est :

$$E_n^{(0)} = -\frac{\alpha^2}{2n^2} mc^2 = -\frac{e'^2}{2n^2 a_0} . \quad (1.246)$$

En choisissant l'axe Oz le long du champ de module \mathcal{E} , la perturbation est :

$$V = -e\mathcal{E}z \quad (1.247)$$

puisque la force électrique agissant sur l'électron est $\vec{F} = -\vec{\nabla}V = e\vec{\mathcal{E}}$.

Avant de commencer les calculs, examinons brièvement les conditions de validité de la méthode de perturbation. Une condition nécessaire est :

$$\frac{|\langle n, l, m | V | n', l', m' \rangle|}{|E_n^{(0)} - E_{n'}^{(0)}|} \ll 1 \iff |\langle n, l, m | e\mathcal{E}z | n', l', m' \rangle| \ll |E_n^{(0)} - E_{n'}^{(0)}| . \quad (1.248)$$

L'ordre de grandeur du premier membre est $|e\mathcal{E}a_0|$; le second est de l'ordre de $\frac{e'^2}{a_0}$. La condition (1.248) se réécrit comme suit :

$$\mathcal{E} \ll \frac{|e|}{4\pi\epsilon_0 a_0^2} \equiv \mathcal{E}_{\text{atomique}} . \quad (1.249)$$

Le sens de cette dernière condition est évident : pour que l'on puisse parler de perturbation, il est nécessaire que le champ électrique appliqué de l'extérieur soit très petit devant le champ intra-atomique ; ce dernier est gigantesque (environ 10^9 V/cm). La condition *nécessaire* d'application de la méthode de perturbation est donc satisfaite pour les champs ordinaires facilement réalisables.

Avant de mettre en œuvre la théorie des perturbations, quelques considérations de symétrie sont utiles. H_0 est à symétrie sphérique et invariant par renversement du temps. Si V possède cette dernière symétrie, en revanche le couplage avec le champ électrique brise la symétrie sphérique : seule subsiste une symétrie de révolution autour de l'axe du champ, Oz . En outre, V est invariant dans toute réflexion par rapport à un plan contenant l'axe Oz ; dans cette opération, L_z change de signe⁷². Les états propres communs⁷³ à (H, \vec{L}^2, L_z) , $|\tau, l, m\rangle$, qui ne diffèrent que par le signe de la valeur propre associée à L_z auront donc la même énergie. L'énergie en présence du champ est donc fonction de τ et de $|m|$ et on peut donc prévoir que la levée de dégénérescence ne sera que partielle.

⁷¹De fait, il s'agit de deux applications classiques de la théorie des perturbations.

⁷²tout comme la composante L_X située dans le plan de réflexion, cependant que la composante perpendiculaire au plan, L_Y , est inaltérée.

⁷³De toute évidence, \vec{L}^2 est encore une constante du mouvement.

Une dernière remarque concernant la symétrie : V est changé en son opposé par parité⁷⁴ :

$$V' \equiv \Pi V \Pi^\dagger = -V . \quad (1.250)$$

Par ailleurs :

$$\Pi |n, l, m\rangle = (-1)^l |n, l, m\rangle . \quad (1.251)$$

L'élément de matrice $\langle n, l, m | z | n', l', m' \rangle$ est multiplié par $(-1)^{l+l'+1}$ par parité ; il est donc nul si l et l' sont de même parité :

$$\langle n, l, m | z | n', l', m' \rangle = 0 \quad \text{si } l \text{ et } l' \text{ de même parité} . \quad (1.252)$$

Examinons maintenant quantitativement l'effet du champ sur les premiers niveaux. En ce qui concerne le fondamental (non-dégénéré), $|1, 0, 0\rangle$, il n'y a pas de correction au premier ordre puisque $\langle 1, 0, 0 | V | 1, 0, 0 \rangle = 0$. La correction du second ordre est⁷⁵ :

$$\Delta E_{\text{fond}}^{(2)} = - \sum_{n=2}^{+\infty} \sum_{l,m} \frac{|\langle 1, 0, 0 | -e\mathcal{E}z | n, l, m \rangle|^2}{E_n^{(0)} - E_1^{(0)}} < 0 , \quad (1.253)$$

où les $E_n^{(0)}$ sont données par (1.246). Grâce à une astuce très ingénieuse, il est possible de sommer cette série ([5], p. 266) ; on trouve :

$$\varepsilon_{\text{fond } 2} = -\frac{1}{2} \alpha \mathcal{E}^2 \quad \alpha = 18\pi\epsilon_0 a_0^3 . \quad (1.254)$$

L'écriture⁷⁶ $-\frac{1}{2}\alpha\mathcal{E}^2$ est choisie pour rappeler qu'il s'agit d'un effet de polarisabilité, la correction d'énergie quadratique en champ devant pouvoir s'écrire $-\int_0^{\mathcal{E}} d\mathcal{E}' \alpha \mathcal{E}'$, une fois admis que le moment induit est proportionnel au champ.

Considérons maintenant le premier groupe d'états excités $n = 2$, définissant un sous-espace dégénéré de dimension égale à 4 en l'absence du spin (qui ici ne joue aucun rôle) ; il convient donc de mettre en œuvre la technique de perturbation spécifique à ce cas. Notons que ce sous-espace est couplé au fondamental, puisque l'élément de matrice $\langle 1, 0, 0 | z | 2, 1, 0 \rangle$ est différent de zéro ; comme on le sait, la méthode de perturbation pour un niveau dégénéré néglige ce type de couplage (au plus bas ordre) et ne retient que la *projection* de l'opérateur V dans le sous-espace $n = 2$. Il faut donc calculer les éléments d'une matrice 4×4 ; en fait, compte tenu des considérations de symétrie ci-dessus, un seul élément de matrice est non-nul, c'est $\langle 2, 0, 0 | V | 2, 1, 0 \rangle$. Un calcul simple d'intégrale donne :

$$\langle 2, 0, 0 | V | 2, 1, 0 \rangle = 3e\mathcal{E}a_0 \equiv -v < 0 . \quad (1.255)$$

Il en résulte que les valeurs propres de l'opérateur projeté sont $\varepsilon_{21} = 0, \pm v$; les vecteurs propres s'en déduisent :

$$|+v\rangle = \frac{1}{\sqrt{2}} (|2, 0, 0\rangle + |2, 1, 0\rangle) \quad |-v\rangle = \frac{1}{\sqrt{2}} (-|2, 0, 0\rangle + |2, 1, 0\rangle) . \quad (1.256)$$

Ces deux états correspondent à la levée de dégénérescence. Deux autres états restent dégénérés en gardant, à cet ordre, l'énergie $E_2^{(0)}$, ce sont les $|2, 1, \pm 1\rangle$.

En utilisant la notation spectroscopique, les vecteurs propres (1.256) s'écrivent :

$$|+v\rangle = \frac{1}{\sqrt{2}} (|2S\rangle + |2P_0\rangle) \quad |-v\rangle = \frac{1}{\sqrt{2}} (-|2S\rangle + |2P_0\rangle) . \quad (1.257)$$

Ainsi, le champ électrique mélange deux états de symétrie différente vis-à-vis du moment cinétique orbital⁷⁷. Ceci a des conséquences importantes, en particulier pour la durée de vie de certains états excités (contamination

⁷⁴Le champ extérieur n'appartient pas au système quantique et se factorise de tous les éléments de matrice.

⁷⁵Ceci constitue le résultat élémentaire : on fait comme si le spectre ne contenait que des états liés. En toute rigueur, la sommation discrète doit être complétée par une intégrale sommant sur tous les états non-liés d'énergie positive.

⁷⁶La notation α est traditionnelle pour la polarisabilité. Ce n'est pas ici la constante de structure fine !

⁷⁷Un tel mélange s'appelle hybridation, terme que l'on retrouve également en Chimie à propos des orbitales atomiques ou moléculaires.

Stark), puisque les règles de sélection E1 sont précisément établies sur la base d'une symétrie "pure" vis-à-vis de \vec{L} .

Ainsi, l'état $|2S\rangle$ a une durée de vie extraordinairement longue, de l'ordre de la seconde⁷⁸, au contraire de l'état $|2P\rangle$ dont la durée de vie est de l'ordre de 10^{-9} s. La raie la plus "rouge" de la série de Lyman⁷⁹, appelée L_α , correspond précisément à la transition $|2P\rangle \rightarrow |2S\rangle$. Le calcul ci-dessus montre que le champ électrique, parce qu'il contamine l'état $2S$ par l'état $2P$, a un effet spectaculaire sur la durée de vie de l'état métastable : tout champ électrique, même petit, est donc capable de détruire le caractère métastable. Si l'atome est préparé dans l'état $2S$ et qu'un champ électrique est ajouté, volontairement ou non, la métastabilité est anéantie et l'atome peut rayonner presque normalement. Ce phénomène porte le nom de *quenching* de métastabilité.

Remarque

Le calcul qui vient d'être fait est tout à fait correct, dans le sens où il rend très bien compte, qualitativement et quantitativement, de l'effet Stark tel qu'il est observé. Il est utile de remarquer que cette adéquation fait fi d'une "irrégularité" commise depuis le début : la perturbation est ici singulière et l'hypothèse de l'existence d'une série de perturbation convergente – cruciale pour l'établissement des formules donnant les diverses corrections – est prise en défaut. On peut s'en convaincre sur des arguments purement physiques, esquissés ci-dessous.

Dès que l'atome est soumis à un champ électrique, si faible soit-il, il n'y a plus d'états liés au sens strict du terme. En effet, l'énergie potentielle totale de l'électron est V_{tot} :

$$V_{\text{tot}} = -\frac{e^2}{r} - e\mathcal{E}z . \quad (1.258)$$

En supposant le champ $\vec{\mathcal{E}}$ dirigé dans le sens des z positifs, et compte tenu du signe de la charge e , on a :

$$\lim_{z \rightarrow +\infty} V_{\text{tot}} = +\infty \quad \lim_{z \rightarrow -\infty} V_{\text{tot}} = -\infty \quad (1.259)$$

avec toujours $\lim_{|z| \rightarrow 0} V_{\text{tot}} = -\infty$. Un potentiel ayant de tels comportements n'a pas d'états liés ; ceci traduit la possibilité pour l'électron de passer par effet-tunnel de n'importe quel état localisé à un état dissocié. Dit autrement, supposons l'électron préparé, en l'absence de champ, dans l'état fondamental d'énergie $E_1 \simeq -13.6$ eV. Cet état est stable, son énergie (négative) est inférieure à la valeur limite du potentiel Coulombien à l'infini, l'électron est piégé et reste lié. En revanche, dès que l'on branche le champ, le même électron voit des états de *continuum* accessible à énergie constante : sa probabilité de passage tunnel devient finie, quoique petite. Ceci est l'expression quantique d'un phénomène banal : un champ électrique peut ioniser un atome.

Bien sûr, toute la question est de savoir combien vaut cette probabilité d'ionisation. Plus le champ est faible, plus la barrière-tunnel est épaisse ; la probabilité de passage par effet-tunnel a une dépendance exponentielle par rapport à l'épaisseur⁸⁰ : pour des champs faibles, cette probabilité sera exponentiellement petite, donnant un temps moyen d'ionisation exponentiellement grand. Sans mettre des nombres plus ou moins précis dans le problème⁸¹, on se doute que le calcul ci-dessus est validé par le fait que des temps macroscopiques sont des temps "infinis" vis-à-vis des échelles de temps pertinentes lors d'une expérience de spectroscopie atomique menée dans les conditions ordinaires.

Effet Zeeman

À nouveau, on choisit l'axe Oz le long du champ magnétique statique, noté $\vec{\mathcal{B}}$. On conçoit qu'une description cohérente exige de prendre en compte tous les effets magnétiques simultanément et c'est pourquoi le Hamiltonien

⁷⁸Ceci tient au fait que, compte tenu de la symétrie orbitale des deux états initial et final, la transition $2S \rightarrow 1S$ se fait par un processus à deux photons dont la probabilité est très faible. Cette durée de vie justifie que l'on déclare l'état $2S$ métastable.

⁷⁹toute entière dans l'U.V.

⁸⁰d'où l'extraordinaire sensibilité du microscope à effet-tunnel.

⁸¹Si on tient absolument à mettre des nombres : avec un champ de 10^4 V/m, l'épaisseur L de la barrière à l'énergie $E_1 \simeq -13.6$ eV est à peu près égale à 3 mm, une longueur astronomique à l'échelle atomique. Si on admet une dépendance exponentielle de la probabilité, elle ne peut être que $e^{-L/(\text{une autre longueur du problème})}$; l'autre longueur ne peut être que a_0 , ce qui donne une probabilité de l'ordre de $e^{-6 \times 10^7}$, autant dire zéro !

d'ordre zéro H_0 doit ici contenir le couplage spin-orbite⁸². Ceci étant, H_0 s'écrit :

$$H_0 = \frac{\vec{p}^2}{2m} - \frac{e'^2}{r} + \xi(r) \vec{L} \cdot \vec{S} . \quad (1.260)$$

La fonction $\xi(r)$ apparaît dans (1.140). En présence du champ \vec{B} , et en négligeant le terme diamagnétique, ce Hamiltonien doit être complété par V :

$$H_{\text{Zeeman}} = -\frac{e}{2m} (L_z + g_e S_z) \mathcal{B} \quad (g_e = 2.0023 \dots \simeq 2) , \quad (1.261)$$

conformément à ce qui résulte de la limite faiblement relativiste de Pauli. Il ne fait aucun doute que si le champ est très faible, le couplage H_{Zeeman} est très petit devant le terme spin-orbite, même si ce dernier est lui-même très inférieur aux deux premiers termes de H_0 ; c'est cette situation que l'on va traiter et qui porte précisément le nom d'effet Zeeman⁸³.

Il est facile d'obtenir l'ordre de grandeur du champ magnétique pour lequel on bascule d'un régime à l'autre. Le rapport des deux termes satisfait (voir (1.142)) :

$$\frac{H_{\text{Zeeman}}}{V_{\text{magn}}} \sim \frac{\frac{|e|\hbar}{m}}{\alpha^2 |E_{n=1}|} \mathcal{B} \sim \alpha^{-2} \frac{|e|a_0^2}{\hbar} \mathcal{B} \sim 10^{-2} \mathcal{B}_{\text{tesla}} , \quad (1.262)$$

où l'expression intermédiaire vient de $|E_{n=1}| \sim \frac{e'^2}{a_0}$. Il est clair que l'effet Paschen - Back se manifeste seulement pour des champs très intenses.

Le choix de se cantonner à la description de l'effet Zeeman étant fait, les états propres à l'ordre zéro sont ceux de H_0 et se notent $|n, j, m_j, l, s\rangle$, communs à $(H_0, \vec{J}^2, J_z, \vec{L}^2, \vec{S}^2)$. Hors dégénérescence accidentelle, chaque niveau est donc dégénéré $2j + 1$ fois ($m_j = -j, -j + 1, \dots, j - 1, +j$) et c'est à nouveau la théorie pour un niveau dégénéré qu'il faut appliquer : considérer exclusivement la matrice de H_{Zeeman} dans le sous-espace auquel on s'intéresse, oublier le reste et diagonaliser. Pour trouver la matrice de H_{Zeeman} dans ce sous-espace, deux méthodes peuvent être suivies :

- H_{Zeeman} s'exprime à l'aide de L_z et de S_z , donc sa matrice sur la base $\{|n, l, m_l, m_s\rangle\}$ se détermine facilement. Il reste ensuite à utiliser les relations entre cette base et la base $\{|n, j, m_j, l, s\rangle\}$ – qui impliquent les coefficients de Clebsch - Gordan – pour en déduire la matrice de H_{Zeeman} sur la base propre de H_0 .
- On peut aussi invoquer le théorème de Wigner - Eckart, d'après lequel, à l'intérieur du sous-espace $\{|n, j, m, l, s\rangle\}$, $\vec{L} + g_e \vec{S}$ est un opérateur vectoriel proportionnel à \vec{J} :

$$\vec{L} + g_e \vec{S} = \frac{\langle (\vec{L} + g_e \vec{S}) \cdot \vec{J} \rangle}{\hbar^2 j(j+1)} \vec{J} \equiv g_{lsj} \vec{J} . \quad (1.263)$$

g_{lsj} désigne un nombre pur qui s'exprime à l'aide des nombres quantiques l, s et j . En prenant $g_e = 2$, le calcul donne précisément :

$$g_{lsj} = \frac{1}{2j(j+1)} [3j(j+1) - l(l+1) + s(s+1)] . \quad (1.264)$$

De (1.263) et (1.261), il résulte immédiatement que la matrice du H_{Zeeman} projeté est déjà diagonale. Les corrections au premier ordre pour l'énergie se lisent sur la diagonale et sont donc :

$$\varepsilon_{nj1} = -g_{lsj} \frac{eB}{2m} m_j \hbar \equiv -g_{lsj} m_j \mu \equiv g_{lsj} m_j \hbar \omega_L , \quad \omega_L = \frac{|e|\hbar}{2m} . \quad (1.265)$$

⁸²et tous les autres effets relativistes du même ordre afin d'avoir des résultats quantitatifs cohérents, comme expliqué plus haut. Par souci de simplicité formelle, ces derniers sont sous-entendus dans la suite.

⁸³L'autre cas où V_{magn} est petit devant le couplage avec \vec{B} donne lieu à l'effet dit Paschen - Back, qui correspond donc à un champ magnétique intense. Il est bien clair que l'on passe continûment d'un cas à l'autre en augmentant graduellement le champ extérieur.

Le champ magnétique lève complètement la dégénérescence et fait éclater chaque niveau atomique en $2j + 1$ niveaux équidistants, l'écart entre deux niveaux consécutifs étant proportionnel à l'intensité \mathcal{B} du champ appliqué. Pour l'atome d'hydrogène avec son unique électron de spin $\frac{1}{2}$, j est forcément demi-entier et chaque niveau engendre un nombre pair de sous-niveaux Zeeman.

Cette levée complète résulte du fait que le champ magnétique, au contraire du champ électrique, brise l'invariance par renversement du temps : deux états se distinguant par le signe de m_j ont des énergies opposées par rapport à l'énergie en champ nul⁸⁴.

1.7 Interaction avec un champ électromagnétique (description semi-classique)

On va calculer⁸⁵ au premier ordre les probabilités de transition induites dans un atome (à un seul électron, de masse μ) par un champ électromagnétique considéré comme extérieur⁸⁶ et représenté par une simple onde plane monochromatique polarisée. Le but ultime est de montrer, le plus simplement possible, l'existence de règles de sélection gouvernant les intensités des raies des spectres atomiques.

Pour une telle onde plane, le potentiel scalaire peut toujours être pris égal à zéro⁸⁷. Les deux champs sont :

$$\vec{\mathcal{E}}(\vec{r}, t) = \mathcal{E} \vec{a} \cos(\vec{k} \cdot \vec{r} - \omega t) , \quad \vec{\mathcal{B}}(\vec{r}, t) = \mathcal{B} \vec{b} \cos(\vec{k} \cdot \vec{r} - \omega t) , \quad (1.266)$$

\vec{a} et \vec{b} sont les polarisations des deux champs, $(\vec{b}, \vec{n} = \frac{\vec{k}}{k}, \vec{a})$ est un repère orthonormé direct (en particulier $\vec{k} \cdot \vec{a} = 0$) ; dans la suite, on suppose l'axe Ox choisi le long de \vec{b} , Oz selon \vec{a} et donc Oy le long du vecteur d'onde \vec{k} . Les modules des champs sont reliés par :

$$\mathcal{E} = \mathcal{B}c , \quad (1.267)$$

c désignant la vitesse de la lumière dans le vide. On vérifie facilement que le potentiel-vecteur :

$$\vec{A}(\vec{r}, t) = A_0 \vec{a} \sin(\vec{k} \cdot \vec{r} - \omega t) , \quad A_0 = \frac{\mathcal{E}}{\omega} = \frac{\mathcal{B}}{k} , \quad \mathcal{E} = \mathcal{B}c \quad (1.268)$$

restitue bien les deux champs (1.266) par les relations $\vec{\mathcal{E}} = -\frac{\partial \vec{A}}{\partial t}$ et $\vec{\mathcal{B}} = \vec{\nabla} \times \vec{A}$. Le Hamiltonien se forme comme d'habitude ; complété à la main par l'interaction entre le champ externe et le spin électronique – mais en négligeant les termes de structure fine, pour simplifier – il s'écrit :

$$H = \frac{1}{2\mu} (\vec{p} - e\vec{A})^2 + V(\vec{r}) - g_e \frac{e}{2\mu} \vec{S} \cdot \vec{\mathcal{B}} . \quad (1.269)$$

$V(\vec{r})$ est le potentiel intra-atomique et $g_e \simeq 2$. En développant l'expression (1.269) et en négligeant le terme diamagnétique quadratique en \vec{A} , on trouve :

$$H = \left[\frac{\vec{p}^2}{2\mu} + V(\vec{r}) \right] - \left[\frac{e}{2\mu} (\vec{A} \cdot \vec{p} + \vec{p} \cdot \vec{A}) + g_e \frac{e}{2\mu} \vec{S} \cdot \vec{\mathcal{B}} \right] \equiv H_0 + V(t) . \quad (1.270)$$

H_0 a pour vecteurs et énergies propres $|\psi_n\rangle$ et E_n . $V(t)$ est la somme de deux termes ; comme $\vec{\nabla} \cdot \vec{A} \propto \vec{k} \cdot \vec{a} = 0$, $\vec{A} \cdot \vec{p} = \vec{p} \cdot \vec{A}$ et on peut écrire :

$$V = V_1 + V_2 , \quad \text{avec : } V_1 = -\frac{e}{\mu} \vec{A} \cdot \vec{p} , \quad V_2 = -g_e \frac{e}{2\mu} \vec{S} \cdot \vec{\mathcal{B}} . \quad (1.271)$$

⁸⁴Le renversement du temps $t \rightarrow -t$ change le signe de la vitesse, donc inverse le moment cinétique, d'où $m_j \rightarrow -m_j$. Comme il se doit, ceci revient (en l'absence de spin) à remplacer la fonction d'onde par sa complexe conjuguée.

⁸⁵Voir [8], complément AXIII.

⁸⁶donc non-quantifié.

⁸⁷[8], Appendice III, 4-b- α .

Ces deux termes ne sont pas du même ordre de grandeur ; en effet :

$$\frac{\langle V_2 \rangle}{\langle V_1 \rangle} \sim \frac{\langle \vec{S} \rangle \cdot \vec{B}}{\vec{A} \cdot \langle \vec{p} \rangle} \sim \frac{\hbar k A_0}{A_0 p} = \frac{\hbar k}{p} \sim k a_0 \quad (1.272)$$

où on a posé, en ordre de grandeur : $p \sim \hbar/a_0$. Dans le domaine optique ($\lambda \sim 5000 \text{ \AA}$), $k a_0 \sim (1/1000) \times 0.5 \ll 1$. Une première approximation consiste donc à oublier purement et simplement le terme V_2 , au moins dans un premier temps.

V_1 peut d'ailleurs lui-même être simplifié. En effet, d'après (1.268), on a :

$$V_1 = -\frac{e}{\mu} A_0 \vec{p} \cdot \vec{a} \sin(\vec{k} \cdot \vec{r} - \omega t) . \quad (1.273)$$

Quand on calcule des éléments de matrice de V_1 avec des fonctions d'onde atomiques, l'intégrand ne prend de valeurs significatives que pour $\|\vec{r}\| \sim a_0$; dans ces conditions, le produit $k y$ venant du produit scalaire $\vec{k} \cdot \vec{r}$ est toujours très petit devant 1 sur le domaine pertinent d'intégration (là où la fonction d'onde est sensiblement non-nulle) et on peut remplacer V_1 par :

$$V_1 \simeq +\frac{e}{\mu} A_0 \vec{p} \cdot \vec{a} \sin \omega t \equiv V_{E1} . \quad (1.274)$$

Ceci constitue l'approximation dite dipolaire électrique⁸⁸ et revient notamment à négliger, dans le domaine optique, le déphasage spatial de l'onde sur l'espace atomique effectif.

Remarques

1. Il est instructif d'écrire les équations de Heisenberg en présence de V_{E1} ; on trouve :

$$\frac{d}{dt} \vec{r}_H = \frac{\vec{p}_H}{\mu} + \frac{e\mathcal{E}}{\mu\omega} \vec{a} \sin \omega t , \quad \frac{d}{dt} \vec{p}_H = -(\vec{\nabla} V)_H . \quad (1.275)$$

Ainsi, au mouvement intra-atomique se superpose une oscillation forcée le long du champ électrique appliqué.

2. La forme de l'interaction dipolaire électrique définie en (1.274) peut surprendre : on aurait plutôt attendu une expression du genre $V'_{E1} = -\vec{d} \cdot \vec{\mathcal{E}}$ où \vec{d} est le moment dipolaire électrique. En fait, (1.274) résulte directement du choix de la jauge (1.268) (le potentiel scalaire étant par ailleurs pris égal à zéro). Un changement de jauge approprié permet d'arriver, au même niveau d'approximation, à l'expression V'_{E1} .

Afin de calculer les amplitudes de transition dans l'approximation E1, il faut connaître les éléments de matrice de $\vec{a} \cdot \vec{p} = p_z$ entre les états propres de H_0 :

$$\langle \psi_f | V_{E1} | \psi_i \rangle = +\frac{e\mathcal{E}}{\mu\omega} \langle \psi_f | \vec{a} \cdot \vec{p} | \psi_i \rangle \sin \omega t \equiv +\frac{e\mathcal{E}}{\mu\omega} \langle \psi_f | p_z | \psi_i \rangle \sin \omega t . \quad (1.276)$$

Comme précédemment, $|\psi_i\rangle$ et $|\psi_f\rangle$ désignent les états extrêmes de la transition et sont supposés être propres de H_0 . En utilisant l'équation :

$$i\hbar \frac{d}{dt} \vec{p} = [\vec{p}, H_0] , \quad (1.277)$$

on obtient :

$$\langle \psi_f | V_{E1} | \psi_i \rangle = ie\mathcal{E} \frac{\omega_{fi}}{\omega} \sin \omega t \langle \psi_f | \vec{r} \cdot \vec{a} | \psi_i \rangle . \quad (1.278)$$

Finalement, les éléments de matrice de V_{E1} sont bien proportionnels à ceux du moment dipolaire électrique.

Pour que la transition $|\psi_i\rangle \rightarrow |\psi_f\rangle$ puisse se produire dans l'approximation E1, il est donc nécessaire et suffisant⁸⁹ que $\langle \psi_f | \vec{r} \cdot \vec{a} | \psi_i \rangle$ soit différent de zéro. Pour un champ central, les états $|\psi_i\rangle$ et $|\psi_f\rangle$ sont de la forme

⁸⁸approximation symbolisée par E1.

⁸⁹tant que l'on reste au premier ordre de la théorie des perturbations.

$R_{nl}Y_{lm}$; les harmoniques sphériques étant de parité déterminée et \vec{r} étant impair, une condition nécessaire est que $|\psi_i\rangle$ et $|\psi_f\rangle$ soient de parité *opposée*. Très précisément, on peut écrire les composantes de \vec{r} à l'aide de Y_{1m} :

$$z = r \sqrt{\frac{4\pi}{3}} Y_{10} , \quad x \pm iy = \mp r \sqrt{\frac{8\pi}{3}} Y_{1\pm 1} . \quad (1.279)$$

Il en résulte que l'intégrale angulaire de $\langle \psi_f | V_{E1} | \psi_i \rangle$ contient un facteur de la forme :

$$\int d\Omega Y_{l_f m_f}^* Y_{1 m} Y_{l_i m_i} \quad (m = 0, \pm 1) . \quad (1.280)$$

Une telle intégrale est nulle si $|l_i - l_f|$ n'est pas égal à 1 et si $-m_f + 1 + m_i$ n'est pas nul. Ceci permet d'énoncer les règles de sélection dipolaires électriques donnant les symétries des états extrêmes pouvant donner lieu à ces transitions :

$$\Delta l = \pm 1 , \quad \Delta m = 0, \pm 1 . \quad (1.281)$$

Pour une polarisation linéaire du champ électrique, il faut $\Delta m = 0$; $\Delta m = \pm 1$ vaut pour les champs polarisés circulairement dans un plan perpendiculaire à l'axe de quantification. Les relations (1.281) étant satisfaites, les intensités relatives des raies permises dans l'approximation E1 sont (essentiellement) fixées relativement par la valeur de l'intégrale radiale $\int r^2 dr R_{n_f l_f} r R_{n_i l_i}$.

Quand une transition est interdite dans l'approximation E1 parce que les règles (1.281) ne sont pas satisfaites, ceci ne signifie pas qu'elle est inobservable, mais seulement qu'elle est *a priori* de faible intensité. En pareil cas, il convient d'aller un cran plus loin dans les approximations ; par exemple – et si on décide de rester au 1^{er} ordre de la théorie des perturbations –, on peut introduire la première correction au déphasage spatial de l'onde en développant le sinus ; on trouve ainsi^{90, 91} :

$$V_1 \simeq V_{E1} - \frac{e\mathcal{B}}{\mu} (\vec{p} \cdot \vec{a}) (\vec{n} \cdot \vec{r}) \cos \omega t . \quad (1.282)$$

L'apparition du module du champ magnétique montre que la cohérence du calcul exige d'introduire simultanément le couplage direct entre le spin et ce champ, donnant lieu au terme V_2 ; en effet, d'après (1.282), l'écart à V_{E1} est d'ordre $\frac{e\mathcal{B}}{\mu} pr \sim \frac{e\mathcal{B}}{\mu} \hbar$, alors que $V_2 \sim \frac{e}{\mu} \hbar \mathcal{B}$. Toujours pour la cohérence, il faut développer V_2 au premier ordre en $\vec{k} \cdot \vec{r}$. Quelques manipulations simples sur le terme correctif dans (1.282) font notamment apparaître le moment cinétique orbital \vec{L} . Finalement, après regroupement de tous les termes, l'opérateur $V(t)$ complet (1.271) apparaît sous la forme approchée :

$$V(t) \simeq V_{E1} - \frac{e}{2\mu} (\vec{L} + g_e \vec{S}) \cdot \vec{\mathcal{B}} \cos \omega t - \frac{e}{2\mu\omega} [(\vec{k} \cdot \vec{r})(\vec{\mathcal{E}} \cdot \vec{p}) + (\vec{k} \cdot \vec{p})(\vec{\mathcal{E}} \cdot \vec{r})] \cos \omega t \equiv V_{E1} + V_{M1} + V_{E2} , \quad (1.283)$$

où $\vec{\mathcal{E}}$ et $\vec{\mathcal{B}}$ représentent ici simplement $\vec{\mathcal{E}}(\vec{0}, t)$ et $\vec{\mathcal{B}}(\vec{0}, t)$. Le terme contenant les moments cinétiques est appelé dipolaire magnétique⁹², puisqu'il représente le couplage direct entre le champ magnétique et le moment (dipolaire) magnétique total de l'atome. Le deuxième terme, V_{E2} , implique des éléments de matrice de combinaisons quadratiques des coordonnées⁹³ et est appelé pour cette raison couplage quadrupolaire électrique, symboliquement E2.

Il est maintenant possible d'énoncer de nouvelles règles de sélection. En ce qui concerne les transitions M1, pour qu'elles soient possibles, il faut d'abord $\Delta l = 0$ puisque $(\vec{L} + g_e \vec{S})$ ne saurait changer le nombre quantique l . Il faut en outre $\Delta m_l = 0, \pm 1$ ou $\Delta m_s = 0, \pm 1$.

Pour les transitions E2, il faut bien sûr $\Delta m_s = 0$. En considérant l'intégrale angulaire $\langle Y_{l_f m_f} | Y_{2m} | Y_{l_i m_i} \rangle$, des arguments analogues à ceux utilisés plus haut pour les transitions E1 montrent que les transitions E2 ne sont pas interdites si :

$$\Delta l = 0, \pm 2 , \quad \Delta m = 0, \pm 1, \pm 2 . \quad (1.284)$$

⁹⁰ Au passage, on utilise la relation $kA_0 = \mathcal{B}$.

⁹¹ \vec{n} désigne un vecteur unitaire le long de \vec{k} .

⁹² Pour dipolaire magnétique, l'acronyme consacré est M1.

⁹³ On s'en convainc en utilisant (1.277).

Ces considérations se généralisent à un ordre multipolaire arbitraire. D'un autre côté, il faut se souvenir que l'analyse théorique des raies atomiques contient deux approximations emboîtées : d'une part le développement multipolaire amorcé ci-dessus, d'autre part les ordres successifs de la théorie des perturbations. Dans toute situation précise, il convient de comparer entre elles les différentes contributions provenant de ces deux types de développements afin de conduire un calcul cohérent, et en relation avec la précision expérimentale.

1.8 États de diffusion

Le potentiel Coulombien conduit également à des états propres d'énergie positive E , représentant la diffusion d'un électron non-lié par un centre chargé positivement en l'absence de tout écrantage⁹⁴. Il s'agit de ce que l'on appelle académiquement la diffusion Rutherford, étant entendu que, en pratique, au moins en phase dense, il n'est pas correct de conserver le potentiel Coulombien nu pour décrire la diffusion aux petits angles, laquelle correspond aux grandes valeurs du paramètre d'impact⁹⁵. La longue portée du potentiel Coulombien exige d'ailleurs un traitement spécifique, alors que la diffusion par un potentiel central décroissant plus vite que $\frac{1}{r}$ peut être formalisée sans référence explicite aux détails définissant ce dernier. Dans ce dernier cas, il est possible de deviner la forme des solutions à très grande distance sur la base de quelques arguments physiques simples.

Dans la situation la plus simple, on considère la diffusion d'un projectile par un centre diffuseur infiniment massif, que l'on peut dès lors remplacer par son champ de forces, donnant lieu à l'énergie potentielle $V(\vec{r})$ pour le projectile. Ceci étant, il faut résoudre l'équation aux valeurs propres pour un certain Hamiltonien H :

$$H = \frac{\vec{p}^2}{2m} + V(\vec{r}) . \quad (1.285)$$

Le premier terme est l'énergie cinétique d'un projectile unique, de masse m , $V(\vec{r})$ est l'énergie potentielle de ce projectile dans le champ d'un diffuseur de la cible. L'un des enjeux de la théorie des collisions est l'obtention de la section efficace différentielle de diffusion, $\sigma_d(\theta, \phi)$. Cette quantité est définie comme suit ; soit dn le nombre de projectiles diffusés par unité de temps dans l'angle solide $d\Omega$ déterminé par les angles θ (déviation) et ϕ (azimut) et soit j_i le courant incident (nombre de projectiles traversant l'unité de section droite du faisceau par unité de temps). Par définition, on a :

$$dn = j_i \sigma_d(\theta, \phi) d\Omega . \quad (1.286)$$

σ_d est homogène à une surface. La section efficace totale de collision s'obtient par intégration sur les angles :

$$\sigma = \int d\Omega \sigma_d(\theta, \phi) = \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta \sigma_d(\theta, \phi) . \quad (1.287)$$

1.8.1 Nature des états de diffusion

Les états stationnaires de diffusion se construisent à partir des états propres ψ du Hamiltonien (1.285) :

$$\left[\frac{\vec{p}^2}{2m} + V(\vec{r}) \right] \psi(\vec{r}) = E\psi(\vec{r}) . \quad (1.288)$$

En posant :

$$E = \frac{\hbar^2 k^2}{2m} \quad (k \in \mathbb{R}) , \quad V = \frac{\hbar^2}{2m} U , \quad (1.289)$$

on obtient :

$$[\Delta + k^2 - U(\vec{r})] \psi(\vec{r}) = 0 . \quad (1.290)$$

⁹⁴En pratique, il faut tenir compte du recul de la cible et donc relier précisément la section efficace dans le repère du laboratoire et la section efficace dans le repère du centre de masse ; à ce sujet, voir [5] p.113.

⁹⁵On sait bien d'ailleurs que la section efficace différentielle de Rutherford se comporte comme $1/\sin^4 \frac{\theta}{2}$ près de $\theta = 0$ et n'est donc pas intégrable. Cette difficulté est précisément réglée par l'introduction d'un écrantage, représentant le "blindage" de la charge nucléaire par les électrons situés près du noyau.

On sait résoudre exactement cette équation pour quelques potentiels remarquables, les fonctions propres différant bien évidemment d'un potentiel à l'autre. Ces différences sont de peu d'intérêt, puisque le compteur se trouve toujours, en pratique, à une distance gigantesque comparée à l'échelle caractéristique de variation de l'interaction due à la cible et décrite par $V(\vec{r})$. En d'autres termes, ce qui importe ce sont les propriétés *asymptotiques* ("à l'infini") des solutions propres $\psi(\vec{r})$, dont on peut deviner la forme générale sur la base de considérations physiques simples.

Remarquons d'abord que le problème physique contient maintenant une direction privilégiée, définie par la localisation de la source de particules – supposée ponctuelle – et celle de la cible. Cette direction définit l'axe Oz dans la suite ; afin de fixer les idées, la source est placée en $z = -\infty$, la cible étant à l'origine (voir fig. 1.3).

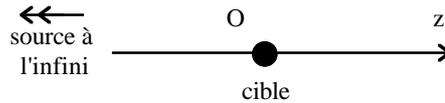


Figure 1.3: Géométrie utilisée dans le texte. La source est située à l'infini en $z = -\infty$.

Bien évidemment, s'agissant de trouver les états propres, il n'y a nulle part de dépendance en temps (celle-ci apparaît quand on introduit des paquets d'ondes construits avec les états stationnaires déduits des ψ et représentant une particule émise par la source à un instant très ancien). Ceci étant précisé, on peut affirmer que, pour $z \rightarrow -\infty$, le terme dominant est e^{ikz} ($k > 0$) puisque, la source étant infiniment éloignée, les particules sont libres ($V = C^{ste}$ à l'infini). Un terme analogue est aussi présent en $z = +\infty$, représentant les particules transmises sans modification (diffusion vers l'avant). Par ailleurs, à grande distance, il doit aussi exister une onde sphérique du genre $\frac{1}{r} e^{ikr}$, représentant l'onde diffusée : c'est elle qui porte la signature du potentiel, et elle satisfait l'équation :

$$(\Delta + k^2) \frac{1}{r} e^{ikr} = 0 \quad (r \rightarrow +\infty) , \quad (1.291)$$

qui est l'équation aux valeurs propres loin de la zone d'action du potentiel V .

Ces considérations incitent à poser que la forme asymptotique de $\psi(\vec{r})$ est une combinaison linéaire des deux termes ci-dessus. Dans la géométrie précisée, on pose ainsi (délaissant la question de la normalisation) :

$$\psi_{\vec{k}}(\vec{r}) \sim C \left(e^{ikz} + f_k(\theta, \phi) \frac{e^{ikr}}{r} \right) , \quad (1.292)$$

où le coefficient f_k (homogène à une longueur) dépend des angles en général et constitue l'image de la cible. f_k s'appelle amplitude de diffusion ; l'indice k rappelle que c'est une fonction du vecteur d'onde, donc de l'énergie. C est une constante de normalisation, qui sera calée sur le courant émis par la source. Il est essentiel de se souvenir que l'expression de $\psi_{\vec{k}}$ prend en compte la position de la source dans l'espace.

En pratique, l'onde plane e^{ikz} n'est pas infiniment étendue dans la direction transverse, le faisceau incident ayant une ouverture finie, que l'on retrouve peu ou prou dans l'onde transmise. En pratique, pour ne pas être "aveuglé", on place le détecteur en-dehors de la direction d'incidence de sorte que, dans la région où se trouve ce dernier, il n'y a pas d'interférences possibles entre onde transmise et onde diffusée. Les interférences ne se produisent que dans la région avant et elles sont d'ailleurs forcément destructives puisque, au total, le nombre de particules est conservé et qu'une partie de ces dernières est diffusées à θ non-nul.

Il existe une relation très simple entre l'amplitude f_k et la section efficace différentielle σ_d . En effet, si on se place en $z = -\infty$, seule subsiste l'onde incidente ($\frac{1}{r} = 0$) et le courant associé⁹⁶ à e^{ikz} est $\frac{\hbar k}{m}$. En réalité, en cet endroit, l'onde plane est $C e^{ikz}$ où C est une constante. Le courant incident, parallèle à Oz , a donc pour module :

$$j_i = |C|^2 \frac{\hbar k}{m} . \quad (1.294)$$

⁹⁶Le courant est généralement donné par :

$$\vec{j} = \frac{\hbar}{2im} (\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^*) = \frac{\hbar}{m} \Im (\psi^* \vec{\nabla} \psi) . \quad (1.293)$$

La constante C s'obtient en connaissant le taux d'émission de la source.

Le courant diffusé, hors la région avant, se calcule uniquement avec la partie onde diffusée (puisque dans cette région, il n'y a pas d'onde transmise). Les composantes du gradient sont :

$$\vec{\nabla} = \left(\frac{\partial}{\partial r}, \frac{1}{r} \frac{\partial}{\partial \theta}, \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \right) . \quad (1.295)$$

Pour l'onde sphérique $C f_k(\theta, \phi) r^{-1} e^{i k r}$, on trouve le courant diffusé suivant :

$$\vec{j}_d = |C|^2 \left(\frac{\hbar k}{m r^2} |f_k|^2, \frac{\hbar}{m r^3} \Im \left(f_k^* \frac{\partial f_k}{\partial \theta} \right), \frac{\hbar}{m r^3 \sin \theta} \Im \left(f_k^* \frac{\partial f_k}{\partial \phi} \right) \right) . \quad (1.296)$$

Dans la zone d'observation, seule subsiste en pratique la composante radiale, les deux autres étant beaucoup plus petites du fait de leur dépendance en r^{-3} . On a donc :

$$\vec{j}_d \simeq \vec{j}_{\text{radial}} = |C|^2 \frac{\hbar k}{m r^2} |f_k|^2 \frac{\vec{r}}{r} . \quad (1.297)$$

Le nombre de particules par unité de temps, dn , qui frappent le détecteur orienté suivant la direction (θ, ϕ) , est $j_d dS$, où $dS = r^2 d\Omega$ est l'ouverture du détecteur⁹⁷. Il en résulte, compte tenu de (1.294) :

$$dn = j_d r^2 d\Omega = |C|^2 \frac{\hbar k}{m r^2} |f_k|^2 r^2 d\Omega \equiv j_i |f_k|^2 d\Omega , \quad (1.298)$$

d'où, en se référant à (1.286), la relation simple entre amplitude de diffusion et section efficace différentielle :

$$\sigma_d(\theta, \phi) = |f_k(\theta, \phi)|^2 . \quad (1.299)$$

Ainsi, le module carré de l'amplitude de l'onde sphérique diffusée n'est rien d'autre que la section efficace cherchée. L'enjeu est donc de trouver f_k .

1.8.2 Méthode des déphasages

Cette méthode est construite par comparaison systématique entre la diffusion par un potentiel $V(r)$, supposé à symétrie sphérique⁹⁸, et la résolution du problème de la particule libre dans \mathbb{R}^3 en coordonnées sphériques ; ce dernier est un cas particulier de champ central. Sa résolution permet d'établir une relation importante :

$$e^{i k z} = \sum_{l=0}^{+\infty} (2l+1) i^l j_l(kr) P_l(\cos \theta) , \quad (1.300)$$

et comme :

$$P_l(\cos \theta) = \sqrt{\frac{4\pi}{2l+1}} Y_{l0}(\theta) , \quad (1.301)$$

on a aussi :

$$e^{i k z} = \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} j_l(kr) Y_{l0}(\theta) , \quad (1.302)$$

où j_l une fonction de Bessel sphérique⁹⁹. Cette relation n'exprime rien d'autre que la décomposition de $e^{i k z}$ – qui, pour $k > 0$, représente une onde plane se propageant le long de Oz –, vecteur propre commun à \vec{p} et $H_0 = \frac{\vec{p}^2}{2m}$, sur la base de vecteurs propres de (H_0, \vec{L}^2, L_z) . Ces derniers sont de la forme :

$$\psi_{k l m}^{(0)}(\vec{r}) = C_k j_l(kr) Y_{l m}(\theta, \phi) , \quad (1.304)$$

⁹⁷puisque $d\Omega$ est par définition la surface de sphère de rayon unité délimitée par le petit cône d'axe (θ, ϕ) .

⁹⁸Dès lors, l'amplitude f_k , et donc aussi la section efficace, ne dépend pas de l'angle azimutal ϕ .

⁹⁹Les j_l sont plus précisément les fonctions de Bessel sphériques de première espèce, et sont reliées aux fonctions de Bessel ordinaires J_ν selon :

$$j_l(z) = \sqrt{\frac{\pi}{2z}} J_{n+\frac{1}{2}}(z) \quad (n = 0, \pm 1, \pm 2, \dots) . \quad (1.303)$$

où C_k est une constante de normalisation¹⁰⁰ et Y_{lm} une harmonique sphérique. k est le label quantique représentant l'énergie :

$$E = \frac{\hbar^2 k^2}{2m} . \quad (1.306)$$

Les $j_l(kr)$ introduisent la dépendance en E par l'intermédiaire du vecteur d'onde k (voir (1.289)).

On sait donc tout des états $\psi_{klm}^{(0)}$, appelés ondes sphériques libres, et aussi ondes partielles dans le contexte de la théorie des collisions. En particulier, on connaît le développement asymptotique de j_l :

$$j_l(kr) \simeq \frac{1}{kr} \sin\left(kr - l\frac{\pi}{2}\right) \quad (r \gg k^{-1}) , \quad (1.307)$$

ce qui permet, toujours dans le cas libre, d'écrire :

$$e^{ikz} \simeq \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} \frac{1}{kr} \sin\left(kr - l\frac{\pi}{2}\right) Y_{l0}(\theta) \quad (kz \gg 1) , \quad (1.308)$$

pour une onde plane issue de $z = -\infty$ ($k > 0$). En définitive, l'expression (1.308) est la forme asymptotique (exacte) de la solution libre satisfaisant les conditions aux limites relatives à une source située à l'infini du côté $z < 0$.

Introduisons maintenant le potentiel diffuseur. L'idée de la méthode des déphasages consiste à envisager les changements minimaux dans le développement asymptotique exact du cas libre : le plus simple est d'ajouter un déphasage. Autrement dit, on pose *a priori* que l'effet du potentiel se réduit à donner un déphasage spécifique à chaque onde partielle de la particule libre. D'où l'écriture :

$$\psi_k(\vec{r}) \simeq C' \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} \frac{1}{kr} \sin\left(kr - l\frac{\pi}{2} + \delta_l\right) Y_{l0}(\theta) . \quad (1.309)$$

Bien évidemment, il ne s'agit pour l'instant que d'un *ansatz* ; s'il est correct, on doit retrouver¹⁰¹ le comportement asymptotique deviné pour ψ_k , (1.292). Les déphasages δ_l sont la signature du potentiel $V(r)$ quand l'onde diffusée est analysée à l'infini ; ils dépendent de l'énergie *via* le vecteur d'onde k .

Il reste maintenant à vérifier que l'idée ayant conduit au développement (1.309) est correcte – ce que confirme un calcul explicite ([8], VIII-4-a- β). On a :

$$\begin{aligned} \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} \frac{1}{kr} \sin\left(kr - l\frac{\pi}{2} + \delta_l\right) Y_{l0}(\theta) = \\ e^{-i\delta_l} \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} \frac{e^{ikr} e^{-il\pi/2} e^{2i\delta_l} - e^{-ikr} e^{+il\pi/2}}{2ikr} Y_{l0}(\theta) . \end{aligned} \quad (1.310)$$

En utilisant :

$$e^{2i\delta_l} = 1 + 2i e^{i\delta_l} \sin \delta_l , \quad (1.311)$$

le numérateur dans (1.310) est :

$$e^{ikr} e^{-il\pi/2} (1 + 2i e^{i\delta_l} \sin \delta_l) - e^{-ikr} e^{+il\pi/2} , \quad (1.312)$$

d'où :

$$\psi_k(\vec{r}) \simeq C' e^{-i\delta_l} \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} \left[\frac{e^{-il\pi/2} e^{i\delta_l}}{kr} \sin \delta_l e^{ikr} + \frac{e^{i(kr-l\pi/2)} - e^{-i(kr-l\pi/2)}}{2ikr} \right] Y_{l0}(\theta) . \quad (1.313)$$

¹⁰⁰ Avec $C_k = \sqrt{2k^2/\pi}$, on a :

$$\langle \psi_{klm}^{(0)} | \psi_{k'l'm'}^{(0)} \rangle = \delta(k-k') \delta_{ll'} \delta_{mm'} . \quad (1.305)$$

¹⁰¹ Dans tous les cas, on a bien une onde plane issue de $z = -\infty$.

Le deuxième terme du crochet donne, d'après (1.308), l'expression asymptotique de l'onde plane e^{ikz} ; d'où :

$$\psi_k(\vec{r}) \simeq C' e^{-i\delta_l} \left[\sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} e^{-il\pi/2} \frac{e^{ikr}}{kr} e^{i\delta_l} \sin \delta_l + e^{ikz} \right] Y_{l0}(\theta) . \quad (1.314)$$

Ceci est bien du type (1.292) (avec $C = C' e^{-i\delta_l}$), ce qui valide l'idée première et, surtout, fournit l'expression de l'amplitude f_k en fonction des déphasages – donc aussi, par (1.299), la section efficace en fonction des δ_l :

$$f_k(\theta) = \frac{1}{k} \sum_{l=0}^{+\infty} \sqrt{4\pi(2l+1)} e^{i\delta_l} \sin \delta_l Y_{l0}(\theta) . \quad (1.315)$$

Par intégration sur les angles, on en déduit la section efficace totale σ ; compte tenu de l'orthonormalisation des Y_{lm} , il vient :

$$\sigma = \frac{4\pi}{k^2} \sum_{l=0}^{+\infty} (2l+1) \sin^2 \delta_l(k) . \quad (1.316)$$

En pratique, un tel développement n'est vraiment utile que pour les potentiels donnant un petit nombre de déphasages sensiblement différents de zéro¹⁰². Les δ_l peuvent en principe se calculer, si le potentiel $V(r)$ est connu, en résolvant l'équation aux valeurs propres correspondante pour chaque valeur de l . Il se trouve que le nombre de déphasages pertinents peut être relié à la portée spatiale du potentiel ; en effet, revenant au cas libre ($V \equiv 0$), et donc aux états propres contenant $j_l(kr)$, il suffit de considérer le comportement de ces fonctions près de l'origine. On sait que :

$$j_l(kr) \simeq \frac{(kr)^l}{(2l+1)!!} \quad (kr \ll 1) . \quad (1.317)$$

Donc, plus l est grand, plus la fonction j_l est "écrasée" dans le voisinage de l'origine ; un calcul précis montre que j_l décolle pour $kr \sim l$, ce qui permet de définir une longueur typique $r_0 = l/k$. Il en résulte qu'un potentiel de portée ξ petite par rapport à r_0 , donc tel que $k\xi \ll 1$, ne donnera qu'un déphasage sensiblement non-nul, δ_0 . Plus généralement, les δ_l significatifs auront un indice l satisfaisant :

$$l \leq k\xi . \quad (1.318)$$

Sans surprise, le nombre de déphasages pertinents augmente avec l'énergie.

Terminons par un exemple simple, mais qui contient un résultat un peu inattendu quand on le confronte avec le cas classique : la diffusion par un puits infini sphérique. Dans ce cas :

$$V(r) = \begin{cases} 0 & \text{si } r > \xi \\ +\infty & \text{si } r < \xi \end{cases} . \quad (1.319)$$

D'après ce qui précède, seuls les déphasages satisfaisant (1.318) sont pertinents. En choisissant une énergie suffisamment faible (donc k petit), on peut se placer dans les conditions où seule l'onde s ($l = 0$) compte. Il reste alors, suivant (1.315) :

$$f_k(\theta) = \frac{\sqrt{4\pi}}{k} e^{i\delta_0} \sin \delta_0 Y_{00} = \frac{1}{k} e^{i\delta_0(k)} \sin \delta_0(k) . \quad (1.320)$$

f_k est indépendant de l'angle de diffusion (une conséquence du fait que seule l'onde sphérique importe vraiment) et la section efficace totale est donc :

$$\sigma = \frac{4\pi}{k^2} \sin^2 \delta_0(k) . \quad (1.321)$$

Il reste à trouver $\delta_0(k)$, ce que l'on peut faire en résolvant l'équation radiale avec $l = 0$. En introduisant comme d'habitude la fonction $u(r) = rR(r)$, il faut résoudre :

$$u'' + k^2 u = 0 \quad (\forall r > \xi) , \quad (1.322)$$

¹⁰²sauf bien sûr si, par bonne fortune, on sait resommer la série.

puisque $u(r)$ satisfait une équation aux valeurs propres à une dimension (réduite à \mathbb{R}_+) avec le potentiel effectif $V_{\text{eff}} = V(r) + \frac{\hbar^2 l(l+1)}{2mr^2} = 0$ si $l = 0$ et $r > \xi$). En outre, par la continuité de la fonction d'onde – identiquement nulle pour $r < 0^-$, on doit avoir :

$$u(r = \xi) = 0 . \quad (1.323)$$

La solution est donc :

$$u(r) = \begin{cases} A \sin[k(r - \xi)] & \text{si } r > \xi \\ 0 & \text{si } r < \xi \end{cases} . \quad (1.324)$$

Par ailleurs, le déphasage est défini par le comportement asymptotique de la fonction radiale $R(r)$ (voir (1.309)) :

$$R(r) \simeq \sqrt{4\pi} \frac{C'}{kr} \sin(kr + \delta_0) \iff u(r) \simeq C'' \sin(kr + \delta_0) \quad (r \rightarrow +\infty) . \quad (1.325)$$

Par comparaison avec (1.324), il vient $\delta_0 = -k\xi$ et donc, suivant (1.316) :

$$\sigma = \frac{4\pi}{k^2} \sin^2 k\xi . \quad (1.326)$$

Comme on a supposé $k\xi \ll 1$, le résultat significatif est :

$$\sigma \simeq 4\pi \xi^2 . \quad (1.327)$$

Ce résultat, valide notamment à très basse énergie, montre que dans une telle limite la section efficace est indépendante de l'énergie. L'expression (1.327) n'est rien d'autre que la section efficace totale classique ($\hbar = 0$) d'une sphère dure de rayon ξ . . . à un facteur 4 près ! ($\sigma_{\hbar=0} = \pi \xi^2$). On peut voir dans cette différence un effet quantique irréductible lié à la variation brutale du potentiel sur une échelle de longueur toujours plus petite que toute autre longueur¹⁰³, puisque V a un saut (variation sur une longueur nulle). Par ailleurs, on sait du cas unidimensionnel qu'il existe, à basse énergie, des résonances de diffusion, liées au fait que la particule, se déplaçant lentement, a bien le temps de sonder le potentiel. Ces résonances existent aussi dans \mathbb{R}^3 et sont visibles dès l'onde p.

À l'inverse, à haute énergie, les longueurs d'onde deviennent très petites et un grand nombre de déphasages sont importants ; dans ce cas, on trouve ([8], C_{VIII}, eq. (22)) :

$$\sigma \simeq 2\pi \xi^2 . \quad (1.328)$$

À nouveau, les effets quantiques subsistent, bien que l'on soit dans la limite des faibles longueurs d'onde où l'on s'attend à retrouver la mécanique "géométrique". Ceci est à nouveau lié à l'existence d'une discontinuité du potentiel. Dans tous les cas, on a :

$$\sigma > \sigma_{\hbar=0} , \quad (1.329)$$

un résultat que l'on peut interpréter comme une conséquence de la dualité onde-corpuscule : une particule est inévitablement délocalisée et sonde l'espace sur une région plus vaste au sens classique que celle occupée par les obstacles qu'elle rencontre – ces derniers étant traités classiquement¹⁰⁴.

1.8.3 Le cas du potentiel Coulombien

S'agissant du potentiel Coulombien, la résolution de l'équation aux valeurs propres (1.57) pour $E > 0$ démarre exactement comme pour les états liés. En conservant les mêmes définitions que précédemment, notamment en gardant :

$$k = \hbar^{-1} \sqrt{-2\mu E} \equiv i \hbar^{-1} \sqrt{2\mu E} \equiv i \kappa , \quad (1.330)$$

une différence essentielle apparaît : maintenant, k est imaginaire pur, tout comme le paramètre λ défini en (1.65). Ceci étant, la récurrence (1.67) reste vraie et fournit des coefficients complexes, mais il n'y a plus aucun moyen de l'arrêter (puisque λ est imaginaire pur) – ce qui n'est plus nécessaire d'ailleurs : en effet, les solutions ne divergent plus à l'infini, puisque qu'elles contiennent le facteur oscillant $e^{\pm i \kappa r}$. La conséquence immédiate

¹⁰³comme par exemple la longueur d'onde de de Broglie $\hbar/\sqrt{2mE} = k^{-1}$.

¹⁰⁴faute de quoi, leur position ne saurait être parfaitement définie.

est que l'énergie E n'est plus quantifiée et peut prendre toute valeur de 0 à $+\infty$. Au total, le spectre de l'atome d'hydrogène est bien composé de deux parties : une partie discrète obtenue plus haut, associée aux états liés, et une partie continue couvrant \mathbb{R}_+ associée aux états propres non-liés.

Les fonctions propres s'expriment maintenant à l'aide de séries que l'on identifie avec une fonction spéciale connue (la fonction hypergéométrique), ce qui permet d'obtenir notamment leur comportement asymptotique et de mettre en évidence la particularité du potentiel Coulombien (sa longue portée). En effet, l'analyse asymptotique révèle l'existence de termes du genre :

$$\frac{1}{r} e^{\pm i [\kappa r - (\ln \kappa r) / (\kappa a_0)]} . \quad (1.331)$$

La comparaison avec (1.292) permet de constater que l'onde diffusée contient, même à grande distance, un déphasage dépendant (faiblement) de la distance. Ainsi, en raison de la portée infinie du champ Coulombien, l'onde diffusée ne coïncide pas avec une onde sphérique à déphasage constant.

Remarque

Sans rentrer dans la jungle des fonctions spéciales, il est facile de mettre en évidence le déphasage logarithmique apparaissant dans (1.331). L'équation aux valeurs propres est :

$$[\Delta + \kappa^2 - \mathcal{U}(r)] u(r) = 0 , \quad (1.332)$$

où :

$$\mathcal{U}(r) = -\frac{l(l+1)}{r^2} + \frac{2\mu e'^2}{\hbar^2 r} \equiv -\frac{l(l+1)}{r^2} + \frac{2}{a_0 r} . \quad (1.333)$$

On pose :

$$u(r) = F(r) e^{i\kappa r} , \quad (1.334)$$

et on suppose que $F(r)$ a un comportement en loi-puissance à l'infini :

$$F(r) \simeq \frac{1}{r^\lambda} \quad (r \rightarrow +\infty) , \quad (1.335)$$

où λ est à déterminer. Reportant dans (1.332) et ne retenant que les termes dominants, on trouve :

$$\lambda = \frac{i}{\kappa a_0} , \quad (1.336)$$

d'où $F(r) \sim r^{-i/(\kappa a_0)} = e^{-[i/(\kappa a_0)] \ln(\kappa r)}$. Le déphasage de l'onde dépend de r et c'est pourquoi les déphasage δ_l de la théorie des collisions usuelle perdent ici une grande partie de leur intérêt. Pour terminer, signalons que l'existence d'une direction privilégiée rend naturel l'usage des coordonnées paraboliques, ce qui permet d'obtenir assez simplement l'amplitude de diffusion Coulombienne (voir [7], ch. 11) :

$$A_C(\theta) = \frac{1}{2\kappa^2 a_0 \sin^2 \frac{\theta}{2}} , \quad (1.337)$$

dont le module au carré fournit la section efficace de Rutherford :

$$\sigma_{dC}(\theta) = \left(\frac{e'^2}{4E} \right)^2 \frac{1}{\sin^4 \frac{\theta}{2}} . \quad (1.338)$$

Comme on s'en doute, la divergence aux petits angles¹⁰⁵ (particule passant loin du noyau) peut être soignée en invoquant un phénomène d'écrantage toujours présent dans la réalité.

¹⁰⁵La divergence en $\sim \theta^{-4}$ a pour conséquence que la section efficace intégrée est infinie.

Chapitre 2

Particules identiques

Ce chapitre expose les problèmes particuliers que soulève en Mécanique Quantique la description des systèmes composés de particules identiques, débouchant sur la nécessité d'un nouveau postulat. Après l'énoncé de celui-ci, on montrera le rôle joué par les opérateurs de permutation, – notamment par le traitement de quelques exemples – et une initiation à la Seconde Quantification sera exposée.

2.1 Indiscernabilité des particules identiques en Mécanique Quantique

On a déjà rencontré un concept spécifiquement quantique, le spin, dont il est impossible de donner une image classique en termes de rotation d'une particule sur elle-même. Le spin s'évanouit à la limite classique : s'il est possible de concevoir des moments cinétiques orbitaux arbitrairement élevés, on ne connaît pas de moment de spin aussi grand que l'on veut.

Un autre concept¹ s'évanouissant à la limite classique est l'indiscernabilité des particules identiques en Mécanique Quantique, provenant de la disparition de la notion de trajectoire : dans un cadre quantique, on ne peut plus, après avoir numéroté les particules dans l'état de départ, les suivre à la trace pour les identifier dans l'état final. Il en résulte notamment qu'une question comme "quelle est la probabilité de trouver l'électron 1 en \vec{r}_1 et l'électron 2 en \vec{r}_2 ?" n'a pas de sens. À l'inverse, la question "quelle est la probabilité de trouver *un* électron en \vec{r}_1 et *un* en \vec{r}_2 " est sensée. L'impossibilité de distinguer plusieurs particules identiques cohabitantes au sein d'un même système – c'est-à-dire finalement l'invariance des propriétés physiques dans tout *échange* des variables dynamiques des particules – se traduira par une symétrie de *permutation* de la fonction d'onde. Cette invariance sera énoncée sous la forme d'un nouveau postulat, aux conséquences extraordinaires : il est l'un des fondements de l'explication de la stabilité de la matière constituée et de ses propriétés².

Deux particules sont dites identiques si toutes leurs grandeurs intrinsèques (masse, charge, spin, *etc.*) sont les mêmes. Un proton et un électron ont même spin mais sont distincts par leurs masses. Un électron et un positron ont même masse et même spin mais ont des charges opposées. Aucune expérience ne permet de distinguer deux particules identiques – sauf si l'on sait qu'elles sont par exemple dans deux régions distinctes

¹Comme on le verra, ces deux concepts spécifiquement quantiques sont intimement liés.

²Un exemple : c'est parce que la matière est formée de fermions – les bosons étant des particules associées aux interactions (champs) – que la limite thermodynamique existe ; en particulier, pour qu'un système de particules de charges opposées soit stable, il faut que l'une des espèces soit constituée de fermions. Le Principe de Pauli sauve la situation : sans lui, l'énergie de l'état fondamental varierait comme $N^{7/5}$ en fonction du nombre N de particules et la limite thermodynamique (qui exige $E \propto N$) tout simplement n'existerait pas ! ([13], § 2. 1).

de l'espace³. Aucune propriété d'un système contenant des particules identiques n'est modifiée si l'on procède à l'échange de deux telles particules.

En Mécanique Classique, la description d'un système composé de particules identiques ne pose aucun problème particulier. Il est possible de numéroter (ou de colorier) les particules dans l'état initial, puis de les suivre une à une sur leur trajectoire. A tout instant, il est possible d'identifier par son dossard la particule portant au départ un numéro donné ; c'est bien ce suivi de proche en proche qui fait de leur numéro (ou de leur couleur) une "constante du mouvement" individuelle.

En Mécanique Quantique, il en va tout autrement : comme il n'existe plus de trajectoire, on ne peut plus suivre les particules à la trace. Même s'il est encore parfois possible de les numéroter au départ, à la condition expresse qu'elles se trouvent dans des régions distinctes de l'espace, il n'est plus possible de retrouver leur numéro dans l'état final si leurs paquets d'ondes se sont enchevêtrés entre temps. L'impossibilité de distinguer deux particules identiques disparaît strictement dans la limite classique : alors, les paquets d'ondes restent parfaitement localisés, ne s'étalent pas et ne se mélangent pas ; leur centre suit la trajectoire classique, qui reprend alors tout son sens. En définitive, le non-recouvrement des paquets d'onde restitue *de facto* la discernabilité classique et l'impossibilité de distinguer deux particules identiques disparaît⁴.

Considérons la collision, observée dans le repère du centre de masse, de deux particules identiques venant à la rencontre l'une de l'autre. Lorsqu'elles sont à l'infini l'une de l'autre, il est possible de numéroter 1 celle qui vient de gauche et 2 celle qui vient de droite. Après la collision, l'observation de l'une d'entre elles porte sur son énergie, son impulsion, son angle de diffusion, *etc.* Mais l'identification s'arrête ici : on ne peut retrouver son numéro puisqu'il est impossible de retracer sa trajectoire et d'en déduire que la particule observée est celle venue d'un côté ou de l'autre. Tant que les deux paquets d'ondes sont distincts, on peut schématiser les déplacements de leurs centres ; dès qu'ils s'interpénètrent, il n'y a plus deux paquets d'ondes identifiables mais une zone diffuse dans l'espace où la probabilité de trouver les *deux* particules est non-nulle. Après séparation, le "flou" lié à l'intrication des paquets d'ondes lors de la collision introduit une forme d'irréversibilité, et rend impossible toute autre identification que les grandeurs physiques intrinsèques de chaque particule (masse, charge, spin, *etc.*), lesquelles sont identiques puisque précisément les particules le sont.

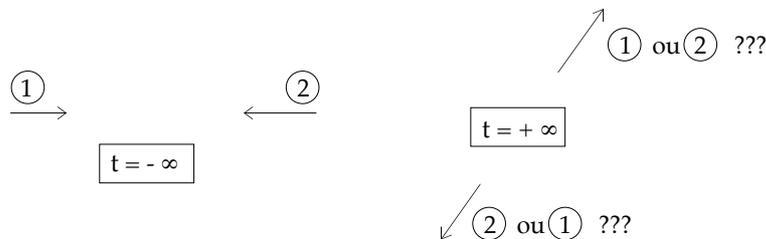


Figure 2.1: Illustration de la collisions de deux spins $\frac{1}{2}$.

Un exemple ([8], § XIV. 3) permet alors de saisir comment l'impossibilité de distinguer deux particules identiques conduit à une difficulté pour l'application des postulats déjà énoncés. Soit deux particules identiques de spin $S = \frac{1}{2}$, dont on ignore pour simplifier les degrés de liberté orbitaux. L'espace des états est de dimension 4, et est engendré par les vecteurs $|m_1\rangle \otimes |m_2\rangle \equiv |m_1, m_2\rangle$ avec $m_i = \pm\frac{1}{2}$. En l'absence de tout autre degré de liberté, le spin de chaque particule est la seule et unique observable pouvant faire l'objet d'une mesure individuelle. En ce sens, connaître les deux spins constitue l'information maximale permise par la Mécanique Quantique, telle qu'elle a été exposée jusqu'à maintenant.

Supposons que l'on sait que l'un des spins est dans l'état $+\frac{1}{2}$, l'autre dans l'état $-\frac{1}{2}$. Cette connaissance peut résulter par exemple d'une mesure individuelle de S_{1z} et de S_{2z} qui a donné $+\frac{\hbar}{2}$ pour un spin, $-\frac{\hbar}{2}$ pour

³voir à ce sujet la discussion de Messiah [11] "Faut-il toujours antisymétriser la fonction d'onde ?"

⁴La même conclusion subsiste encore approximativement si les paquets d'onde ne s'étalent pas trop et restent bien séparés. Par ailleurs, si le potentiel extérieur ne varie pas trop vite dans l'espace, leur centre suit en gros la trajectoire classique (Théorème d'Ehrenfest).

l'autre⁵. Juste après cette mesure, l'état est $|+\frac{1}{2}, -\frac{1}{2}\rangle$, ou $|-\frac{1}{2}, +\frac{1}{2}\rangle$, ou n'importe quelle combinaison linéaire des deux. Il est en effet certain que tout état du genre :

$$|\psi\rangle = a|+\frac{1}{2}, -\frac{1}{2}\rangle + b|-\frac{1}{2}, +\frac{1}{2}\rangle, \quad |a|^2 + |b|^2 = 1. \quad (2.1)$$

sur lequel on refait immédiatement une mesure de S_{1z} et de S_{2z} donne avec probabilité 1 ce qu'a donné la première mesure. Les postulats énoncés jusqu'à présent ne permettent nullement de fixer⁶ les coefficients a et b . En définitive, tout état du type (2.1) à la fois représente la connaissance maximale du système et est dégénéré (la dégénérescence correspondante est appelée dégénérescence d'échange)⁷. Clairement, cette ambiguïté sur l'état initial va conduire à l'impossibilité de faire des prévisions sur ce système.

À titre d'exemple, soit à trouver la probabilité $P(a, b)$ pour qu'une mesure simultanée de S_{1x} et S_{2x} , effectuée juste après la préparation ci-dessus, donne la valeur $+\frac{\hbar}{2}$ pour chaque spin. L'état propre $|+\frac{1}{2}\rangle_x$ de S_x se décompose comme suit sur les états propres de S_z , $|\pm\frac{1}{2}\rangle$:

$$|\frac{1}{2}\rangle_x = \frac{1}{\sqrt{2}} \left(|+\frac{1}{2}\rangle + |-\frac{1}{2}\rangle \right). \quad (2.2)$$

Pour les deux spins, le produit tensoriel se développe suivant :

$$|\frac{1}{2}\rangle_{1x} \otimes |\frac{1}{2}\rangle_{2x} = \frac{1}{2} \left(|+\frac{1}{2}, \frac{1}{2}\rangle + |+\frac{1}{2}, -\frac{1}{2}\rangle + |-\frac{1}{2}, +\frac{1}{2}\rangle + |-\frac{1}{2}, -\frac{1}{2}\rangle \right). \quad (2.3)$$

Suivant les postulats déjà énoncés, la probabilité $P(a, b)$ de trouver $+\frac{\hbar}{2}$ pour chaque spin le long de Ox en effectuant une mesure sur l'état $|\psi\rangle$ (2.1) est le module au carré du produit scalaire :

$$P(a, b) = |(\langle\psi|, |\frac{1}{2}\rangle_{1x} \otimes |\frac{1}{2}\rangle_{2x})|^2. \quad (2.4)$$

Le calcul à partir de (2.3) et (2.1) donne immédiatement $P(a, b) = |\frac{1}{2}(a+b)|^2$. Sans surprise, cette probabilité est *indéterminée*, elle dépend des deux coefficients arbitraires a et b : en raison de l'identité des particules, aucun moyen n'existe pour fixer ces derniers, on ne peut d'aucune façon préciser davantage l'état de départ : cette ambiguïté conduit à l'impossibilité de prévoir le résultat de la deuxième mesure.

Cette difficulté peut être résolue en adoptant une prescription, qui sera en fait un postulat supplémentaire devant permettre de fixer a et b et d'obtenir alors une réponse unique à la question posée. L'énoncé de ce postulat sera donné au paragraphe suivant. Avant d'en venir là, quelques remarques méritent d'être faites.

Soit un système de N particules identiques ; oubliant pour l'instant leurs interactions mutuelles, il est possible de définir l'état de chacune d'entre elles par un jeu de nombres quantiques, α . L'état du système sera spécifié complètement en disant qu'il y a n_1 particules dans l'état α_1 , n_2 dans l'état α_2 , etc. On ne peut pas en dire plus en précisant par exemple que les n_1 premières sont dans α_1 , les n_2 suivantes dans α_2 etc. Pour un ensemble d'électrons, il est possible de calculer la probabilité de trouver un électron en un certain point avec un spin donné, mais c'est n'importe lequel des électrons constituant le système ; l'identification ne peut être poussée plus loin. En définitive, pour des particules identiques sans interaction, les bonnes variables quantiques sont les nombres d'occupation des différents états à une particule⁸. Cette reformulation porte le nom (impropre) de *Seconde Quantification*.

⁵Les deux opérateurs S_{1z} et S_{2z} commutent puisqu'ils agissent sur des variables différentes.

⁶L'arbitraire vient précisément de l'identité des particules : toute action extérieure (au moyen d'un champ par exemple) ne peut produire un effet sélectif puisque tous les attributs intrinsèques des deux particules sont strictement les mêmes.

⁷On peut objecter à l'argument en disant qu'il existe un moyen de fixer a et b : il suffit de procéder à une mesure du spin *total*. Si l'on a trouvé $S = 1$ alors l'état issu de la mesure est l'état triplet pour lequel $a = b = \frac{1}{\sqrt{2}}$; si on a trouvé $S = 0$, c'est l'état singulet, caractérisé par $a = -b = \frac{1}{\sqrt{2}}$. Ce faisant, la connaissance acquise porte sur un système de deux particules, où les caractéristiques individuelles de chacune se fondent dans le collectif.

⁸Ce formalisme n'est heureusement pas caduc en présence d'interactions. Si celles-ci sont faibles, on parle alors de quasi-particules, ou de particules habillées ; si elles sont fortes, on redéfinit de nouveaux modes pour lesquels il existe à nouveau des nombres d'occupation.

L'indiscernabilité des particules identiques a des conséquences extraordinaires, constituant le fondement de l'explication de la stabilité de la matière et de ses propriétés⁹.

D'une façon générale, à la notion d'indiscernabilité s'attachent deux grandes familles de particules, les fermions (spin demi-entier, $S = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$) et les bosons (spin entier, $S = 0, 1, 2, \dots$) ; chaque famille obéit à une statistique radicalement différente de celle à laquelle l'autre est soumise (Fermi-Dirac *vs* Bose-Einstein), donnant lieu à des comportements collectifs aux différences spectaculaires ; très schématiquement : les fermions n'aiment pas cohabiter dans la même case quantique, les bosons au contraire ont tendance à se rassembler (d'où, à basse température, le phénomène de condensation de Bose).

La structure des atomes, et la classification périodique, résulte¹⁰ du "Principe d'exclusion de Pauli", affirmation de principe énoncée à l'aube de la Mécanique Quantique mais qui est en réalité une conséquence du fait que les électrons au sein d'un même atome sont des fermions indiscernables. La distinction entre isolants et conducteurs résulte également de l'identité des électrons dans un métal et de leur sujétion à la statistique de Fermi - Dirac. Il convient d'ailleurs de signaler dès à présent que la nature fermionique ou bosonique d'un système composite dépend subtilement des interactions entre ses composants élémentaires : à basse température, les électrons d'un métal peuvent s'associer par paires¹¹ ayant toutes les caractéristiques de bosons¹² (paires de Cooper). Ces bosons composites peuvent se condenser et produisent l'état supraconducteur.

2.2 Le postulat de symétrisation

La difficulté stigmatisée dans la section précédente est éliminée par l'adoption du postulat suivant :

Pour un système contenant N particules identiques, les seuls états physiques sont soit symétriques (pairs) soit antisymétriques (impairs) dans l'échange de deux quelconques de ces particules.

Autrement dit, soit $\Psi(1, 2, \dots, N)$ la fonction d'onde d'un tel système où i dénote l'ensemble des degrés de liberté (spin et coordonnées d'espace au sens large) ; notons que i n'est pas le numéro d'une particule, mais le numéro d'un point de l'espace-spin d'une particule. Le postulat ci-dessus affirme que les seuls états possibles, suivant la nature des particules composant le système, sont :

1. les états symétriques tels que :

$$\Psi(\dots, i, \dots, j, \dots) = +\Psi(\dots, j, \dots, i, \dots) , \quad (2.5)$$

que l'on note génériquement Ψ_S et appartiennent à un sous-espace \mathcal{E}_S de l'espace des états \mathcal{E} .

2. les états antisymétriques tels que :

$$\Psi(\dots, i, \dots, j, \dots) = -\Psi(\dots, j, \dots, i, \dots) , \quad (2.6)$$

que l'on note génériquement Ψ_A et appartiennent à un sous-espace \mathcal{E}_A de l'espace des états \mathcal{E} .

On appelle bosons les particules pour lesquelles la fonction d'onde doit être symétrique, fermions celles pour lesquelles la fonction d'onde doit être antisymétrique. Toutes les particules connues actuellement vérifient en outre la règle empirique suivante :

⁹Voir la note 2.

¹⁰Il en va de même du noyau, tel qu'il est décrit par le modèle en couches.

¹¹On peut trouver énigmatique que deux électrons qui, suivant la loi de Coulomb, se repoussent fortement, peuvent néanmoins s'associer pour former un édifice stable. Suivant la théorie standard, dite théorie BCS, il existe de surcroît une interaction *attractive* colportée par les vibrations du réseau (phonons) permettant la formation de paires liées, dont on peut se faire la représentation physique suivante. Lorsqu'un électron se propage dans le réseau d'ions, il le déforme ; la déformation du réseau entraînant avec elle le fond continu de densité électronique, une lacune positive apparaît, *via* la conservation locale de la charge. Un autre électron survenant alors est attiré par cette polarisation positive et le résultat net est bien une interaction attractive dynamique entre les deux électrons, celui qui a créé la lacune positive et celui qui est attiré par elle. La supraconductivité n'existe qu'à basse température car les paires de Cooper sont fragiles et se cassent dès que l'énergie thermique $k_B T$ devient comparable à leur énergie de liaison.

¹²Compte tenu des règles relatives à l'addition des moments cinétiques, il faut évidemment un nombre *pair* de fermions pour former un spin total entier.

- les particules de spin entier sont des bosons, et sont donc décrits par des états symétriques satisfaisant (2.5)
- les particules de spin demi-entier sont des fermions, et sont donc décrits par des états antisymétriques satisfaisant (2.6).

Selon ce que l'on sait actuellement, la matière est faite de fermions, tandis que les bosons se chargent de médier les interactions. Dès que ce postulat est admis pour les particules considérées comme élémentaires, il s'applique également aux particules composites. En effet, échanger deux telles particules complexes revient à échanger simultanément toutes les particules composant la première avec toutes les particules composant la deuxième. L'état est inchangé si les constituants élémentaires sont des bosons (aucun changement de signe) ou s'il s'agit de fermions en nombre pair, un nombre pair de fermions donnant un spin entier donc un spin de boson. A l'inverse, des particules complexes formées d'un nombre impair de fermions sont des fermions (spin total demi-entier) et l'échange de deux telles particules par l'échange¹³ de leurs constituants élémentaires donnera globalement un signe $-$.

Il est important de remarquer que, dans tous les cas, l'échange de deux particules laisse toujours invariant le *module carré* de la fonction d'onde. De fait, échanger deux particules identiques ne doit pas modifier l'information contenue dans la fonction d'onde ; il faut donc que l'on ait :

$$|\Psi(\dots, i, \dots, j, \dots)|^2 = |\Psi(\dots, j, \dots, i, \dots)|^2 . \quad (2.7)$$

soit :

$$\Psi(\dots, i, \dots, j, \dots) = e^{i\alpha} \Psi(\dots, j, \dots, i, \dots) . \quad (2.8)$$

En définitive, le postulat de symétrisation affirme que la phase α ne peut être¹⁴ que 0 (bosons) ou π (fermions).

Remarques importantes

1. La propriété de symétrie dans l'échange se réfère essentiellement à la fonction d'onde *complète*, dépendant des degrés de liberté orbitaux ("coordonnées", c'est-à-dire coordonnées ou impulsions suivant la représentation choisie) et de ceux de spin. Pour en revenir à l'exemple des deux spins $\frac{1}{2}$ traité plus haut – en ne considérant toujours que les variables de spin –, on verra que le postulat de symétrisation permet de retenir, pour la partie de la fonction d'onde ne dépendant que du spin, deux combinaisons linéaires et deux seulement (définies comme suit, à une phase près) :

- (a) $a = +1$ et $b = +1$ correspondant à la composante $M_S = 0$ de l'état triplet (spin total $S = 1$) et donnant la combinaison :

$$|\chi_1\rangle = \frac{1}{\sqrt{2}} \left(\left| +\frac{1}{2}, -\frac{1}{2} \right\rangle + \left| -\frac{1}{2}, +\frac{1}{2} \right\rangle \right) . \quad (2.9)$$

- (b) $a = +1$ et $b = -1$ correspondant à l'état singulet (spin total $S = 0$) et donnant la combinaison :

$$|\chi_0\rangle = \frac{1}{\sqrt{2}} \left(\left| +\frac{1}{2}, -\frac{1}{2} \right\rangle - \left| -\frac{1}{2}, +\frac{1}{2} \right\rangle \right) . \quad (2.10)$$

La fonction de spin de l'état singulet est antisymétrique dans l'échange, celle de l'état triplet est au contraire symétrique – et pourtant il s'agit dans tous les cas d'un système de deux fermions dont l'état global doit être antisymétrique. La difficulté n'est qu'apparente : les vecteurs $|\chi_0\rangle$ et $|\chi_1\rangle$ ne prennent pas en compte les degrés de liberté orbitaux, il s'agit seulement de ce que l'on appelle parfois "l'état de spin", qui n'est que l'un des "morceaux" de l'état total (pour N quelconque, un tel état de spin peut en fait avoir

¹³C'est pourquoi les deux isotopes de l'Hélium ^3He et ^4He ont des propriétés radicalement différentes. ^3He est un fermion, alors que ^4He est un boson ; la superfluidité de ce dernier à basse température est la manifestation de la condensation de Bose.

¹⁴Voir aussi la remarque 4 p. 54.

une symétrie¹⁵ très différente : elle n'est pas quelconque mais entièrement déterminée par la valeur de M_S , et peut s'obtenir à l'aide de graphiques appelés *tableaux d'Young* [11]). La probabilité $P(a, b)$ (2.4) vaut 1 pour l'état triplet, 0 pour l'état singulet. Quoi qu'il en soit, le postulat de symétrisation supprime bien la difficulté présentée ci-dessus à propos des deux spins $\frac{1}{2}$, exemple qui doit être repris et complété afin de tenir compte des degrés orbitaux.

2. Pour des fermions, la fonction d'onde change de signe à chaque échange de deux particules – ce que l'on appelle une transposition dans la théorie du groupe des permutations. Si l'on fait deux échanges, la fonction d'onde au total ne change pas de signe. Par exemple :

$$\Psi_A(1, 2, 3, 4) = (-1)^2 \Psi_A(2, 1, 4, 3) . \quad (2.11)$$

Plus généralement, pour une permutation quelconque, Ψ acquiert un signe qui est la signature de la permutation, voir ci-dessous.

3. On rappellera ci-dessous qu'une permutation quelconque est factorisable en un produit de transpositions – ce qui permet de dire que les transpositions sont les “atomes” des permutations. Ceci étant réalisé, on voit qu'il faut, et il suffit, de dire ce que fait une fonction d'onde sous l'action d'une transposition, quelle que soit cette dernière. Pour des fermions, le postulat d'antisymétrisation affirme que toute transposition change le signe de la fonction d'onde
4. On rencontre parfois l'argument suivant. Considérons le cas $N = 2$; soit $f(1, 2)$ une fonction quelconque, et T la transposition qui échange 1 et 2 :

$$Tf(1, 2) = f(2, 1) ; \quad (2.12)$$

si on effectue à nouveau la transposition sur la fonction de droite, il vient :

$$Tf(2, 1) = f(1, 2) , \quad (2.13)$$

d'où l'on déduit $T^2 = \mathbf{1}$ – ce qui n'est pas un *scoop*. (2.13) montre aussi que les valeurs propres de T sont ± 1 , ce qui est indéniable, et que les fonctions propres de T , $\phi(1, 2)$, sont soit symétriques, soit antisymétriques puisqu'elles satisfont :

$$T\phi(1, 2) = (\pm 1)\phi(1, 2) . \quad (2.14)$$

Effectuons maintenant ces opérations avec une fonction d'onde $\Psi(1, 2)$; on est alors tenté d'écrire :

$$T\Psi(1, 2) = (\pm 1)\Psi(1, 2) , \quad (2.15)$$

ce qui tendrait à montrer que l'on peut se passer du postulat d'antisymétrisation – cette conclusion imposant l'idée que le raisonnement a une faille, ou n'est pas pertinent pour les états physiquement acceptables.

De fait, il n'en est rien et le postulat est bien nécessaire : imposer (2.15) à $\Psi(1, 2)$, c'est contraindre les états physiques à être propres de T , or rien ne l'exige : la seule nécessité est que Ψ et $T\Psi$ représentent la *même état physique*. Or si effectivement le carré de T est l'identité (l'opération qui ne fait rien), ceci n'exclut pas pour autant que l'on ait :

$$T^2\Psi(1, 2) = e^{i\alpha}\Psi(1, 2) . \quad (2.16)$$

¹⁵En outre, dans le cas où le Hamiltonien ne dépend pas du spin, on pourra toujours construire la fonction à symétriser conformément au postulat à partir du *produit* d'une fonction d'espace par une fonction de spin. La symétrisation détruit en général la séparation des variables espace-spin : après cette opération, la fonction d'onde n'est plus un simple produit de deux facteurs, l'un dépendant des coordonnées seules, l'autre des variables de spin, mais est une combinaison linéaire de tels produits ; ceux-ci s'échangent mutuellement avec les bons signes lors d'une permutation, en sorte que la fonction d'onde totale a la symétrie requise. La séparation en un seul produit ne subsiste que dans deux cas : pour $N = 2$ d'une part, et pour les composantes $M_S = \pm S$ de l'état de spin maximum $S = N\frac{\hbar}{2}$ pour N fermions de spin $\frac{1}{2}$. Dans tous les autres cas, la fonction d'onde est une combinaison linéaire de produits : c'est un exemple d'*intrication* (en anglais : *entanglement*), ici c'est une intrication entre degrés de liberté d'espace et degrés de liberté de spin.

En effet, T^2 ne fait rien à Ψ mais cette fonction et $e^{i\alpha}\Psi$ représentent un seul et même état¹⁶ : il n'y a donc *a priori* aucune contradiction à prendre α quelconque. Le postulat de symétrisation impose en fait $\alpha = 0$ ou π . En définitive, l'argument tendant à prouver que ces deux valeurs sont une *conséquence* de $T^2 = \mathbf{1}$ est spécieux¹⁷ : on a bel et bien besoin du postulat d'antisymétrisation.

Il est facile de voir que la symétrie de permutation de l'état d'un système donné ne change pas au cours du temps. Remarquons d'abord que pour un système de particules identiques, le Hamiltonien est inaltéré si on échange deux ou plusieurs particules : puisqu'elles sont identiques, elles figurent exactement de la même façon dans H . Il en va d'ailleurs de même pour toute autre observable de ce système : s'il n'en était pas ainsi, la mesure de cette observable permettrait justement de distinguer les particules or celles-ci sont précisément réputées identiques¹⁸. Par exemple, si \vec{S}_i désigne le spin de l'une quelconque des particules d'un système de N particules identiques, le spin total est :

$$\vec{S} = \sum_{i=1}^N \vec{S}_i \quad (2.17)$$

et est manifestement un opérateur invariant dans toute permutation. De même, soit le Hamiltonien d'un système de N particules identiques en interaction deux à deux par l'opérateur $W(r_{ij})$ où $r_{ij} = \|\vec{r}_i - \vec{r}_j\|$; $W(r_{ij})$ est par exemple l'interaction de Coulomb $\frac{e'^2}{r_{ij}}$ entre deux électrons. Toutes les particules ont notamment la même masse m par hypothèse et sont en général soumises individuellement à un champ de force donnant à chacune l'énergie potentielle $V(\vec{r}_i)$. Dans ces conditions, le Hamiltonien s'écrit :

$$H = \sum_{i=1}^N \left[\frac{\vec{p}_i^2}{2m} + V(\vec{r}_i) \right] + \frac{1}{2} \sum_{i=1}^N \sum_{j=1, j \neq i}^N W(r_{ij}) ; \quad (2.18)$$

le facteur $\frac{1}{2}$ évite de compter deux fois l'interaction entre deux particules données¹⁹. À nouveau, H est visiblement symétrique dans toute permutation des indices.

Montrons maintenant que la propriété de symétrie de permutation (symétrie ou antisymétrie) se conserve au cours du temps. Ceci assure qu'il ne peut y avoir de contradiction entre les postulats généraux énoncés antérieurement et le postulat de symétrisation, étant entendu que toute observable pouvant faire l'objet d'une mesure est nécessairement totalement symétrique. Chaque espace \mathcal{E}_A ou \mathcal{E}_S , où se trouve le vecteur d'état d'un système à un instant donné, est stable par l'évolution dans le temps et tout ce qui a été dit précédemment reste vrai dans ce sous-espace dont le système ne peut sortir.

Soit un état initial $\Psi(1, 2, \dots, N ; t)$ de symétrie donnée (symétrique ou antisymétrique). Par l'équation de Schrödinger, on a :

$$\Psi(1, 2, \dots, N ; t + dt) = \Psi(1, 2, \dots, N ; t) + \frac{dt}{i\hbar} H(1, \dots, N) \Psi(1, 2, \dots, N ; t) + \mathcal{O}(dt^2) . \quad (2.20)$$

Effectuons l'échange de deux quelconques des particules dans les deux membres. Il vient :

$$\Psi(\dots, j, \dots, i, \dots ; t + dt) = \Psi(\dots, j, \dots, i, \dots ; t) + \frac{dt}{i\hbar} H(1, \dots, N) \Psi(\dots, j, \dots, i, \dots ; t) + \mathcal{O}(dt^2) . \quad (2.21)$$

H ne change pas. Si $\Psi(\dots ; t)$ est une fonction paire (symétrique) Ψ_S , le second membre ne change pas et est donc égal à $\Psi_S(\dots, i, \dots, j, \dots ; t + dt)$, d'où :

$$\Psi_S(\dots, j, \dots, i, \dots ; t + dt) = \Psi_S(\dots, i, \dots, j, \dots ; t + dt) \quad (2.22)$$

¹⁶La phase α est ici bien sûr supposée constante. Une phase variable (fonction de la coordonnée) se rencontre par exemple lorsque l'on effectue une transformation de jauge, voir [8], ch. III, complément H_{III}.

¹⁷Merci à Gaëtan Borot pour une discussion très éclairante sur ce sujet.

¹⁸Notons que la réciproque est fautive : pour deux particules non-identiques, le Hamiltonien peut néanmoins être symétrique. Ainsi, le Hamiltonien du couple électron-positron libre (réduit à l'interaction électrostatique) est symétrique, et pourtant l'électron et le positron, différant par leur charge, ne sont pas deux particules identiques.

¹⁹De ce point de vue, le terme à deux corps dans (2.18) s'écrit tout autant :

$$\sum_{i=1}^N \sum_{j=i+1}^N W(r_{ij}) . \quad (2.19)$$

et la fonction d'onde à l'instant $t + dt$ est encore paire. Au contraire, si $\Psi(\dots; t)$ est une fonction impaire (antisymétrique) Ψ_A , on a maintenant :

$$\Psi_A(\dots, j, \dots, i, \dots; t+dt) = -\Psi_A(\dots, i, \dots, j, \dots; t) - \frac{dt}{i\hbar} H(1, \dots, N) \Psi_A(\dots, i, \dots, j, \dots; t) + \mathcal{O}(dt^2) , \quad (2.23)$$

puisque les deux fonctions au second membre changent de signe par hypothèse mais pas H . Il en résulte :

$$\Psi_A(\dots, j, \dots, i, \dots; t + dt) = -\Psi_A(\dots, i, \dots, j, \dots; t + dt) . \quad (2.24)$$

Ainsi, de proche en proche, le caractère de symétrie de l'état initial se perpétue au cours du temps : c'est ce que l'on appelle parfois une règle de super-sélection. Ceci a des conséquences importantes : par exemple, un système de bosons ne peut évoluer en se fragmentant en un nombre impair de fermions. Tout morcellement ne peut donc se faire que par émission de fermions par paires de façon à préserver le caractère bosonique de l'état. De même, un système de fermions ne peut créer un boson "libre" que si celui-ci est formé de deux fermions. D'un autre côté, deux fermions peuvent échanger des bosons virtuels, particules véhiculant l'interaction entre les deux fermions.

2.3 Permutations. Opérateurs de symétrisation et d'antisymétrisation

Soit $\mathcal{S}_N \equiv \{P_\lambda\}$ l'ensemble des permutations de N objets. Une permutation peut se noter :

$$\begin{pmatrix} 1 & 2 & \dots & N \\ i_1 & i_2 & \dots & i_n \end{pmatrix} , \quad (2.25)$$

la deuxième ligne donnant le résultat de la permutation des N premiers entiers écrits dans l'ordre naturel dans la première ligne. Cet ensemble est visiblement un groupe ; le produit de deux permutations est une permutation, que l'on détermine comme suit :

$$\begin{pmatrix} 1 & 2 & \dots & N \\ i_1 & i_2 & \dots & i_n \end{pmatrix} \begin{pmatrix} 1 & 2 & \dots & N \\ j_1 & j_2 & \dots & j_n \end{pmatrix} = \begin{pmatrix} j_1 & j_2 & \dots & j_N \\ k_1 & k_2 & \dots & k_n \end{pmatrix} \begin{pmatrix} 1 & 2 & \dots & N \\ j_1 & j_2 & \dots & j_n \end{pmatrix} . \quad (2.26)$$

Le produit au second membre n'est autre que :

$$\begin{pmatrix} 1 & 2 & \dots & N \\ k_1 & k_2 & \dots & k_n \end{pmatrix} . \quad (2.27)$$

L'élément neutre est la permutation identité $\begin{pmatrix} 1 & 2 & \dots & N \\ 1 & 2 & \dots & N \end{pmatrix}$; manifestement, tout élément a un inverse. Il y a $N!$ éléments dans \mathcal{S}_N , qui est appelé groupe symétrique.

Les P_λ agissent sur les fonctions d'onde comme suit ; si :

$$P_\lambda = \begin{pmatrix} 1 & 2 & \dots & N \\ i_1 & i_2 & \dots & i_N \end{pmatrix} , \quad (2.28)$$

alors :

$$P_\lambda \Psi(1, 2, \dots, N) = \Psi(i_1, i_2, \dots, i_N) , \quad (2.29)$$

Les P_λ sont visiblement des opérateurs unitaires ; en effet, un produit scalaire tel que :

$$(P_\lambda \Psi, P_\lambda \Phi) \quad (2.30)$$

est une intégration sur les variables d'espace continues et une sommation sur les variables de spin discrètes ; toutes ces sommations portent sur des variables muettes et le résultat ne dépend en aucune façon de toute permutation effectuée sur *l'ensemble* de ces variables. Il vient donc :

$$(P_\lambda \Psi, P_\lambda \Phi) = (\Psi, \Phi) \quad (2.31)$$

qui est la relation caractéristique d'un opérateur unitaire. Toute permutation étant unitaire, son adjoint est son inverse :

$$P_\lambda^\dagger = P_\lambda^{-1} . \quad (2.32)$$

En général, une permutation n'est pas égale à son inverse, et n'est donc pas hermitique – sauf dans le cas des permutations particulières appelées transpositions, voir ci-dessous.

Un lemme sera utile, dit “lemme de réarrangement”. Soit une permutation P_μ donnée ; En effectuant tous les produits $P_\mu P_\lambda$ où λ décrit le groupe *entier*, on obtient une fois et une seule les $N!$ permutations. En effet : le produit $P_\mu P_\lambda$ est par définition de la structure de groupe une certaine permutation P_ν . Quand on effectue ces $N!$ produits, si on trouvait deux fois la même permutation P_ν , cela signifierait qu'il existe λ_1 et λ_2 tels que :

$$P_\mu P_{\lambda_1} = P_\nu \quad P_\mu P_{\lambda_2} = P_\nu \iff P_{\lambda_1} = P_\mu^{-1} P_\nu \quad P_{\lambda_2} = P_\mu^{-1} P_\nu \quad (2.33)$$

d'où $P_{\lambda_1} = P_{\lambda_2}$. Le lemme de réarrangement peut s'écrire formellement :

$$P_\lambda \mathcal{S}_N = \mathcal{S}_N . \quad (2.34)$$

Certaines permutations jouent un rôle particulièrement important, ce sont les transpositions. Une transposition échange deux objets seulement ; notant T_{ij} la permutation qui se borne à échanger i et j , on a donc par définition :

$$T_{ij} = \begin{pmatrix} 1 & 2 & 3 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & N \\ 1 & 2 & 3 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & N \end{pmatrix} , \quad (2.35)$$

que l'on peut noter plus simplement :

$$T_{ij} = \begin{pmatrix} i & j \\ j & i \end{pmatrix} . \quad (2.36)$$

Comme toute permutation, les transpositions sont unitaires, mais comme leur carré est l'identité, elles sont également hermitiques :

$$T_{ij}^2 = \mathbf{1} \iff T_{ij} = T_{ij}^{-1} = T_{ij}^\dagger . \quad (2.37)$$

Enfin, évidemment, $T_{ij} = T_{ji}$.

Un résultat important est le suivant : toute permutation peut se décomposer en produit de transpositions, la décomposition n'étant pas unique (on peut en écrire autant qu'on veut) mais le nombre σ de transpositions apparaissant dans le produit a une parité parfaitement définie, que l'on appelle la parité de la permutation ; le nombre $(-1)^\sigma$ est appelé signature de la permutation – la signature d'une transposition est égale à -1 . Par exemple, la permutation circulaire :

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad (2.38)$$

peut se décomposer comme suit :

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = T_{12} T_{23} . \quad (2.39)$$

Pour cette permutation, σ vaut 2 (ou n'importe quel nombre pair), il s'agit donc d'une permutation paire. Les deux permutations circulaires, pour $N = 3$, ont pour signature $(-1)^2 = +1$.

La signature de la permutation P_λ , $(-1)^{\sigma_\lambda}$, permet d'exprimer simplement le résultat d'une permutation quelconque agissant sur une fonction d'onde de fermions :

$$P_\lambda \Psi_A(1, 2, \dots, N) = (-1)^{\sigma_\lambda} \Psi_A(1, 2, \dots, N) \quad \forall P_\lambda \quad (\text{fermions}) ; \quad (2.40)$$

si on permute deux arguments (ou un nombre pair d'arguments) d'une fonction antisymétrique, celle-ci au total ne change pas de signe. Evidemment, pour des bosons, on a :

$$P_\lambda \Psi_S(1, 2, \dots, N) = \Psi_S(1, 2, \dots, N) \quad \forall P_\lambda \quad (\text{bosons}) . \quad (2.41)$$

Ces définitions étant données, on peut maintenant montrer comment construire les fonctions physiquement acceptables, c'est-à-dire ayant la symétrie requise vis-à-vis des permutations. L'idée est de savoir engendrer, à partir d'une fonction Φ donnée obtenue d'une façon ou d'une autre, la bonne fonction Φ_A ou Φ_S , suivant la nature du système considéré.

Avant de donner la méthode générale, on peut en comprendre l'idée centrale en considérant le cas simple de $N = 2$ particules. Une fonction $\Phi(1, 2)$ étant donnée, il est facile d'en déduire les fonctions ayant la bonne symétrie ; ainsi, la combinaison linéaire $\Phi(1, 2) + \Phi(2, 1)$ est symétrique dans l'échange, en revanche $\Phi(1, 2) - \Phi(2, 1)$ est antisymétrique dans l'échange ; ces deux fonctions sont donc respectivement du type Φ_S et Φ_A et peuvent en outre s'écrire²⁰ :

$$\Phi_S = (1 + T_{12}) \Phi(1, 2) \quad \Phi_A = (1 - T_{12}) \Phi(1, 2) \quad (2.42)$$

On voit ainsi qu'en faisant agir sur une fonction quelconque la bonne combinaison linéaire d'opérateurs de permutation, on obtient la fonction ayant la symétrie voulue. C'est cette idée que l'on généralise maintenant, en définissant les deux projecteurs Y_S et Y_A :

$$Y_S = \frac{1}{N!} \sum_{\lambda=1}^{N!} P_\lambda \quad Y_A = \frac{1}{N!} \sum_{\lambda=1}^{N!} (-1)^{\sigma_\lambda} P_\lambda \quad (2.43)$$

que l'on peut collectivement dénoter Y_I :

$$Y_I = \frac{1}{N!} \sum_{\lambda=1}^{N!} \varepsilon_\lambda P_\lambda \quad \varepsilon_\lambda = \begin{cases} 1 & (\text{I=S, bosons}) \\ (-1)^{\sigma_\lambda} & (\text{I=A, fermions}) \end{cases} . \quad (2.44)$$

Il est facile de voir que $[Y_S, Y_A] = 0$ et que, plus précisément :

$$Y_S Y_A = Y_A Y_S = 0 . \quad (2.45)$$

En effet, notons d'abord que :

$$P_\nu = P_\lambda P_\mu \implies \sigma_\nu = \sigma_\lambda + \sigma_\mu \iff \varepsilon_\nu = \varepsilon_\lambda \varepsilon_\mu \iff \varepsilon_\lambda = \varepsilon_\mu \varepsilon_\nu , \quad (2.46)$$

par définition des σ : le nombre de transpositions en lesquelles se décompose P_ν est la somme des deux nombres correspondants pour P_λ et P_μ ; la dernière égalité dans (2.46) s'obtient en multipliant la précédente membre à membre par ε_μ et sachant que $\varepsilon_\mu^2 = 1$ dans tous les cas. Maintenant :

$$Y_A Y_S = \frac{1}{N!^2} \sum_{\lambda, \mu=1}^{N!} \varepsilon_\lambda P_\lambda P_\mu = \frac{1}{N!^2} \sum_{\nu, \mu=1}^{N!} \varepsilon_\mu \varepsilon_\nu P_\nu = \frac{1}{N!^2} \left(\sum_{\mu=1}^{N!} \varepsilon_\mu \right) \sum_{\nu=1}^{N!} \varepsilon_\nu P_\nu . \quad (2.47)$$

On trouve visiblement la même expression en effectuant le produit dans l'autre sens, soit $Y_S Y_A$, d'où $[Y_S, Y_A] = 0$. Maintenant, soit T une transposition quelconque ; comme $T Y_A = -Y_A$:

$$T Y_A Y_S = -Y_A Y_S = -Y_S Y_A , \quad (2.48)$$

où la dernière égalité vient de la commutation des Y_I entre eux. En multipliant à nouveau à gauche par T , et puisque $T^2 = \mathbf{1}$:

$$T^2 Y_A Y_S = -T Y_S Y_A \iff Y_A Y_S = -T Y_S Y_A ; \quad (2.49)$$

par ailleurs, $T Y_S = Y_S$, d'où finalement :

$$Y_A Y_S = -Y_S Y_A , \quad (2.50)$$

²⁰Ces combinaisons linéaires ne sont pas normalisées, même si Φ l'est.

ce qui, allié au fait que les Y_I en fait commutent, établit (2.45). Par comparaison avec (2.47), on en déduit aussi que :

$$\sum_{\mu=1}^{N!} \varepsilon_{\mu} = 0 . \quad (2.51)$$

De plus, Y_I est bien un projecteur :

$$Y_I^2 = \frac{1}{N!^2} \sum_{\lambda=1}^{N!} \sum_{\mu=1}^{N!} \varepsilon_{\lambda} \varepsilon_{\mu} P_{\lambda} P_{\mu} = \frac{1}{N!^2} \sum_{\nu=1}^{N!} \varepsilon_{\nu} P_{\nu} \sum_{\lambda=1}^{N!} 1 . \quad (2.52)$$

La dernière forme est obtenue à partir du lemme de réarrangement. La deuxième sommation donne le simple facteur $N!$, de sorte que :

$$Y_I^2 = Y_I . \quad (2.53)$$

Y_I est donc idempotent ; en outre, il s'agit visiblement d'un opérateur hermitique puisque chaque P_{λ} est unitaire : dans la sommation dont on a pris l'adjoint, P_{λ}^{\dagger} peut être remplacé par P_{λ}^{-1} , qui est un autre terme de la même sommation, avec évidemment²¹ le même ε_{λ} . Comme Y_I est idempotent et hermitique, c'est un projecteur.

Calculons maintenant le produit $P_{\mu} Y_I$, où P_{μ} est une permutation quelconque mais donnée. Il vient :

$$P_{\mu} Y_I = \frac{1}{N!} \sum_{\lambda=1}^{N!} \varepsilon_{\lambda} P_{\mu} P_{\lambda} = \frac{1}{N!} \sum_{\nu=1}^{N!} \varepsilon_{\mu} \varepsilon_{\nu} P_{\nu} \equiv \varepsilon_{\mu} Y_I . \quad (2.54)$$

Soit maintenant une fonction quelconque Φ ; le résultat de l'action de Y_I sur Φ produit soit une fonction de symétrie I, soit une fonction identiquement nulle, voir plus loin. En effet, par exemple avec Y_A , on trouve, compte tenu de (2.54) :

$$P_{\mu} Y_A \Phi = (-1)^{\sigma_{\mu}} Y_A \Phi . \quad (2.55)$$

La fonction $Y_A \Phi$ est donc une fonction antisymétrique dans toute transposition. De la même façon, on a :

$$P_{\mu} Y_S \Phi = Y_S \Phi . \quad (2.56)$$

Ainsi, Y_I projette toute fonction donnée de symétrie quelconque sur le sous-espace \mathcal{E}_I ayant la symétrie voulue. Cette opération engendre donc la bonne fonction, sauf toutefois dans un cas : celui où la fonction Φ de départ a déjà la symétrie "orthogonale" à celle voulue. Alors, l'action de Y_I projette un vecteur orthogonal au sous-espace de projection et on trouve zéro²² (usuellement, un peu de discernement permet de voir d'emblée si l'on est dans un tel cas particulier !). Notons enfin que deux fonctions de symétrie différente A et S sont, d'après (2.45), orthogonales par construction.

Remarquons que la fonction construite par $Y_I \Phi$ n'est pas en général normalisée à l'unité, même si Φ l'est, et qu'il convient de la normaliser après coup. D'une façon générale, les fonctions normalisées de symétrie requise seront :

$$\Psi_A = C_A Y_A \Phi \quad \Psi_S = C_S Y_S \Phi , \quad (2.57)$$

où les C_I sont des constantes de normalisation. Compte tenu des propriétés de Y_I , on a :

$$\| \Psi_I \|^2 = |C_I|^2 \langle \Phi | Y_I \Phi \rangle = \frac{1}{N!} |C_I|^2 \sum_{\lambda=1}^{N!} \varepsilon_{\lambda} \langle \Phi | P_{\lambda} \Phi \rangle ; \quad (2.58)$$

l'expression de C_I dépend donc de tous les produits scalaires $\langle \Phi | P_{\lambda} \Phi \rangle$.

²¹La factorisation de P_{λ}^{-1} s'obtient en renversant l'ordre des facteurs dans celle de P_{λ} : ceci ne change pas le nombre de facteurs !
²²Autre façon de voir : si Φ a déjà une symétrie I ou S, alors $Y_I \Phi = \Phi$. Dès lors $Y_I \Phi = Y_I Y_I \Phi = \delta_{II'} \Phi$ en vertu de (2.45).

Traitons un exemple avec $N = 3$. Soit $\Phi(1, 2, 3)$ une fonction quelconque. La fonction symétrique engendrée à partir de Φ est :

$$\begin{aligned}\Psi_S &= C_S \left[\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right. \\ &\quad \left. + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right] \Phi(1, 2, 3) \\ &= C_S \frac{1}{3!} [\Phi(1, 2, 3) + \Phi(2, 1, 3) + \Phi(3, 2, 1) + \Phi(1, 3, 2) + \Phi(2, 3, 1) + \Phi(3, 1, 2)] .\end{aligned}\quad (2.59)$$

Il est manifeste que cette dernière fonction est complètement symétrique dans toute permutation. De même :

$$\begin{aligned}\Psi_A &= C_A \left[\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right. \\ &\quad \left. + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right] \Phi(1, 2, 3) \\ &= C_A \frac{1}{3!} [\Phi(1, 2, 3) - \Phi(2, 1, 3) - \Phi(3, 2, 1) - \Phi(1, 3, 2) + \Phi(2, 3, 1) + \Phi(3, 1, 2)] .\end{aligned}\quad (2.60)$$

Il est non moins évident que Ψ_A change de signe dans toute transposition mais ne change pas si l'on applique l'une des deux permutations circulaires de \mathcal{S}_3 . D'un autre côté, on voit bien sur cet exemple que si Φ est déjà complètement symétrique, la fonction Ψ_A ainsi fabriquée est identiquement nulle : il y a autant de signes + que de signes -. Enfin, il est visible que Ψ_A et Ψ_S sont orthogonales²³.

Pour illustrer le plus simplement possible la différence fondamentale entre bosons et fermions, considérons un système formé de deux particules identiques dont l'état est construit à partir d'un simple produit de deux fonctions ψ et ϕ supposées orthonormalisées. Compte tenu du postulat de symétrisation, les seuls états physiquement acceptables sont de la forme :

$$\Psi(1, 2) = C [\mathbf{1} + \varepsilon T_{12}] \psi(1)\phi(2) \quad (2.61)$$

où $\varepsilon = +1$ s'il s'agit de bosons, $\varepsilon = -1$ s'il s'agit de fermions ; comme ψ et ϕ sont orthonormalisées, on a simplement (à une phase près) $C = \frac{1}{\sqrt{2}}$. Ceci peut s'écrire en notation de Dirac :

$$|\Psi\rangle = \frac{1}{\sqrt{2}} [\mathbf{1} + \varepsilon T_{12}] |\psi \phi\rangle \quad (2.62)$$

à condition de convenir implicitement que dans le ket (et plus tard dans les bras) les variables sont écrites dans l'ordre naturel de gauche à droite. Soit A une observable à une particule, $A(1, 2) = a(1) + a(2)$, dont le spectre est supposé entièrement discret, pour simplifier :

$$a |a_n\rangle = a_n |a_n\rangle \quad A |a_n a_m\rangle = (a_n + a_m) |a_n a_m\rangle . \quad (2.63)$$

Cherchons maintenant la probabilité pour qu'une mesure de A sur les deux particules fournisse la valeur a_n pour l'une et a_m pour l'autre. L'état issu de la mesure ayant fourni ces valeurs est lui aussi un état de deux particules identiques. En tant que tel, il doit satisfaire le postulat de symétrisation et est donc de la forme :

$$|\Psi'\rangle = \frac{1}{\sqrt{2}} [\mathbf{1} + \varepsilon T_{12}] |a_n a_m\rangle . \quad (2.64)$$

La probabilité de trouver ce résultat est donc :

$$P = |\langle \Psi' | \Psi \rangle|^2 = \left| \frac{1}{2} \langle a_n a_m | [\mathbf{1} + \varepsilon T_{12}^\dagger] [\mathbf{1} + \varepsilon T_{12}] |\psi \phi\rangle \right|^2 . \quad (2.65)$$

²³Dans le langage de la Théorie de la Représentation (linéaire) des Groupes, on dit que Ψ_S et Ψ_A se transforment respectivement suivant les représentations irréductibles symétrique et antisymétrique. Deux fonctions se transformant suivant deux représentations irréductibles inéquivalentes sont forcément orthogonales.

Il apparaît deux termes : celui venant de $\mathbf{1}$ est dit “terme direct”, celui venant de T_{12} est appelé “terme d'échange”. Comme T_{12} est hermitique et que son carré est l'identité, il vient simplement :

$$P = |\langle a_n a_m | [\mathbf{1} + \varepsilon T_{12}] | \psi \phi \rangle|^2 = |\langle a_n a_m | \psi \phi \rangle + \varepsilon \langle a_n a_m | \phi \psi \rangle|^2 . \quad (2.66)$$

Ainsi, dans l'amplitude donnant la probabilité P , les bosons interfèrent avec le signe $+$, les fermions avec le signe $-$. Bien évidemment, la probabilité P est nulle pour deux fermions.

Remarque

Les opérateurs Y_I permettent de fabriquer automatiquement les fonctions ayant la symétrie voulue, à partir d'une fonction Φ convenablement choisie. Par exemple, Φ peut être sélectionnée en tant que fonction propre d'un ECOC ; dans le cas le plus simple, Φ est seulement propre du Hamiltonien, $H\Phi = E\Phi$.

Il est bien clair que l'action de Y_I sur Φ ne modifie pas son statut d'état propre ; en effet, H est invariant dans toute permutation, donc $[P_\lambda, H] = 0 \forall \lambda$, donc également $[Y_I, H] = 0$. On a donc :

$$Y_I H \Phi = Y_I E \Phi = E Y_I \Phi , \quad (2.67)$$

d'une part ; d'autre part, en vertu de $Y_I H = H Y_I$, il vient :

$$Y_I H \Phi = H (Y_I \Phi) = E (Y_I \Phi) , \quad (2.68)$$

où les parenthèses ne sont là que pour la clarté. (2.68) montre que si Φ est propre de H , alors $Y_I \Phi$ est encore propre²⁴ de H . Le même argument vaut pour toute observable d'un système de N particule identiques. En définitive, les projecteurs fabriquent les fonctions de bonne symétrie sans altérer leur éventuel caractère d'état propre.

2.4 Etats d'un système de particules indépendantes : différence fondamentale entre bosons et fermions

On dit que N particules sont indépendantes s'il n'existe aucune interaction entre deux quelconques d'entre elles ; autrement dit, le Hamiltonien du système qu'elles constituent ne contient aucun terme dépendant simultanément des grandeurs dynamiques (coordonnées, impulsions, spins, *etc.*) de deux d'entre elles. Il n'empêche que chacune d'entre elles peut être soumise individuellement à un champ de forces dérivant d'un potentiel et lui donnant l'énergie potentielle $V(\vec{r}_i)$. Dans ces conditions, H s'écrit :

$$H(1, 2, \dots, N) = \sum_{i=1}^N \left[\frac{\vec{p}_i^2}{2m} + V(\vec{r}_i) \right] \equiv \sum_{i=1}^N H_1(i) . \quad (2.69)$$

H est la somme de Hamiltoniens à une particule, H_1 , tous identiques et ne différant dans la sommation que par l'indice i des variables dont ils dépendent. Désignons maintenant par ψ_k et E_k les modes propres de H_1 :

$$H_1 \psi_k(\rho) = E_k \psi_k(\rho) \quad (2.70)$$

où ρ dénote collectivement les coordonnées d'espace et la variable de spin. Si H_1 ne dépend pas du spin, la fonction ψ_k peut toujours être prise comme le produit d'une fonction ϕ_k par une fonction de spin χ , ϕ_k étant propre de H_1 :

$$\psi_k(\rho) \equiv \phi_k(\vec{r}) \chi(m_s) \quad H_1 \phi_k(\vec{r}) = E_k \phi_k(\vec{r}) . \quad (2.71)$$

Suivant la terminologie usuelle, la fonction $\psi_k(\rho)$ est appelée spin-orbitale, la fonction $\phi_k(\vec{r})$ est appelée orbitale. Dans la suite, quand c'est possible, on continue à raisonner en toute généralité avec les spin-orbitales.

²⁴D'ailleurs, comme toutes les fonctions $P_\lambda \Phi$ sont propres de H avec la même valeur propre E , n'importe quelle combinaison $[\sum_\lambda c_\lambda P_\lambda] \Phi$ reste propre. Les combinaisons résultant des deux projecteurs $Y_I \Phi$ ne sont que deux combinaisons particulières.

Comme H est une somme d'opérateurs H_1 à un corps, tout produit de spin-orbitales ψ_k sera fonction propre de H avec la valeur propre égale à la somme des valeurs propres associées aux spin-orbitales :

$$\begin{aligned} H \prod_{j=1}^N \psi_{k_j}(\rho_j) &= \sum_{i=1}^N \prod_{j=1}^N H_1(\rho_i) \psi_{k_j}(\rho_j) = \sum_{i=1}^N \left[\prod_{j=1, j \neq i}^N \psi_{k_j}(\rho_j) \right] H_1(\rho_i) \psi_{k_i}(\rho_i) \\ &= \sum_{i=1}^N \left[\prod_{j=1, j \neq i}^N \psi_{k_j}(\rho_j) \right] E_{k_i} \psi_{k_i}(\rho_i) = \sum_{i=1}^N E_{k_i} \prod_{j=1}^N \psi_{k_j}(\rho_j) \end{aligned} \quad (2.72)$$

On a donc :

$$H \prod_{j=1}^N \psi_{k_j}(\rho_j) = E_{k_1, k_2, \dots, k_N} \prod_{j=1}^N \psi_{k_j}(\rho_j) \quad E_{k_1, k_2, \dots, k_N} = \sum_{j=1}^N E_{k_j} , \quad (2.73)$$

un résultat qui est physiquement évident : en l'absence d'interactions, l'énergie du système des N particules est simplement la somme des énergies individuelles. Ceci étant, le produit $\prod_{j=1}^N \psi_{k_j}$, qui joue le rôle de la fonction notée Φ dans la discussion générale, n'est visiblement pas une fonction physiquement acceptable vis-à-vis du postulat de symétrisation ; il convient maintenant de lui appliquer l'un des opérateurs Y_1 (2.44) afin d'obtenir une fonction (non normalisée) de symétrie convenable. Considérons d'abord le cas des bosons. Une bonne fonction est :

$$\Psi_S(1, 2, \dots, N) = C_S Y_S \prod_{i=1}^N \psi_{k_i}(\rho_i) \equiv C_S Y_S \Phi(1, 2, \dots, N) \quad (2.74)$$

l'énergie associée à Ψ_S étant la somme des E_{k_i} , conformément à (2.73). En particulier, l'état fondamental des N bosons sera construit en prenant pour chaque ψ_{k_i} l'état fondamental de H_1 , ψ_1 , d'énergie E_1 donnant une énergie totale $E_{1, 1, \dots, 1} = N E_1$. Ainsi, l'état fondamental²⁵ d'un système de bosons sans interactions est un état dans lequel tous les bosons sont chacun dans son état fondamental ; on dit que tous les bosons sont *condensés* dans le même état. Cette propriété remarquable ouvre la possibilité, à température finie et pour un système à nombre de bosons fixé, au phénomène appelé condensation de Bose²⁶. L'état fondamental des N bosons est alors :

$$\Psi_{\text{fond. bosons}} = \psi_1(\rho_1) \psi_1(\rho_2) \dots \psi_1(\rho_N) . \quad (2.75)$$

Cette fonction est visiblement symétrique et il ne servirait à rien de lui appliquer Y_S . Si les spin-orbitales $\psi_k(\rho)$ sont orthonormalisées, $\Psi_{\text{fond. bosons}}$ est *de facto* normalisée à l'unité.

Pour les fermions, il en va tout autrement. Comme la fonction d'onde totale doit être antisymétrique dans l'échange de deux fermions, on voit tout de suite que si deux fermions sont placés dans le même état quantique, la fonction d'onde est identiquement nulle. Pour bien voir ce fait fondamental, soit un système de deux fermions indépendants et plaçons-les tous deux dans le même état ψ_k (même partie orbitale, même valeur pour la projection du spin). La fonction antisymétrique convenable s'obtient en appliquant Y_A qui, pour $N = 2$, s'écrit simplement :

$$Y_A = \frac{1}{2!} \left[\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right] . \quad (2.76)$$

D'où il résulte :

$$\Psi_A(1, 2) = C_A Y_A \psi_k(1) \psi_k(2) = \frac{C_A}{2!} [\psi_k(1) \psi_k(2) - \psi_k(2) \psi_k(1)] \equiv 0 . \quad (2.77)$$

²⁵ Il s'agit bien sûr de l'état à température *nulle*.

²⁶ À température finie, le jeu des fluctuations thermiques peut supprimer la condensation de Bose si la dimensionnalité d est faible : l'argument standard montre qu'il n'y a pas de condensation de Bose si $d \leq 2$ (il donne $T_C \propto [\zeta(d/2)]^{-2/d}$ où ζ est la fonction de Riemann). Pour $d > 2$, la condensation se produit à une certaine température T_C : pour $T < T_C$, une fraction *macroscopique* de bosons est dans l'état fondamental ; cette fraction tend vers 1, quand la température tend vers zéro, suivant la loi-puissance $1 - \left(\frac{T}{T_C}\right)^{\frac{d}{2}}$ ($d > 2$).

Il est tout à fait remarquable qu'une transition de phase se produise pour un système de particules sans interactions au sens classique. En réalité, le postulat de symétrisation introduit une "interaction" (quantique !), que l'on peut schématiquement qualifier d'attractive pour les bosons et de répulsive pour les fermions (*trou de Fermi*, voir ci-dessous) ; en atteste aussi la première correction quantique de la fonction de partition du gaz parfait classique, qui se traduit par un terme additionnel que l'on peut interpréter comme une interaction effective (dépendant de la température), attractive pour les bosons, répulsive pour les fermions. Par ailleurs, la question de la condensation en présence d'interaction entre les bosons est actuellement un problème largement ouvert.

Ainsi, pour des fermions, l'antisymétrie exige que deux d'entre eux n'occupent pas le même état quantique, faute de quoi l'état est identiquement nul – en d'autres termes n'existe pas en tant qu'état possible ; c'est ce qui fut historiquement (et empiriquement) énoncé sous la forme du "Principe d'exclusion de Pauli", pour rendre compte des spectres atomiques, de la structure en couches des atomes et de la classification périodique.

Pour fabriquer l'état fondamental d'un système de fermions sans interactions, il faut donc prendre successivement tous les états propres de H_1 et placer dans chacun d'entre eux un et un seul électron, compte tenu du spin. En notant $\psi_1, \psi_2, \dots, \psi_n, \dots$ les états classés par ordre d'énergies croissantes $E_1 \leq E_2 \leq \dots \leq E_n, \dots$, l'état fondamental se construit à l'aide de Y_A et du simple produit $\psi_1 \psi_2 \dots \psi_N$: c'est bien l'état de plus basse énergie compatible avec le principe d'exclusion. Toutes choses égales par ailleurs²⁷, l'énergie totale d'un système de N fermions est donc considérablement plus grande que celle d'un système de N bosons puisque l'on est dans l'obligation d'aller chercher des états excités de H_1 , chaque état propre de H_1 ne pouvant recevoir qu'un seul fermion. En tout cas, le fondamental s'écrit comme suit à l'aide des spin-orbitales :

$$\Psi_A(1, 2, \dots, N) = C_A Y_A [\psi_1(\rho_1) \psi_2(\rho_2) \dots \psi_N(\rho_N)] . \quad (2.78)$$

Comme déjà mentionné, il arrive souvent que les états propres de H_1 ne dépendent pas de la variable de spin, m_s . En pareil cas, chaque ψ_k est le produit d'une orbitale $\phi_k(\vec{r})$ et d'une fonction de spin χ . Dans le cas d'électrons, $S = \frac{1}{2}$, la spin-orbitale est soit $\phi_k(\vec{r})\alpha(m_s)$, soit $\phi_k(\vec{r})\beta(m_s)$, où les deux fonctions α et β sont définies comme suit :

$$\alpha(m_s) \equiv \langle m_s | \alpha \rangle , \quad \beta(m_s) \equiv \langle m_s | \beta \rangle ; \quad (2.79)$$

avec ces définitions, l'état complet d'un électron (spin-orbitale) peut alors se noter :

$$\psi(\vec{r}, m_s; t) = \psi_{+\frac{1}{2}}(\vec{r}; t) \alpha(m_s) + \psi_{-\frac{1}{2}}(\vec{r}; t) \beta(m_s) . \quad (2.80)$$

En effet, si on choisit $m_s = +\frac{1}{2}$ au premier membre, on obtient bien l'égalité puisque, au second membre, seul le premier terme subsiste. Ainsi écrit, $\psi(\vec{r}, m_s; t)$ apparaît manifestement comme une combinaison linéaire d'états fabriqués en formant le produit tensoriel des états orbitaux ("d'espace") par les états de spin. Notons que les deux fonctions α et β sont orthogonales :

$$\sum_{m_s = \pm \frac{1}{2}} \alpha(m_s) \beta(m_s) = 0 . \quad (2.81)$$

On peut donc utiliser une fonction (orbitale) ϕ_k deux fois au plus, l'associant soit avec α , soit avec β . Ainsi, si le système contient un nombre pair d'électrons, tous ceux-ci sont appariés dans l'état fondamental et une bonne fonction pour cet état s'écrira :

$$\Psi_A(1, 2, \dots, N) = C_A Y_A \prod_{i=1}^{N/2} \phi_i(\vec{r}_{2i-1}) \alpha(m_{s\ 2i-1}) \phi_i(\vec{r}_{2i}) \beta(m_{s\ 2i}) . \quad (2.82)$$

Pour un nombre impair d'électrons, s'agissant toujours de l'état fondamental, le dernier état à une particule (celui de plus haute énergie) contiendra un électron non apparié, responsable d'un moment cinétique de spin total pour le système égal à $\frac{1}{2}$. L'état occupé de plus haute énergie s'appelle le niveau de Fermi. Il joue un rôle prééminent dans un métal, comme on le verra ci-dessous.

Ainsi, même s'ils sont sans interaction mutuelle au sens classique, deux fermions d'un même système sont malgré tout corrélés puisqu'ils ne peuvent se trouver dans le même état quantique. Par référence aux états orbitaux spécifiés par les trois nombres quantiques (n, l, m_l) , auxquels il convient d'ajouter le nombre $m_s (= \pm \frac{1}{2})$ pour $S = \frac{1}{2}$ associé au spin, on dit que deux électrons d'un même atome ne peuvent avoir leurs quatre nombres quantiques identiques.

Revenons au cas de deux électrons, dont le spin total peut être 0 ou 1, suivant la composition des moments cinétiques ; avec deux fonctions d'espace ϕ_1 et ϕ_2 linéairement indépendantes, normalisées et orthogonales, on

²⁷notamment si la loi de dispersion est la même.

peut fabriquer un état singulet de spin total $S = 0$ et un état triplet de spin $S = 1$, ce dernier ayant trois composantes distinctes $M_S = 0, \pm 1$; la composante normalisée $M_S = +1$ du triplet est :

$$\Psi_{S=1, M_S=1}(1, 2) = \frac{1}{\sqrt{2}} [\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) - \phi_2(\vec{r}_1)\phi_1(\vec{r}_2)] \alpha(m_{s_1})\alpha(m_{s_2}) , \quad (2.83)$$

et la probabilité de trouver les deux fermions l'un près de \vec{r}_1 , l'autre près de \vec{r}_2 est :

$$dP_{S=1}(\vec{r}_1, \vec{r}_2) = \frac{1}{2} |\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) - \phi_2(\vec{r}_1)\phi_1(\vec{r}_2)|^2 d^3r_1 d^3r_2 . \quad (2.84)$$

En particulier, la probabilité élémentaire est nulle de trouver les deux électrons au même point :

$$dP_{S=1}(\vec{r}_1 = \vec{r}_2) = 0 . \quad (2.85)$$

C'est ce que l'on appelle le "trou de Fermi", que l'on traduit par l'expression : il y a une probabilité nulle de trouver deux électrons au même point avec la même projection de spin ; la présence d'un électron (ou d'un fermion) en un point interdit l'accès de ce point à un autre électron de même spin. Notons que cette propriété reste vraie pour les deux autres composantes $M_S = 0, -1$ du triplet, puisqu'elle ne dépend que de la partie spatiale. On devrait donc plutôt dire : *dans l'état triplet, il y a une probabilité nulle de trouver deux électrons au même point.*

Ce trou de densité n'existe pas pour l'état singulet (où les deux électrons ont forcément des spins contraires), puisqu'alors la fonction d'onde est :

$$\Psi_{S=0, M_S=0}(1, 2) = \frac{1}{\sqrt{2}} [\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) + \phi_2(\vec{r}_1)\phi_1(\vec{r}_2)] \frac{1}{\sqrt{2}} [\alpha(m_{s_1})\beta(m_{s_2}) - \beta(m_{s_1})\alpha(m_{s_2})] . \quad (2.86)$$

et la densité $dP_{S=0}(\vec{r}_1, \vec{r}_2)$ ne s'annule pas en $\vec{r}_1 = \vec{r}_2$. C'est pourquoi on dit que deux électrons "de même spin" se "repoussent" davantage que deux électrons de spin contraire. Il est remarquable que cet antagonisme est indépendant de l'existence ou non d'une interaction ordinaire : il s'exerce en toute circonstance, et ne repose que sur le caractère fermionique. Par ailleurs, comme la répulsion électrostatique entre deux électrons est essentiellement positive et est donc un facteur déstabilisant sur le plan énergétique, on peut s'attendre à ce que, toutes choses égales par ailleurs, les états triplets soient plus bas en énergie que les états singulets correspondants : le principe de Pauli atténue la répulsion classique de Coulomb pour deux électrons formant un état triplet de spin.

Ainsi, l'interaction électrostatique, alliée au spin, assure que pour une configuration donnée, l'état triplet est plus stable que l'état singulet : bien que le Hamiltonien ne contienne pas le spin, l'énergie totale dépend de S ! C'est là une propriété fondamentale permettant de comprendre l'existence du paramagnétisme et plus généralement de mettre le doigt sur les briques élémentaires du magnétisme : les ampériens ne sont pas les fruits de forces magnétiques au sens classique d'une interaction, ils résultent fondamentalement du caractère fermionique des électrons et existent même si ces derniers sont en interaction purement électrostatique, à l'exclusion de toute autre force magnétique.

Les corrélations dues à l'échange et à l'antisymétrisation, traduites par l'expression imagée "trou de Fermi", portent le nom de corrélations dynamiques ; on retiendra qu'elles sont présentes même lorsque toute interaction classique directe est négligée ou absente. En ce sens, et même s'il n'y a pas d'interactions à deux corps dans le Hamiltonien, des particules identiques – qu'il s'agisse d'ailleurs de fermions ou de bosons – sont *forcément* corrélées.

A titre de transition vers le cas général de N particules, il est utile de remarquer que, pour $N = 2$, toute fonction antisymétrique peut s'écrire sous la forme d'un déterminant – qui est par définition une fonction (multilinéaire) antisymétrique. Ainsi, toute fonction de deux fermions construites sur deux spin-orbitales ψ_{k_1} et ψ_{k_2} (orthonormalisées) peut s'écrire :

$$\Psi_A(1, 2) = \frac{1}{\sqrt{2!}} \begin{vmatrix} \psi_{k_1}(1) & \psi_{k_2}(1) \\ \psi_{k_1}(2) & \psi_{k_2}(2) \end{vmatrix} . \quad (2.87)$$

Il est facile de se convaincre que la fonction d'onde antisymétrique (2.78) construite sur un produit de N fonctions²⁸ à une particule $\{\psi_{k_i}\}$ peut de même s'écrire sous la forme d'un déterminant, historiquement appelé déterminant de Slater :

$$\Psi_A(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{k_1}(1) & \psi_{k_2}(1) & \dots & \psi_{k_N}(1) \\ \psi_{k_1}(2) & \psi_{k_2}(2) & \dots & \psi_{k_N}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{k_1}(N) & \psi_{k_2}(N) & \dots & \psi_{k_N}(N) \end{vmatrix} . \quad (2.88)$$

On vérifie aisément que $\frac{1}{\sqrt{N!}}$ est la bonne constante assurant que Ψ_A est normalisée à l'unité quand les spin-orbitales sont elles-mêmes orthonormalisées. En effet, partant de (2.78) et en posant $\Phi = \prod_{i=1}^N \psi_{k_i}$, on a :

$$\langle \Psi_A | \Psi_A \rangle = |C_A|^2 \langle Y_A \Phi | Y_A \Phi \rangle = |C_A|^2 \langle \Phi | Y_A^\dagger Y_A | \Phi \rangle . \quad (2.89)$$

Puisque Y_A est hermitique et idempotent, ceci est égal à :

$$\langle \Psi_A | \Psi_A \rangle = |C_A|^2 \langle \Phi | Y_A \Phi \rangle = \frac{|C_A|^2}{N!} \langle \Phi | \sum_{\lambda=1}^{N!} \varepsilon_\lambda P_\lambda | \Phi \rangle . \quad (2.90)$$

Le produit scalaire contient $N!$ termes, mais seul celui venant de la permutation identité donne une contribution non-nulle, grâce à l'orthogonalité des spin-orbitales. D'où :

$$\langle \Psi_A | \Psi_A \rangle = \frac{|C_A|^2}{N!} \prod_{i=1}^N \langle \psi_{k_i} | \psi_{k_i} \rangle = \frac{|C_A|^2}{N!} = 1 \iff |C_A| = \sqrt{N!} . \quad (2.91)$$

Finalement :

$$|\Psi_A\rangle = \frac{\sqrt{N!}}{N!} \sum_{\lambda=1}^{N!} \varepsilon_\lambda P_\lambda |\Phi\rangle = \frac{1}{\sqrt{N!}} \sum_{\lambda=1}^{N!} (-1)^{\sigma_\lambda} P_\lambda \left| \prod_{i=1}^N \psi_{k_i} \right\rangle , \quad (2.92)$$

qui reproduit bien l'expression (2.88).

Sur la forme (2.88), on voit immédiatement que si deux fonctions ψ_{k_i} coïncident, la fonction d'onde est identiquement nulle : alors, le déterminant a deux colonnes identiques. Enfin, signalons la notation abrégée courante pour l'expression (2.88) :

$$\Psi_A(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \text{Det} [\psi_{k_1} \psi_{k_2} \dots \psi_{k_N}] . \quad (2.93)$$

Remarques

1. Le déterminant ci-dessus peut être construit à partir d'un modèle de particules indépendantes ; alors, chaque fonction ψ_{k_i} est connue d'avance. D'un autre côté, pour des fermions en interaction, on peut se poser la question de la meilleure fonction d'onde du type déterminant, au sens de la méthode variationnelle. Alors, les ψ_{k_i} sont d'avance inconnues et doivent être déterminées. L'écriture du principe variationnel conduit alors à des équations intégro-différentielles pour les fonctions ψ_{k_i} : ce sont les célèbres équations de Hartree-Fock, que l'on ne peut résoudre – hormis quelques cas triviaux sans grand intérêt – qu'avec une machine, par un processus itératif. Cette méthode est typiquement une méthode de champ moyen : en effet, comme fondamentalement la fonction d'onde globale retient une forme *produit* de fonctions à une particule, chaque électron (fermion) se meut en réalité dans le champ *moyen* de tous les autres ; de fait, les équations de Hartree-Fock font apparaître pour les ψ_{k_i} une équation aux valeurs propres contenant un terme potentiel représentant l'effet moyen de tous les autres électrons sur l'un d'entre eux²⁹. En pareil

²⁸ Bien noter que le déterminant est construit avec des spin-orbitales, pas avec des orbitales !

²⁹ De façon *caractéristique* pour un système de fermions, le potentiel de champ moyen fait apparaître *deux* termes ; le premier (terme direct) a une interprétation simple : c'est l'interaction d'un électron avec les autres, distribués suivant une densité donnée par le module carré des spin-orbitales. En revanche, le second (terme d'échange) résulte directement du postulat d'antisymétrisation et ne peut recevoir d'interprétation imagée. L'approximation dite de Hartree consiste à oublier le terme d'échange.

cas, bien sûr, l'énergie des N électrons n'est pas la somme des valeurs propres (énergies) de Hartree-Fock – alors on compterait deux fois l'interaction au sein d'une même paire. La méthode de Hartree-Fock s'emploie pour toute assemblée de fermions (matière stellaire, noyaux, atomes, molécules, gaz de Fermi dans les solides, etc.)

2. Montrons en détail que toutes les fonctions antisymétriques obtenues pour deux électrons peuvent bien s'écrire sous la forme de déterminant(s). En effet, la fonction (2.83) est identique à $\frac{1}{\sqrt{2}} \text{Det}[\phi_1\alpha \phi_2\alpha]$. Maintenant, comme tout déterminant est une fonction multilinéaire (alternée), l'application de l'opérateur $S_- = S_{1-} + S_{2-}$ produit³⁰ :

$$(S_{1-} + S_{2-}) \frac{1}{\sqrt{2}} \text{Det}[\phi_1\alpha \phi_2\beta] = \frac{\hbar}{\sqrt{2}} [\text{Det}[\phi_1\beta \phi_2\alpha] + \text{Det}[\phi_1\alpha \phi_2\beta]] . \quad (2.94)$$

D'autre part $S_- \Psi_{S=1 M_S=1} = \hbar\sqrt{1(1+1) - 1(1-1)} \Psi_{S=1 M_S=0} \equiv \hbar\sqrt{2} \Psi_{S=1 M_S=0}$, de sorte que :

$$\Psi_{S=1 M_S=0} = \frac{1}{2} [\text{Det}[\phi_1\beta \phi_2\alpha] + \text{Det}[\phi_1\alpha \phi_2\beta]] . \quad (2.95)$$

Il est facile de vérifier que ceci est identique à $\frac{1}{2}(\phi_1\phi_2 - \phi_2\phi_1)(\alpha\beta + \beta\alpha)$; la composante $M_S = -1$ du triplet est évidemment $\frac{1}{\sqrt{2}} \text{Det}[\phi_1\beta \phi_2\beta]$. Pour obtenir la seule et unique composante du singulet $\Psi_{S=0 M_S=0}$, il suffit de se souvenir que cette fonction doit être orthogonale à $\Psi_{S=1 M_S=0}$. Deux déterminants différant par au moins une spin-orbitale sont orthogonaux si les spin-orbitales le sont ; la fonction orthogonale à (2.95) est donc :

$$\Psi_{S=0 M_S=0} = \frac{1}{2} [\text{Det}[\phi_1\beta \phi_2\alpha] - \text{Det}[\phi_1\alpha \phi_2\beta]] , \quad (2.96)$$

obtenue par un simple changement de signe dans la combinaison linéaire des deux déterminants. De toute évidence, toutes les fonctions complètes sont bien antisymétriques dans la transposition T_{12} .

2.5 Introduction à la Seconde Quantification

Compte tenu de l'indiscernabilité des particules identiques, la spécification complète de l'état d'un système s'énonce en affirmant que n_1 particules sont dans l'état ψ_1 , n_2 particules sont dans l'état ψ_2 , etc. Souvent, les états à une particule ψ_k sont ceux définis en l'absence d'interaction, mais ceci n'est nullement obligatoire. k généralement désigne l'ensemble des nombres quantiques associés à un ECOC.

En pratique, on est souvent contraint de considérer des situations où le nombre de particules n'est pas fixé une fois pour toutes. Par exemple, dans le cas du champ électromagnétique couplé à la matière, le nombre de photons n'est pas donné et constant puisque la matière peut absorber ou émettre des photons. Il en va de même pour les phonons, particules (*quanta*) décrivant la quantification des vibrations harmoniques d'un solide : le nombre (moyen) de phonons augmente avec la température puisque l'élévation de celle-ci augmente l'amplitude des vibrations du réseau. En Théorie Quantique des Champs, la non-conservation du nombre de particules au cours du temps est la règle. Une paire électron-positron peut s'annihiler sous forme de photons, étant entendu d'ailleurs que le processus inverse est tout autant possible.

Ainsi naît la nécessité de reformuler la Mécanique Quantique dans un cadre plus large où le nombre de particules est variable. De la sorte, un système quantique évolue dans un espace d'états "infiniment" plus vaste, appelé espace de Fock. On peut se le représenter comme un empilement de feuilles (secteurs) : dans chaque feuille, le nombre de particules est constant ; on passe d'une feuille à l'autre par l'action d'opérateurs de création ou d'annihilation de particules, qui généralisent ceux que l'on a rencontrés à propos de l'oscillateur harmonique, et dont il existe évidemment deux familles, l'une associée aux bosons, l'autre associée aux fermions. On verra que la distinction se tient fondamentalement dans les relations de commutation : les opérateurs de bosons satisfont des règles de commutation (comme dans le cas de l'oscillateur harmonique), les opérateurs de

³⁰étant entendu que $S_- \alpha = \hbar\sqrt{\frac{1}{2}(\frac{1}{2} + 1) - \frac{1}{2}(\frac{1}{2} - 1)}\beta \equiv \hbar\beta$.

fermions obéissent à des relations d'anticommutation. Cette distinction contient à elle seule la différence de symétrie vis-à-vis du groupe de permutations, en conformité avec le postulat de symétrisation.

Un état quelconque de l'espace de Fock est noté :

$$|n_1, n_2, \dots, n_k, \dots\rangle \quad (n_j \in \mathbb{N}) \quad (2.97)$$

et désigne un état produit tensoriel où il y a n_1 particules chacune dans l'état $|1\rangle$, n_2 particules chacune dans l'état $|2\rangle$, \dots ; les n_j sont les nombres d'occupation. On introduit alors des opérateurs de création a_k^\dagger et d'annihilation a_k tels que :

$$a_k^\dagger |n_1, n_2, \dots, n_k, \dots\rangle \propto |n_1, n_2, \dots, n_{k-1}, n_k + 1, n_{k+1}, \dots\rangle, \quad (2.98)$$

$$a_k |n_1, n_2, \dots, n_k, \dots\rangle \propto |n_1, n_2, \dots, n_{k-1}, n_k - 1, n_{k+1}, \dots\rangle. \quad (2.99)$$

Ce sont ces opérateurs qui permettent de passer d'un secteur à l'autre de l'espace de Fock, le nombre total de particules variant au plus d'une unité.

Ceci étant, il est possible de distinguer notamment :

- l'état vide :

$$|\text{vide}\rangle \equiv |n_1 = 0, n_2 = 0, \dots, n_k = 0, \dots\rangle. \quad (2.100)$$

Pour l'état vide, on a :

$$a_k |\text{vide}\rangle = 0; \quad (2.101)$$

cette relation exprime le fait que l'on ne peut détruire une particule qui n'existe pas

- les états à une particule :

$$|\Psi_k\rangle \equiv |n_1 = 0, n_2 = 0, \dots, n_{k-1} = 0, n_k = 1, n_{k+1} = 0, \dots\rangle, \quad (2.102)$$

que l'on peut construire en faisant agir un opérateur de création sur l'état vide, conformément à (2.98) :

$$|\Psi_k\rangle = a_k^\dagger |\text{vide}\rangle. \quad (2.103)$$

Par ailleurs, toujours grâce à l'argument suivant lequel il est impossible de détruire une particule qui n'existe pas, on a :

$$a_k |0, 0, \dots, n_{k'}, 0, \dots\rangle = 0 \quad \forall k \neq k', \quad (2.104)$$

- les états à deux particules, $|n_k, n_l\rangle$. Ils sont engendrés à partir du vide par action d'un produit tel que $a_k^\dagger a_l^\dagger$:

$$|\Psi_{kl}\rangle = a_k^\dagger a_l^\dagger |\text{vide}\rangle. \quad (2.105)$$

Si les particules sont des bosons, l'état doit être symétrique ; l'égalité suivante doit donc être vraie :

$$a_k^\dagger a_l^\dagger |\text{vide}\rangle = a_l^\dagger a_k^\dagger |\text{vide}\rangle \quad \forall k, l. \quad (2.106)$$

Elle doit l'être d'ailleurs quel que soit l'état sur lequel agit le produit $a_k^\dagger a_l^\dagger$. On en déduit, pour des bosons :

$$a_k^\dagger a_l^\dagger = a_l^\dagger a_k^\dagger \quad \forall k, l \iff [a_k^\dagger, a_l^\dagger] = 0 \quad (\text{bosons}). \quad (2.107)$$

La relation hermitique conjuguée est aussi vraie :

$$a_k a_l = a_l a_k \quad \forall k, l \iff [a_k, a_l] = 0 \quad (\text{bosons}). \quad (2.108)$$

Au contraire, s'il s'agit de fermions, l'état doit être antisymétrique dans toute transposition ; il faut que :

$$a_k^\dagger a_l^\dagger |\text{vide}\rangle = -a_l^\dagger a_k^\dagger |\text{vide}\rangle \quad \forall k, l \quad (2.109)$$

et, plus généralement³¹ :

$$a_k^\dagger a_l^\dagger = -a_l^\dagger a_k^\dagger \quad \forall k, l \iff \{a_k^\dagger, a_l^\dagger\} = 0 \quad (\text{fermions}), \quad (2.110)$$

qui entraîne à son tour :

$$a_k a_l = -a_l a_k \quad \forall k, l \iff \{a_k, a_l\} = 0 \quad (\text{fermions}). \quad (2.111)$$

³¹Les accolades désignent l'anticommutateur : $\{A, B\} \stackrel{\text{d\'ef}}{=} AB + BA$.

Les produits $a_k^\dagger a_l^\dagger a_m^\dagger$ agissant sur $|\text{vide}\rangle$ donnent des états à trois particules, et ainsi de suite : le produit d'un nombre quelconque d'opérateurs de création agissant sur le vide permet de construire les états à nombre quelconque de particules. L'ensemble de ces états constitue par définition l'espace de Fock.

Il convient de déterminer les constantes de proportionnalité dans (2.98) et (2.99), relations que l'on peut écrire plus précisément :

$$a_k^\dagger |\dots, n_k, \dots\rangle = C_{n_k+1} |\dots, n_k+1, \dots\rangle, \quad (2.112)$$

$$a_k |\dots, n_k, \dots\rangle = C_{n_k}^* |\dots, n_k-1, \dots\rangle. \quad (2.113)$$

Tous les nombres $n_{k'}$, $k \neq k'$ sont sous-entendus, et inchangés d'un membre à l'autre ; les deux constantes ne dépendent évidemment que du nombre d'occupation du mode que les opérateurs modifient (deux modes distincts ne se voient pas). De ces deux relations on déduit :

$$a_k^\dagger a_k |\dots, n_k, \dots\rangle = C_{n_k}^* a_k^\dagger |\dots, n_k-1, \dots\rangle = |C_{n_k}|^2 |\dots, n_k, \dots\rangle. \quad (2.114)$$

Ainsi, les états de Fock sont propres du produit $N_k \equiv a_k^\dagger a_k$; l'opérateur N_k est appelé nombre d'occupation (de l'état k), il est hermitique et ses valeurs propres $n_k = |C_{n_k}|^2$ donnent le nombre de particules effectivement placées dans l'état k . Les relations (2.112), (2.113) et (2.114) sont vraies qu'il s'agisse de fermions ou de bosons mais, pour la clarté, traitons maintenant séparément chaque famille de particules.

- Pour les bosons, chaque état k peut accepter un nombre quelconque de particules ; la relation $n_k = |C_{n_k}|^2$ permet alors de fixer (à une phase près conventionnelle) la constante C_{n_k} . En décidant conventionnellement de prendre cette dernière réelle positive, il vient :

$$C_{n_k} = \sqrt{n_k}, \quad (2.115)$$

d'où il résulte, pour les bosons ($n_k \in \mathbb{N}$) (voir (2.112) et (2.113)) :

$$a_k^\dagger |\dots, n_k, \dots\rangle = \sqrt{1+n_k} |\dots, n_k+1, \dots\rangle \quad (2.116)$$

$$a_k |\dots, n_k, \dots\rangle = \sqrt{n_k} |\dots, n_k-1, \dots\rangle. \quad (2.117)$$

La combinaison de ces deux équations redonne bien (2.114), compte tenu de (2.115) :

$$a_k^\dagger a_k |\dots, n_k, \dots\rangle = \sqrt{n_k} a_k^\dagger |\dots, n_k-1, \dots\rangle = \sqrt{n_k} \sqrt{1+(n_k-1)} |\dots, (n_k-1)+1, \dots\rangle, \quad (2.118)$$

de sorte que :

$$a_k^\dagger a_k |\dots, n_k, \dots\rangle = n_k |\dots, n_k, \dots\rangle, \quad (2.119)$$

comme il se doit.

- Pour les fermions, il reste vrai que $|C_{n_k}|^2$ est une fonction linéaire de n_k , en accord avec (2.114), mais maintenant n_k ne peut prendre que les deux valeurs 0 ou 1, en conformité avec le principe de Pauli. Un peu de réflexion montre que l'on a ($n_k = 0, 1$) :

$$a_k^\dagger |\dots, n_k, \dots\rangle = \sqrt{1-n_k} |\dots, n_k+1, \dots\rangle \quad (2.120)$$

$$a_k |\dots, n_k, \dots\rangle = \sqrt{n_k} |\dots, n_k-1, \dots\rangle. \quad (2.121)$$

Il est possible de rassembler ces équations sous la forme :

$$a_k^\dagger |\dots, n_k, \dots\rangle = \sqrt{1+\varepsilon n_k} |\dots, n_k+1, \dots\rangle \quad (2.122)$$

$$a_k |\dots, n_k, \dots\rangle = \sqrt{n_k} |\dots, n_k-1, \dots\rangle \quad (2.123)$$

avec :

$$\varepsilon = \begin{cases} +1 & \text{bosons} \\ -1 & \text{fermions} \end{cases} \quad (2.124)$$

Pour finir, ces équations permettent d'obtenir les relations de commutation (ou d'anticommutation) entre un opérateur de création a_k^\dagger et un opérateur d'annihilation a_l . On trouve immédiatement :

$$[a_k, a_l^\dagger] = \delta_{kl} \quad (\text{bosons}) , \quad (2.125)$$

$$\{a_k, a_l^\dagger\} = \delta_{kl} \quad (\text{fermions}) , \quad (2.126)$$

soit :

$$a_k a_l^\dagger - \varepsilon a_l^\dagger a_k = \delta_{kl} . \quad (2.127)$$

Dans tous les cas, $N_k = a_k^\dagger a_k$ est l'opérateur nombre de particules ; ses valeurs propres sont des entiers non-négatifs :

$$N_k = \begin{cases} 0, 1, 2, \dots & (\text{bosons}) \\ 0, 1 & (\text{fermions}) \end{cases} . \quad (2.128)$$

L'usage courant désigne par la lettre b les opérateurs de bosons, par c ceux des fermions. Avec cette convention, et à titre de récapitulation, les relations fondamentales s'écrivent :

- Fermions :

$$\{c_k, c_l^\dagger\} = \delta_{kl} , \quad \{c_k, c_l\} = \{c_k^\dagger, c_l^\dagger\} = 0 \quad \forall k, l \quad (2.129)$$

$$c_k | \dots, n_k, \dots \rangle = \sqrt{n_k} | \dots, n_k - 1, \dots \rangle , \quad c_k^\dagger | \dots, n_k, \dots \rangle = \sqrt{1 - n_k} | \dots, n_k + 1, \dots \rangle \quad (2.130)$$

- Bosons :

$$[b_k, b_l^\dagger] = \delta_{kl} , \quad [b_k, b_l] = [b_k^\dagger, b_l^\dagger] = 0 \quad \forall k, l \quad (2.131)$$

$$b_k | \dots, n_k, \dots \rangle = \sqrt{n_k} | \dots, n_k - 1, \dots \rangle , \quad b_k^\dagger | \dots, n_k, \dots \rangle = \sqrt{1 + n_k} | \dots, n_k + 1, \dots \rangle . \quad (2.132)$$

Le point essentiel à retenir est que la symétrie de permutation est *incluse* dans l'algèbre des opérateurs de création et d'annihilation. Par exemple, un état tel que $|\Psi_{kl}\rangle = a_k^\dagger a_l^\dagger |\text{vac}\rangle$ est un état à deux particules ayant la bonne symétrie. S'il s'agit de bosons, $b_k^\dagger b_l^\dagger |\text{vac}\rangle = +b_l^\dagger b_k^\dagger |\text{vac}\rangle = +|\Psi_{lk}\rangle$; au contraire, pour des fermions $|\Psi_{kl}\rangle = c_k^\dagger c_l^\dagger |\text{vac}\rangle = -c_l^\dagger c_k^\dagger |\text{vac}\rangle = -|\Psi_{lk}\rangle$.

Les opérateurs représentant les observables s'expriment très simplement, et de façon très transparente, en Seconde Quantification. Dans la suite, $|\Psi_k\rangle$ désigne un état propre pour une particule, de Hamiltonien H_1 :

$$H_1 |\Psi_k\rangle = \varepsilon_k |\Psi_k\rangle \quad \Longleftrightarrow \quad H_1 = \sum_k |\Psi_k\rangle \varepsilon_k \langle \Psi_k| . \quad (2.133)$$

On a, par définition :

$$|\Psi_k\rangle = a_k^\dagger |\text{vac}\rangle . \quad (2.134)$$

Considérons maintenant un ensemble de telles particules sans interaction, dont le Hamiltonien est la grandeur additive $H = \sum_i H_1(i)$; quelle est l'énergie moyenne E du système dans un état où il y a n_k particules dans l'état $|\Psi_k\rangle$, $k = 1, 2, \dots$? Visiblement, on a :

$$E = \sum_k n_k \varepsilon_k . \quad (2.135)$$

Le Hamiltonien peut donc s'écrire :

$$H = \sum_k \varepsilon_k a_k^\dagger a_k \equiv \sum_k \varepsilon_k N_k . \quad (2.136)$$

Bien entendu, il est possible de considérer une autre base d'états à une particule (toujours orthonormalisée), $\{|\Phi_\kappa\rangle\}$, sur laquelle H_1 n'est pas diagonal. Il existe une transformation unitaire U permettant d'écrire :

$$|\Psi_k\rangle = \sum_\kappa U_{\kappa k} |\Phi_\kappa\rangle , \quad U_{\kappa k} = \langle \Phi_\kappa | \Psi_k \rangle . \quad (2.137)$$

En considérant (2.134), on voit que la même transformation existe entre les opérateurs associés aux deux bases ; avec des notations évidentes :

$$a_k^\dagger = \sum_{\kappa} U_{\kappa k} A_{\kappa}^\dagger . \quad (2.138)$$

Ceci étant, H (2.136) se transforme comme suit :

$$H = \sum_{\kappa \lambda} \sum_k \varepsilon_k U_{\kappa k} U_{\lambda k}^* A_{\kappa}^\dagger A_{\lambda} ; \quad (2.139)$$

la somme sur k est :

$$\sum_k \varepsilon_k \langle \Phi_{\kappa} | \Psi_k \rangle \langle \Phi_{\lambda} | \Psi_k \rangle^* = \sum_k \langle \Phi_{\kappa} | \Psi_k \rangle \varepsilon_k \langle \Psi_k | \Phi_{\lambda} \rangle \equiv \langle \Phi_{\kappa} | H_1 | \Phi_{\lambda} \rangle , \quad (2.140)$$

de sorte que H s'écrit :

$$H = \sum_{\kappa \lambda} \langle \Phi_{\kappa} | H_1 | \Phi_{\lambda} \rangle A_{\kappa}^\dagger A_{\lambda} ; \quad (2.141)$$

D'où la forme générale d'une observable additive à un corps $\Omega = \sum_i \omega(i)$, exprimée en fonction des opérateurs de création et d'annihilation relatifs à des modes quelconques :

$$\Omega = \sum_{kl} \langle \Psi_k | \omega | \Psi_l \rangle a_k^\dagger a_l \equiv \sum_{kl} \omega_{kl} a_k^\dagger a_l . \quad (2.142)$$

Sur ses modes propres, Ω est diagonal et la somme ne contient alors que les termes $k = l$, ce qui fait réapparaître les nombres d'occupation (de ces modes propres), tout comme pour H en (2.136).

À titre d'illustration, montrons comment fonctionne le formalisme. Partons de la quantité $\langle \Psi_k | \Omega | \Psi_l \rangle$, élément de matrice d'une observable additive à un corps entre deux états à une particule engendrés à partir du vide selon (2.103). On peut de ce fait écrire, tenant compte de (2.142) :

$$\langle \Psi_k | \Omega | \Psi_l \rangle = \langle \text{vac} | a_k \Omega a_l^\dagger | \text{vac} \rangle = \sum_{k' l'} \langle \text{vac} | a_k a_{k'}^\dagger a_{l'} a_l^\dagger | \text{vac} \rangle \omega_{k' l'} . \quad (2.143)$$

L'idée consiste à jouer avec les relations de commutation³² afin de mettre les opérateurs d'annihilation à droite dans le produit, juste devant $|\text{vac}\rangle$: en vertu de (2.101), le terme correspondant sera nul. Le produit des quatre opérateurs dans (2.143) est aussi $a_k a_{k'}^\dagger (\delta_{ll'} + \varepsilon a_l^\dagger a_{l'})$ (voir (2.127)) ; comme $a_{l'} |\text{vac}\rangle = 0$, il reste :

$$\langle \Psi_k | \Omega | \Psi_l \rangle = \sum_{k' l'} \langle \text{vac} | a_k a_{k'}^\dagger \delta_{ll'} | \text{vac} \rangle \omega_{k' l'} ; \quad (2.144)$$

en recommençant la même opération avec les opérateurs restants, il vient :

$$\langle \Psi_k | \Omega | \Psi_l \rangle = \sum_{k' l'} \langle \text{vac} | (\delta_{kk'} + \varepsilon a_k^\dagger a_k) | \text{vac} \rangle \delta_{ll'} \omega_{k' l'} = \sum_{k' l'} \delta_{kk'} \delta_{ll'} \omega_{k' l'} = \omega_{kl} ; \quad (2.145)$$

Quant au premier membre, c'est très précisément :

$$\langle 0, 0, \dots, 0, 1_k, 0, \dots | \sum_i \omega(i) | 0, 0, \dots, 0, 1_l, 0, \dots \rangle ; \quad (2.146)$$

tant que l'indice i porte sur une particule qui n'existe pas (parce que le nombre d'occupation est nul dans le *bra* et dans le *ket*), le résultat est nul ; le seul terme non-nul est bien l'élément de matrice de ω avec une particule dans $|\Psi_k\rangle$ à gauche et dans $|\Psi_l\rangle$ à droite : c'est bien ω_{kl} , en conformité avec (2.145).

Considérons maintenant une observable à deux corps :

$$V = \frac{1}{2} \sum_{i \neq j} v(i, j) . \quad (2.147)$$

³²Le jeu qui consiste à faire commuter les opérateurs d'un produit pour réduire de deux unités le nombre de facteurs s'appelle une *contraction*.

V est par exemple l'interaction entre particules, $V = \frac{1}{2} \sum_{i \neq j} v(\vec{r}_i - \vec{r}_j)$. En généralisation immédiate de (2.142), on écrit³³ :

$$V = \frac{1}{2} \sum_{k, l, k', l'} v_{k k' l l'} a_{k'}^\dagger a_k^\dagger a_l a_{l'} \quad (2.148)$$

et toute la question est d'identifier la quantité $v_{k k' l l'}$. Pour cela, on fait jouer les contractions comme ci-dessus, en partant de la quantité $\langle \Psi_m \Psi_n | V | \Psi_{m'} \Psi_{n'} \rangle$. Un calcul un peu long montre que $v_{k k' l l'}$ n'est autre qu'un élément de matrice à deux corps, donné par l'intégrale :

$$v_{k k' l l'} = \sum_{\text{spins}} \int d^3 r d^3 r' \Psi_k^*(\vec{r}) \Psi_{k'}^*(\vec{r}') v(\vec{r} - \vec{r}') \Psi_l(\vec{r}) \Psi_{l'}(\vec{r}') \equiv \langle \Psi_k \Psi_{k'} | v | \Psi_l \Psi_{l'} \rangle \quad (2.149)$$

Au total, un Hamiltonien ne contenant que des termes à un et à deux corps :

$$H = \sum_{i=1}^N H_1(i) + \frac{1}{2} \sum_{i=1}^N \sum_{j=1, j \neq i}^N H_2(i, j) \quad (2.150)$$

s'écrit comme suit :

$$H = \sum_{kl} \varepsilon_{kl} a_k^\dagger a_l + \frac{1}{2} \sum_{k, l, k', l'} v_{k k' l l'} a_{k'}^\dagger a_k^\dagger a_l a_{l'} \quad , \quad (2.151)$$

où :

$$\varepsilon_{kl} = \langle \Psi_k | H_1 | \Psi_l \rangle \quad , \quad v_{k k' l l'} = \langle \Psi_k \Psi_{k'} | H_2 | \Psi_l \Psi_{l'} \rangle \quad . \quad (2.152)$$

Le point remarquable est que l'expression (2.151) est indépendante du nombre (fixe ou variable) de particules, N .

³³Attention à l'ordre des indices !

Chapitre 3

Atomes à plusieurs électrons

Pour un atome de charge nucléaire Z et possédant N électrons, le Hamiltonien électrostatique est, avec des notations évidentes (et en supposant le noyau infiniment massif) :

$$H_e = \sum_{i=1}^N \left[\frac{\vec{p}_i^2}{2m} - \frac{Ze'^2}{r_i} \right] + \frac{1}{2} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \frac{e'^2}{r_{ij}} \equiv \sum_{i=1}^N H_1(i) + V . \quad (3.1)$$

Le dernier terme représente la répulsion Coulombienne des électrons entre eux et constitue un terme essentiellement déstabilisant (le facteur $\frac{1}{2}$ est là pour ne pas compter deux fois l'interaction au sein d'une même paire), dont on voit qu'il est de l'ordre de plusieurs eV et est loin d'être "petit". D'ailleurs, même s'il l'était, il faut réaliser qu'il y a $N(N-1) \sim N^2$ termes de cette sorte ; ceux-ci, tous positifs, sont donc beaucoup plus nombreux que les $N \sim Z$ termes représentant l'attraction du noyau sur les électrons. Dans tous les cas, le terme à deux corps joue donc visiblement un rôle non-négligeable.

À cause de ce terme à deux corps, on ne sait pas résoudre exactement l'équation aux valeurs propres et il faut recourir à des descriptions approchées – que l'on ne peut d'ailleurs mettre réellement en œuvre dans les cas intéressants qu'en utilisant l'outil numérique. L'atome à N électrons est un premier exemple de système relevant des techniques du problème à N corps (*N-body*, ou *many-body problem*), omniprésent en Physique : très souvent, c'est quand on "branche" les interactions que des propriétés remarquables surgissent¹.

3.1 Modèle à électrons indépendants

Ce modèle est défini en effectuant l'approximation la plus triviale, consistant à oublier purement et simplement la répulsion V dans (3.1). Ce procédé est visiblement brutal et conduira à des résultats très médiocres ; cependant, le modèle n'est pas sans intérêt dans la mesure où il permet d'introduire des idées que l'on peut généraliser dans des contextes plus subtils et aussi parce que, finalement, il constitue le fondement le plus simple du principe de la classification périodique, au moins dans ses grandes lignes. Enfin, sa discussion permet de revenir sur le fait fondamental que deux particules (identiques) indépendantes au sens classique sont malgré tout corrélées en raison de leur indiscernabilité ; une illustration fameuse de cette "interaction" quantique est le trou de Fermi, qui sera brièvement rediscuté ci-dessous.

Dans ces conditions, le Hamiltonien à considérer est :

$$H_{\text{ind}}(\rho_1, \rho_2, \dots, \rho_N) = \sum_{i=1}^N H_1(\rho_i) . \quad (3.2)$$

¹La théorie BCS en est un bel exemple.

H_{ind} est la somme de Hamiltoniens à une particule, H_{ρ_i} , tous identiques et ne différant dans la sommation que par l'indice i des variables dont ils dépendent et sur lesquelles ils agissent. Désignons maintenant par ψ_k et E_k les modes propres de H_1 :

$$H_1 \psi_k(\rho) = E_k \psi_k(\rho) \quad (3.3)$$

où ρ dénote collectivement les coordonnées d'espace et la variable de spin. Si H_1 ne dépend pas du spin, la fonction ψ_k peut toujours être prise comme le produit d'une fonction ϕ_k par une fonction de spin χ , ϕ_k étant propre de H_1 :

$$\psi_k(\rho) \equiv \phi_k(\vec{r}) \chi(m_s) , \quad H_1 \phi_k(\vec{r}) = E_k \phi_k(\vec{r}) . \quad (3.4)$$

Suivant la terminologie usuelle, la fonction $\psi_k(\rho)$ est appelée spin-orbitale, la fonction $\phi_k(\vec{r})$ est appelée orbitale. Dans la suite, quand c'est possible, on continue à raisonner en toute généralité avec les spin-orbitales.

H_{ind} est une somme d'opérateurs H_1 à un corps : comme on l'a vu précédemment, tout produit de spin-orbitales ψ_k sera fonction propre de H_{ind} avec pour valeur propre la somme des valeurs propres associées aux spin-orbitales :

$$\begin{aligned} H \prod_{j=1}^N \psi_{k_j}(\rho_j) &= \sum_{i=1}^N \prod_{j=1}^N H_1(\rho_i) \psi_{k_j}(\rho_j) = \sum_{i=1}^N \left[\prod_{j=1, j \neq i}^N \psi_{k_j}(\rho_j) \right] H_1(\rho_i) \psi_{k_i}(\rho_i) \\ &= \sum_{i=1}^N \left[\prod_{j=1, j \neq i}^N \psi_{k_j}(\rho_j) \right] E_{k_i} \psi_{k_i}(\rho_i) = \sum_{i=1}^N E_{k_i} \prod_{j=1}^N \psi_{k_j}(\rho_j) \end{aligned} \quad (3.5)$$

On a donc :

$$H_{\text{ind}} \prod_{j=1}^N \psi_{k_j}(\rho_j) = E_{k_1, k_2, \dots, k_N} \prod_{j=1}^N \psi_{k_j}(\rho_j) , \quad E_{k_1, k_2, \dots, k_N} = \sum_{j=1}^N E_{k_j} . \quad (3.6)$$

Ceci étant, le produit $\prod_{j=1}^N \psi_{k_j}$ n'est pas une fonction physiquement acceptable vis-à-vis du postulat de symétrisation : les électrons sont des fermions et la transposition des variables de deux quelconques d'entre eux doit donner un changement de signe dans la fonction d'onde. Une façon commode d'engendrer une bonne fonction consiste à appliquer au produit de spin-orbitales l'opérateur d'antisymétrisation, Y_A introduit antérieurement :

$$Y_A = \frac{1}{N!} \sum_{\lambda=1}^{N!} (-1)^{\sigma_\lambda} P_\lambda . \quad (3.7)$$

Dans cette équation, P_λ désigne l'une quelconque des $N!$ permutations du groupe symétrique \mathcal{S}_N et σ_λ est le nombre de transpositions dont le produit reconstitue P_λ ; Y_A est hermitique et idempotent :

$$Y_A^2 = Y_A \quad Y_A^\dagger = Y_A , \quad (3.8)$$

et pour toute fonction $\Phi(\rho_1, \rho_2, \dots, \rho_N)$:

$$P_\mu Y_A \Phi = (-1)^{\sigma_\mu} Y_A \Phi , \quad (3.9)$$

en conformité avec le postulat de symétrisation. Si on prend maintenant pour Φ le produit de spin-orbitales ci-dessus, on obtient une fonction acceptable : elle est fonction propre de H_{ind} et satisfait les exigences liées à l'indiscernabilité des particules identiques. En définitive, toute fonction propre du modèle à électrons indépendants est de la forme :

$$\Psi(\rho_1, \rho_2, \dots, \rho_N) = C Y_A \prod_{j=1}^N \psi_{k_j}(\rho_j) , \quad (3.10)$$

où C est une constante de normalisation.

En particulier, pour fabriquer l'état fondamental d'un système de fermions sans interactions, il faut donc prendre successivement tous les états propres de H_1 et placer dans chacun d'entre eux un et un seul

électron, compte tenu du spin. En notant $\psi_1, \psi_2, \dots, \psi_n, \dots$ les états classés par ordre d'énergies croissantes $E_1 \leq E_2 \leq \dots \leq E_n, \dots$, l'état fondamental se construit à partir du produit $\psi_1 \psi_2 \dots \psi_N$. Toutes choses égales par ailleurs², l'énergie totale d'un système de N fermions est donc considérablement plus grande que celle d'un système de N bosons puisque l'on est dans l'obligation d'aller chercher des états excités de H_1 , chaque état propre de H_1 ne pouvant recevoir qu'un seul fermion. En tout cas, le fondamental s'écrit comme suit à l'aide des spin-orbitales :

$$\Psi_A(\rho_1, \rho_2, \dots, \rho_N) = C Y_A [\psi_1(\rho_1) \psi_2(\rho_2) \dots \psi_N(\rho_N)] . \quad (3.11)$$

Comme déjà mentionné, il arrive souvent que les états propres de H_1 ne dépendent pas de la variable de spin, m_s . En pareil cas, chaque ψ_k est le produit d'une orbitale $\phi_k(\vec{r})$ et d'une fonction de spin χ . Dans le cas d'électrons, $S = 1/2$, la spin-orbitale est soit $\phi_k(\vec{r})\alpha(m_s)$, soit $\phi_k(\vec{r})\beta(m_s)$, où les deux fonctions α et β sont définies comme suit :

$$\alpha(m_s) = \begin{cases} 1 & \text{si } m_s = +1/2 \\ 0 & \text{si } m_s = -1/2 \end{cases} , \quad \beta(m_s) = \begin{cases} 0 & \text{si } m_s = +1/2 \\ 1 & \text{si } m_s = -1/2 \end{cases} . \quad (3.12)$$

On peut donc utiliser une fonction ϕ_k deux fois au plus, l'associant soit avec α , soit avec β . Ainsi, si l'atome contient un nombre pair d'électrons, tous ceux-ci sont appariés dans l'état fondamental et une bonne fonction pour cet état s'écrira :

$$\Psi_A(\rho_1, \rho_2, \dots, \rho_N) = C Y_A \prod_{i=1}^{N/2} \phi_i(\vec{r}_{2i-1})\alpha(m_{s\ 2i-1})\phi_i(\vec{r}_{2i})\beta(m_{s\ 2i}) . \quad (3.13)$$

Pour un nombre impair d'électrons, s'agissant toujours de l'état fondamental, le dernier état à une particule (celui de plus haute énergie) contiendra un électron non apparié, responsable d'un moment cinétique de spin total pour le système égal à $1/2$.

Le modèle à électrons indépendants conduit naturellement au modèle en couches de l'atome. Chaque niveau hydrogénoïde E_n est dégénéré³ n^2 fois ; compte tenu du spin, on peut donc mettre $2n^2$ électrons dans l'ensemble $\{\psi_{nlmm_s}\}_{lmm_s}$. La couche K ($n = 1$) peut donc contenir 2 électrons, et contient une seule sous-couche, notée 1s ($l = 0$). La couche L ($n = 2$) contient les deux sous-couches 2s et 2p ($l = 1$) qui peuvent accommoder respectivement 2 et 6 électrons. La couche M regroupe les trois sous-couches s, p et d ($l = 2$) et peut accueillir $2 + 6 + 10 = 18$ électrons, et ainsi de suite. Au moins au début de la classification, l'oubli de la répulsion électrostatique n'est pas catastrophique et l'ordonnement des énergies hydrogénoïdes suffit à comprendre qualitativement la constitution de l'état fondamental des atomes pas trop lourds. Ainsi, l'état fondamental du carbone (6 électrons), est caractérisé par la configuration électronique⁴ $(1s)^2(2s)^2(2p)^2$. D'un autre côté, en l'absence de toute interaction – s'en tenant donc strictement aux niveaux hydrogénoïdes – on ne comprend pas pourquoi un atome aussi simple que le lithium ($Z = 3$) a pour configuration fondamentale $(1s)^2(2s)$ et pourquoi pas $(1s)^2(2p)$: ces deux configurations ont, de ce point de vue, strictement la même énergie ; le même argument vaut pour le bore ($Z = 4$), dont la configuration fondamentale est $(1s)^2(2s)^2$.

3.2 Termes spectraux. Multiplets

Que l'on prenne ou non en compte la répulsion entre les électrons, il est possible de faire une présentation générale reposant exclusivement sur la symétrie du Hamiltonien. Comme le système est invariant par rotation, le moment cinétique total⁵ est une constante du mouvement. De surcroît, dans l'approximation électrostatique, H commute évidemment avec tous les opérateurs de spin et en particulier avec le spin total $\vec{S} = \sum_{i=1}^N \vec{s}_i$. En définitive, il existe des vecteurs propres communs à $H, \vec{L}^2, L_z, \vec{S}^2$ et S_z .

²notamment si la loi de dispersion est la même.

³Se souvenir de la dégénérescence "accidentelle" du champ Coulombien nu.

⁴Cette configuration ne permet pas de comprendre la valence 4 du carbone ; on rend compte de celle-ci en invoquant la notion d'hybridation, consistant à prendre une orbitale 2s et à la combiner avec les 3 orbitales p (hybridation notée $(sp)^3$ par les chimistes.

⁵ H ne commute évidemment pas avec la somme partielle des moments cinétiques d'une partie des électrons : il n'y a invariance par rotation que si *tous* les électrons tournent en bloc.

Le moment cinétique orbital total s'obtient en combinant les moments cinétiques \vec{l}_i des orbitales, représentés par les nombres quantiques l_i et m_{l_i} , $1 \leq i \leq N$. La discussion générale est sans grand intérêt, mais on peut noter qu'il est facile de savoir quelles valeurs de L sont associées à une configuration donnée $\dots (n_i, l_i)^{p_i} \dots$, sachant que $M_L = \sum_{i=1}^N m_{l_i}$. Notons d'abord que la somme (partielle) des m_{l_i} relatifs à une couche complète est nulle, donc on ne peut trouver d'états $L \neq 0$ qui exigeraient $M_L \neq 0$; il suffit donc de regarder les couches incomplètement remplies et de trouver les valeurs possibles de la somme des m_{l_i} associés, sans oublier le principe de Pauli. Ayant trouvé le M_L maximum, $M_{L, \max}$, on peut affirmer que les valeurs de L – moment cinétique orbital – sont $M_{L, \max}, M_{L, \max} - 1, \dots, 0$. Il n'y a évidemment qu'une seule façon de combiner les m_{l_i} pour obtenir $M_{L, \max}$ (chacun d'entre eux prend sa valeur maximale l_i ; en revanche, il y a plusieurs façons de construire la somme égale à $M_{L, \max} - 1$. L'une de ces combinaisons donnera la composante $M_L - 1$ associée à $L_{\max} \equiv M_{L, \max}$, les autres – orthogonales – donneront les composantes maximum pour $L_{\max} - 1$.

Le même petit jeu peut, sans difficulté, être précisé à propos du spin total $\vec{S} = \sum_{i=1}^N \vec{s}_i$, puisque chaque s_i prend exclusivement la valeur $\frac{1}{2}$. Ainsi, pour N électrons, la valeur maximale de S est :

$$S_{\max} = \frac{N}{2} . \quad (3.14)$$

La composante $M_S = S$ ("ferromagnétique") associée à S_{\max} s'obtient en accolant chaque orbitale une fois occupée avec la fonction de spin α . La fonction d'onde correspondante se forme ainsi à partir de :

$$\Phi_{S_{\max} M_S=S_{\max}} = \left[\prod_{i=1}^N \phi_{n_i l_i m_{l_i}}(\vec{r}_i) \right] \prod_{j=1}^N \alpha(m_{s_j}) . \quad (3.15)$$

La bonne fonction antisymétrique correspondante est alors :

$$\Psi_{S_{\max} M_S=S_{\max}} = Y_A [\Phi_{S_{\max} M_S=S_{\max}}] ; \quad (3.16)$$

comme la partie de spin de l'expression (3.15) est totalement symétrique, il vient simplement :

$$\Psi_{S_{\max} M_S=S_{\max}}(\rho_1, \rho_2, \dots, \rho_N) = \left[Y_A \prod_{i=1}^N \phi_{n_i l_i m_{l_i}}(\vec{r}_i) \right] \prod_{j=1}^N \alpha(m_{s_j}) . \quad (3.17)$$

Pour la composante maximum de l'état de plus haute multiplicité de spin, la fonction d'onde antisymétrisée est encore⁶ le produit d'une fonction d'espace par une fonction de spin⁷. Maintenant, en faisant agir l'opérateur de descente $S_- = \sum_{i=1}^N s_{i-}$, on obtient la composante $M_{S_{\max}} - 1$ de l'états $S_{\max} = \frac{N}{2}$. Compte tenu de la forme de la fonction de spin dans (3.17) et de :

$$s_+ \alpha = s_- \beta = 0 , \quad s_- \alpha \propto \beta , \quad s_+ \beta \propto \alpha , \quad (3.18)$$

cette composante apparaît sous la forme d'une combinaison linéaire de N termes, contenant chacun $(N - 1)$ fonctions α et une fonction β (il y a N façons de basculer un spin) ; on peut donc affirmer qu'il existe $N - 1$ états de spin $S = S_{\max} - 1$. Plus généralement, il n'est pas difficile de montrer⁸ qu'il y a $C_N^p - C_N^{p-1}$ états de spin $S = S_{\max} - p$; la somme de toutes les multiplicités de spin doit bien évidemment être égale à 2^N , ce que l'on vérifie trivialement en additionnant les différences $C_N^p - C_N^{p-1}$.

Au total, choisir une valeur de L et une valeur de S revient à considérer $(2L + 1)(2S + 1)$ états propres constituant ce que l'on appelle un terme spectral, noté comme suit :

$${}^{2S+1}\mathcal{F}(L) \quad (\mathcal{F}(L) = S, P, D, F, \dots \text{ si } L = 0, 1, 2, 3, \dots) . \quad (3.19)$$

Ainsi, un terme spectral $S = 1, L = 2$ est noté 3D ; de même, l'état fondamental d'un atome ayant toutes ses couches complètes et un électron célibataire dans une orbitale s (l'argent par exemple) sera du type 2S . Pour distinguer deux termes de même symétrie, l'usage est d'ajouter un numéro d'ordre ; ainsi, on note 1^2S l'état

⁶dans l'approximation électrostatique.

⁷Cette propriété est toujours vraie dans le cas très simple de $N = 2$ électrons.

⁸ $\frac{N}{2} - p \geq 0$.

fondamental de l'hydrogène et 2^2S le premier état excité ayant $L = 0$. Enfin, rappelons que dans l'approximation à électrons indépendants, tous les termes spectraux issus d'une même configuration électronique construite sur des orbitales hydrogénoïdes sont dégénérés.

Tout ce qui a été dit précédemment concerne exclusivement l'approximation électrostatique. Une description plus précise doit prendre en compte l'interaction spin-orbite responsable, en tant que petite correction⁹, de la structure fine (éclatement des termes spectraux en plusieurs composantes formant ce que l'on appelle un multiplet).

Il convient donc d'inclure notamment les termes du genre $\xi(\vec{r}_i) \vec{l}_i \cdot \vec{s}_i$, en généralisation immédiate de ce qui a été fait pour l'unique électron de l'atome d'hydrogène. Toutefois, en toute rigueur, ceci ne suffit pas : on devrait également ajouter des termes du genre dipôle-dipôle du genre $\vec{s}_i \cdot \vec{s}_j$, représentant les interactions entre les moments magnétiques des électrons, et aussi des termes $\vec{l}_i \cdot \vec{s}_j$ (appelés *spin - autre orbite*) décrivant l'interaction du moment magnétique d'un électron avec les boucles de courant formées par les autres. Dans un souci de simplicité – et aussi parce que, en pratique, c'est souvent suffisant – on admettra que l'interaction spin-orbite pour un atome complexe est bien prise en compte par¹⁰ :

$$W_{\text{so}} = \sum_{i=1}^N \xi(\vec{r}_i) \vec{l}_i \cdot \vec{s}_i . \quad (3.20)$$

Dans cette approximation, le Hamiltonien à prendre en compte désormais est :

$$H_{\text{sf}} = H_e + W_{\text{so}} , \quad (3.21)$$

où H_e est donné en (3.1).

Ceci étant admis, il est clair que les moments cinétiques totaux orbital et de spin, $\vec{L} = \sum_i \vec{l}_i$ et $\vec{S} = \sum_i \vec{s}_i$, ne sont plus des constantes du mouvement. De fait, on a par exemple :

$$[\vec{L}^2, W_{\text{so}}] = \sum_i \sum_{j,k} \sum_{u,v=x,y,z} \xi(\vec{r}_i) [l_{uj} l_{uk}, l_{vi} s_{vi}] = i\hbar \sum_i \sum_{j,k} \sum_{u,v} \xi(\vec{r}_i) \varepsilon_{uvw} (\delta_{ki} l_{uj} l_{wk} + \delta_{ji} l_{wj} l_{uk}) s_{vi} , \quad (3.22)$$

où on a utilisé $[J_u, J_v] = i\hbar \varepsilon_{uvw} J_w$ ($\varepsilon_{uvw} = -\varepsilon_{vuw}$, $\varepsilon_{uvw} = +1$ si uvw se déduit de xyz par une permutation circulaire, -1 dans le cas contraire). D'où :

$$[\vec{L}^2, W_{\text{so}}] = i\hbar \sum_i \sum_{u,v=x,y,z} \xi(\vec{r}_i) s_{vi} \varepsilon_{uvw} \sum_j (l_{uj} l_{wi} + l_{wi} l_{uj}) . \quad (3.23)$$

La parenthèse à droite est toujours symétrique dans l'échange $u \leftrightarrow v$: si $i \neq j$, les opérateurs commutent, si $i = j$ la symétrie saute aux yeux. Comme ε_{uvw} est antisymétrique en u et v , la \sum_u donne zéro, d'où $[\vec{L}^2, W_{\text{so}}] = 0$, et de même $[\vec{S}^2, W_{\text{so}}] = 0$: les vecteurs \vec{L} et \vec{S} ne sont plus des constantes du mouvement, mais leurs modules le sont toujours. On peut donc encore affecter les nombres quantiques L et S aux états de structure fine, mais plus les M_L et M_S , tant que la description simplifiée par W_{so} est adéquate.

À l'inverse, évidemment, $\vec{J} = \vec{L} + \vec{S}$ est une constante du mouvement, c'est le moment cinétique total d'un système isolé (si on fait tourner tous les \vec{l}_i et tous les \vec{s}_i , la somme des produits scalaires dans W_{so} est évidemment invariante) ; cette propriété se vérifie d'ailleurs explicitement en calculant le commutateur de \vec{J} avec H_{sf} . Il en résulte que les vecteurs propres pertinents changent de nature ; il faut maintenant considérer¹¹ les $|\tau J M_J L S\rangle$ au lieu des $(2S+1)(2L+1)$ vecteurs $|\tau L S M_L M_S\rangle$. On passe des uns aux autres par des combinaisons linéaires contenant les coefficients de Clebsch-Gordan, la transformation se faisant à l'intérieur du sous-espace L et S fixés ; les valeurs possibles de J sont :

$$J = L + S, L + S - 1, \dots, |L - S| . \quad (3.24)$$

⁹Le rapport entre l'interaction spin-orbite et la différence d'énergie entre deux termes spectraux est typiquement de l'ordre de α^2 .

¹⁰Pour plus de détails, voir [14], p. 181.

¹¹ τ désigne l'ensemble des autres nombres quantiques nécessaires pour spécifier complètement l'état. Essentiellement, ce sont ceux qui sont associés à l'énergie propre du Hamiltonien électrostatique (sans structure fine).

Une conséquence importante est que tout élément de matrice du Hamiltonien entre deux états du genre $|\tau JM_J LS\rangle$ est nul si l'un au moins des nombres quantiques de moment angulaire est différent du *bra* au *ket* :

$$\langle \tau' J' M_J' L' S' | H_{\text{sf}} | \tau JM_J LS \rangle \propto \delta_{JJ'} \delta_{M_J M_J'} \delta_{LL'} \delta_{SS'} . \quad (3.25)$$

Comme toujours, les considérations de symétrie permettent de simplifier grandement le problème.

La structure fine est une petite correction¹² pour les niveaux de type électrostatique, qu'ils soient traités à l'intérieur de l'approximation des électrons indépendants, ou que la répulsion de Coulomb entre ceux-ci soit partiellement prise en compte par un champ central effectif. Il est donc légitime de la traiter par une technique de perturbation pour un niveau dégénéré. L'approximation consiste donc ici à négliger les éléments de matrice de type¹³ $\langle \tau JM_J LS | H_{\text{sf}} | \tau' JM_J LS \rangle$, $\tau' \neq \tau$, et est valide tant que $|\langle \tau JM_J LS | W_{\text{so}} | \tau' JM_J LS \rangle| \ll |E_\tau - E_{\tau'}|$.

Dans ce cadre, les niveaux de structure fine perturbés sont obtenus par diagonalisation de W_{so} projeté dans le sous-espace dégénéré engendré par les $|\tau LSM_L M_S\rangle$ ou les $|\tau JM_J LS\rangle$ correspondants. La première base fournit une matrice non-diagonale, la seconde une matrice diagonale (voir (3.25)) : il est donc plus simple et plus direct de se placer d'emblée sur la base $|\tau JM_J LS\rangle$; alors, la lecture des éléments diagonaux fournira les corrections d'énergie au premier ordre (comme L et S sont fixés une fois pour toutes, on note désormais éventuellement $|\tau JM_J LS\rangle \equiv |\tau JM_J\rangle$ pour simplifier).

Pour calculer ces éléments diagonaux, il suffit d'invoquer à plusieurs reprises le théorème de Wigner-Eckart, appliqué au cas simple des opérateurs vectoriels. Selon ce théorème, pour tout opérateur vectoriel \vec{V} , on a la relation¹⁴ :

$$\langle \tau JM_J | \vec{V} | \tau JM_J \rangle = \langle \tau JM_J | \vec{J} | \tau JM_J \rangle \frac{\langle \vec{J} \cdot \vec{V} \rangle}{\hbar^2 J(J+1)} . \quad (3.26)$$

On choisit d'abord $\vec{J} = \vec{L}$ et $\vec{V} = \vec{l}_i$ dans (3.26) ; ceci permet de montrer que, sur la base $|\tau LSM_L M_S\rangle$, \vec{l}_i et \vec{L} ont des éléments de matrice proportionnels ; la même opération faite avec $\vec{J} = \vec{S}$ et $\vec{V} = \vec{s}_i$ établit que $\vec{s}_i \propto \vec{S}$. En définitive, sur cette base, W_{so} et $\vec{L} \cdot \vec{S}$ ont des éléments de matrice proportionnels. Cette propriété est évidemment préservée quand on passe aux $|\tau JM_J LS\rangle$ (la transformation (unitaire) se fait à L, S fixés). En définitive, pour un même terme spectral ${}^{2S+1}\mathcal{F}$ (correspondant précisément à L, S donnés), il est équivalent, à un facteur près, de manipuler W_{so} ou l'opérateur $\propto \vec{L} \cdot \vec{S}$, ce que l'on résume par :

$$W_{\text{so}} \rightarrow A(\tau, L, S) \vec{L} \cdot \vec{S} \quad \text{pour un multiplet donné} , \quad (3.27)$$

où $A(\tau, L, S)$ est une constante caractéristique du terme spectral considéré. L'opérateur $\vec{L} \cdot \vec{S}$ est visiblement de trace nulle : la somme des corrections d'énergie pondérée par les dégénérescences est nulle.

En utilisant maintenant l'identité :

$$\vec{L} \cdot \vec{S} = \frac{1}{2} (\vec{J}^2 - \vec{L}^2 - \vec{S}^2) , \quad (3.28)$$

on voit que les corrections d'énergie donnant la structure fine sont données par des expressions du genre :

$$\frac{\hbar^2}{2} A(\tau, L, S) [J(J+1) - L(L+1) - S(S+1)] . \quad (3.29)$$

Ce résultat s'appelle "la règle des intervalles de Landé". En l'absence de champ magnétique, tout multiplet $|\tau JM_J LS\rangle$ reste dégénéré $2J+1$ fois. Par exemple, pour un terme spectral du type 2P , J prend les deux valeurs

¹²Les variations dues à la structure fine sont de l'ordre de 10^3 MHz – comparable d'ailleurs à un élargissement Doppler à l'ambiante. Ces corrections, étant d'origine relativiste, sont d'autant plus importantes que l'atome est lourd : l'énergie de l'atome variant en gros comme Z^2 , le théorème du Viriel permet d'affirmer que, au moins semi-qualitativement, la vitesse quadratique moyenne d'un électron augmente elle aussi comme Z^2 . Pour un atome lourd, les électrons internes vont très vite (ils voient une grande charge nucléaire, et doivent donc tourner très vite pour rester en équilibre) – leurs énergies atteignent donc rapidement quelques dizaines de keV, ce qui n'est pas si petit devant 511 keV ; au contraire, les électrons périphériques voient une charge nucléaire fortement écrantée par les électrons internes, de sorte que leurs énergies typiques restent dans le domaine de l'eV.

¹³Ces éléments de matrice n'ont aucune raison d'être nuls, puisqu'ils ont les mêmes nombres quantiques pour tous les moments cinétiques.

¹⁴Dans le cas des opérateurs vectoriels, cette relation est parfois appelée *théorème de projection*. Il est le fondement théorique du *modèle vectoriel* de l'atome.

$1 \pm \frac{1}{2}$, soit $\frac{1}{2}$ et $\frac{3}{2}$; l'usage est de compléter la notation du terme spectral en indiquant la valeur de J en indice : dans le cas présent, on obtient les deux multiplets ${}^2P_{\frac{1}{2}}$ et ${}^2P_{\frac{3}{2}}$. Les dégénérescences respectives sont 2 et 4, leur somme vaut 6 et est bien égale à la dégénérescence ($= 2 \times 3$) du terme spectral 2P qui a donné naissance aux deux multiplets différant par leur J .

3.3 Théorème du Viriel

Dans l'impossibilité de résoudre exactement le problème à N -corps, on est dans l'obligation de recourir à des méthodes approchées dont il faut savoir contrôler la qualité. Celle-ci peut s'apprécier de diverses façons et par référence au but poursuivi ; elle peut également être mesurée par confrontation avec des propriétés générales que possède la solution exacte inconnue. Ces propriétés peuvent être des symétries requises, ou la satisfaction de relations fondamentales. Un exemple est fourni par le théorème du Viriel, bien connu en Mécanique Classique et qui se généralise au domaine quantique ; la discussion de ce théorème permet d'ailleurs de réaliser que certains garde-fous peuvent se révéler illusoire.

Dans ce qui suit¹⁵, on considère uniquement les interactions électrostatiques de Coulomb, qui sont des fonctions homogènes de degré -1 ; pour tout terme de Coulomb, $v(r)$, on a :

$$v(\lambda r) = \lambda^{-1} v(r) , \quad (3.30)$$

où λ est un facteur d'échelle de contraction ou de dilatation des longueurs (*scaling*). En particulier, si on note $\mathcal{V}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ l'ensemble des termes coulombiens (attraction du noyau et répulsion entre électrons), on a :

$$\mathcal{V}(\lambda \vec{r}_1, \lambda \vec{r}_2, \dots, \lambda \vec{r}_N) = \lambda^{-1} \mathcal{V}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) . \quad (3.31)$$

L'énergie potentielle électrostatique totale est aussi une fonction homogène de degré -1 . Par ailleurs, en représentation- q , tout terme cinétique est du genre $-\frac{\hbar^2}{2m}(\nabla_{\vec{r}})^2$: c'est donc aussi une fonction homogène des coordonnées, mais de degré -2 .

$\Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ désignant une fonction propre exacte normalisée¹⁶, définissons une autre fonction déduite de celle-ci¹⁷ par un *scaling* :

$$\Psi_{\lambda}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \lambda^{\frac{3N}{2}} \Psi(\lambda \vec{r}_1, \lambda \vec{r}_2, \dots, \lambda \vec{r}_N) . \quad (3.32)$$

Le facteur $\lambda^{\frac{3N}{2}}$ assure que Ψ_{λ} est aussi une fonction normalisée. Calculons maintenant les valeurs moyennes de l'énergie cinétique et de l'énergie potentielle dans l'état (fictif) Ψ_{λ} . En changeant les variables muettes dans les intégrales, on établit facilement :

$$\langle T \rangle(\lambda) \equiv \langle \Psi_{\lambda} | T | \Psi_{\lambda} \rangle = \lambda^2 \langle \Psi | T | \Psi \rangle \equiv \lambda^2 \langle T \rangle , \quad (3.33)$$

$$\langle \mathcal{V} \rangle(\lambda) \equiv \langle \Psi_{\lambda} | \mathcal{U} | \Psi_{\lambda} \rangle = \lambda \langle \Psi | \mathcal{V} | \Psi \rangle \equiv \lambda \langle \mathcal{V} \rangle . \quad (3.34)$$

Maintenant, par le Principe Variationnel, on peut affirmer que dans la famille de fonctions $\{\Psi_{\lambda}\}_{\lambda}$ il existe une fonction donnant sa plus petite valeur à l'énergie $E(\lambda) = \langle T \rangle(\lambda) + \langle \mathcal{V} \rangle(\lambda)$, obtenue en minimisant $E(\lambda)$. La condition de minimisation est :

$$2\lambda \langle T \rangle + \langle \mathcal{V} \rangle = 0 . \quad (3.35)$$

Comme la famille variationnelle contient la solution exacte Ψ , qui n'est autre que $\Psi_{\lambda=1}$, la relation (3.35) est vraie en $\lambda = 1$. On en déduit que, pour la fonction exacte, le Théorème du Viriel pour le champ Coulombien s'énonce :

$$2 \langle T \rangle = - \langle \mathcal{V} \rangle . \quad (3.36)$$

On peut donc vouloir exiger que toute fonction approchée satisfasse ce théorème, en gage de sa qualité. Ce critère n'est pas forcément sûr. En effet, supposons que la fonction Ψ ci-dessus ne soit pas la fonction exacte.

¹⁵L'argument suivant est dû à Fock.

¹⁶En pratique, il s'agira de l'état fondamental puisque la suite de l'argument s'appuie sur le Principe Variationnel.

¹⁷ Ψ coïncide formellement avec $\Psi_{\lambda=1}$.

Alors la relation de minimisation, qui donne maintenant la valeur de λ définissant la meilleure fonction au sens du Principe Variationnel, ne fournit plus $\lambda = 1$ (selon (3.35)), la valeur optimale de λ est donnée par le rapport $-\frac{\langle \mathcal{V} \rangle}{2\langle T \rangle}$; cependant, la même relation de minimisation s'écrit aussi :

$$2\lambda^2 \langle T \rangle + \lambda \langle \mathcal{V} \rangle = 0 , \quad (3.37)$$

soit, compte tenu de (3.33) et (3.34) :

$$2\langle T \rangle(\lambda) = -\langle \mathcal{V} \rangle(\lambda) . \quad (3.38)$$

Autrement dit, disposant d'une fonction approchée quelconque, un simple *scaling* permet de construire une autre fonction approchée satisfaisant le Théorème du Viriel : il suffit de faire une nouvelle variation en minimisant l'énergie par rapport au facteur d'échelle. En conclusion, la satisfaction du Viriel n'est pas, en soi, un signe de la qualité d'une fonction d'onde approchée.

En outre, la portée de ce théorème est souvent limitée en pratique du fait que le potentiel "nu" constitue une idéalisation qui n'est pas toujours acceptable. En particulier, en présence d'écrantage par les électrons internes, le potentiel ressenti par les électrons de valence est à grande distance plutôt du genre Yukawa et invalide de ce fait les hypothèses conduisant au théorème du Viriel.

3.4 Au-delà de l'approximation à électrons indépendants

S'agissant de présenter les descriptions approchées, il serait vain de dresser le catalogue de toutes les méthodes et techniques disponibles. On se bornera ici à dégager quelques idées et à ne présenter en détail que l'une d'entre elles, la méthode de Hartree - Fock, dont la portée et l'intérêt sont universels et débordent largement le champ de la Physique Atomique.

D'une façon ou d'une autre, c'est le Principe Variationnel qui est le plus souvent à l'œuvre dans l'élaboration d'une méthode approchée. Ce "principe" n'en est d'ailleurs pas un à proprement parler. On montre [1] qu'il y a équivalence entre équation aux valeurs propres et minimisation d'une certaine fonctionnelle. Plus précisément, soit d'une part un mode propre normalisable (Ψ, E) satisfaisant :

$$H \Psi = E \Psi . \quad (3.39)$$

D'autre part, considérons la fonctionnelle $E[\Psi]$ définie comme :

$$E[\Psi] = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} . \quad (3.40)$$

Il est possible de montrer [1] que tout vecteur minimisant $E[\Psi]$ est vecteur propre de H , associé à une valeur propre égale à la valeur de $E[\Psi]$ à son minimum, et réciproquement. Autrement dit, l'équivalence suivante est vraie :

$$H \Psi = E \Psi \iff \delta E[\Psi] = 0 . \quad (3.41)$$

En soi, ce résultat n'est pas très utile, car résoudre (3.39) ou trouver Ψ tel que $\delta E[\Psi] = 0$ sont deux tâches d'une égale difficulté. En revanche, ce qui est important c'est le corollaire suivant :

$$\forall |\Psi\rangle, \langle \Psi | \Psi \rangle = 1, \quad \langle \Psi | H | \Psi \rangle \geq E_{\text{fond}} , \quad (3.42)$$

inégalité affirmant que l'énergie calculée avec n'importe quel état normalisé à l'unité est toujours *supérieure* à l'énergie exacte du fondamental. La démonstration est élémentaire et se fait comme suit, dans l'hypothèse où le spectre de H est entièrement discret. L'hermiticité de H assure que l'ensemble de ses fonctions propres $|\psi_n\rangle$, associée chacune à la valeur propre E_n , constitue une base complète orthonormée ; quel que soit le vecteur normalisé $|\Psi\rangle$, le développement suivant existe :

$$\forall |\Psi\rangle, \quad |\Psi\rangle = \sum_{n=1}^{+\infty} c_n |\psi_n\rangle \quad \sum_{n=1}^{+\infty} |c_n|^2 = 1 . \quad (3.43)$$

Il en résulte :

$$E \equiv \langle \Psi | H | \Psi \rangle = \sum_{n=1}^{+\infty} |c_n|^2 E_n . \quad (3.44)$$

Supposons maintenant les énergies $\{E_n\}_n$ rangées par ordre croissant ($E_n < E_{n+1}$) ; alors, en minorant chaque énergie de la série (3.44) par E_1 :

$$\langle \Psi | H | \Psi \rangle \geq \sum_{n=1}^{+\infty} |c_n|^2 E_1 = E_1 \equiv E_{\text{fond}} , \quad (3.45)$$

ce qui établit le résultat (3.42). La méthode à suivre est donc transparente dans son principe : on définit une famille de fonctions dépendant d'un ou plusieurs paramètres λ_i . La meilleure fonction de cette famille, au sens de la méthode variationnelle, est par définition celle qui donne à l'énergie sa plus petite valeur possible et sera donc trouvée par minimisation de $E(\lambda_i)$. Il est important de remarquer que cette méthode ne fonctionne que pour l'état fondamental ; s'il est évidemment possible de la généraliser formellement aux états excités¹⁸, cette généralisation est en pratique illusoire.

L'exemple sans doute le plus élémentaire d'application de la méthode variationnelle est la description de l'état fondamental de l'atome d'hélium. Une idée simple consiste à définir une famille de fonctions d'essai dont la forme s'inspire des fonctions du modèle à électrons indépendants. Dans ce cas – en l'absence donc d'interaction $\frac{e'^2}{r_{12}}$ – le produit d'orbitales suivant est propre (et normalisé) :

$$\psi_Z(\vec{r}_1, \vec{r}_2) = \frac{Z^3}{\pi a_0^3} e^{-Z \frac{r_1}{a_0}} e^{-Z \frac{r_2}{a_0}} , \quad (3.46)$$

à condition de choisir $Z = 2$. En fait, on va laisser “flotter” Z , en le prenant comme paramètre variationnel ; l'image physique sous-jacente est l'écrantage de la charge nucléaire par un électron vis-à-vis de l'autre.

Il faut donc calculer la fonctionnelle énergie¹⁹ $E[\psi_Z]$, qui est une certaine fonction $E(Z)$:

$$E(Z) = \langle \psi_Z | \frac{\vec{p}_1^2}{2m} + \frac{\vec{p}_2^2}{2m} - \frac{2e'^2}{r_1} - \frac{2e'^2}{r_2} + \frac{e'^2}{r_{12}} | \psi_Z \rangle . \quad (3.47)$$

Les contributions à un corps ne présentent aucune difficulté ; le terme à deux corps peut se calculer en développant $\frac{e'^2}{r_{12}}$ en harmoniques sphériques²⁰. Au total, on trouve :

$$E(Z) = \frac{e'^2}{a_0} \left(Z^2 - \frac{27}{8} Z \right) . \quad (3.48)$$

Cette fonction présente bien un minimum, survenant pour $Z_0 = \frac{27}{16} \simeq 1.69$ et donnant :

$$E(Z_0) \simeq -2.85 \frac{e'^2}{a_0} \geq -2.904 \frac{e'^2}{a_0} \equiv E_{\text{exp}} . \quad (3.49)$$

L'erreur sur l'énergie est de l'ordre de 2%, ce qui est assez remarquable compte tenu de la rusticité de la méthode²¹. Bien évidemment, cette approche peut être raffinée de bien des façons. Un calcul célèbre est dû à Hylleraas en 1930, qui a considéré des fonctions d'essai dépendant explicitement de la distance r_{12} ; les résultats sont bien meilleurs, mais la méthode n'a pas pu, semble-t-il, être généralisée au cas d'atomes plus complexes et constitue plutôt un remarquable objet de musée.

¹⁸Par exemple, si on connaît le fondamental *exact*, Ψ_1 , n'importe quelle fonction orthogonale à ce fondamental donne une énergie supérieure à celle du 1^{er} état excité. En pratique, il est rare de connaître le fondamental exact...

¹⁹Attention ! Le principe de la méthode consiste à calculer la valeur moyenne du Hamiltonien *exact* avec une fonction d'essai choisie : ne pas mettre Z – le paramètre variationnel – dans les termes d'attraction électron-noyau (voir (3.47))... Noter aussi qu'ici Z joue finalement le rôle d'un facteur d'échelle.

²⁰Pour plus de détails, voir [5], p. 258.

²¹Par comparaison, une méthode ordinaire de perturbation (où Z est fixé à la valeur 2) donne un résultat entaché d'une erreur de 5%.

Une autre méthode, dite de l'interaction de configuration, consiste à élargir l'espace de recherche et à considérer un espace vectoriel de fonctions associées à *plusieurs* configurations électroniques ayant les bonnes symétries. Par exemple, même pour le fondamental de l'hélium, il est sensé de considérer tous les déterminants issus par exemple des deux configurations $(1s)^2$ et $(1s)(2s)$, à condition de retenir pour cette dernière le seul état singulet ($S = 0$) qui en résulte. On trouve alors²², par diagonalisation de la matrice du Hamiltonien, que l'état fondamental est une combinaison linéaire des deux singulets, avec toutefois un poids prépondérant ($\simeq 90\%$) sur le singulet construit avec $(1s)^2$. Il est clair que l'intérêt de cette méthode dépend en grande partie de l'outil numérique disponible. En outre, si l'on peut affirmer qu'en augmentant la dimension de l'espace – c'est-à-dire en augmentant le nombre de configurations de bonne symétrie – on ira forcément vers une amélioration de la description, la "convergence" peut être très lente, en particulier parce que les fonctions hydrogénoïdes liées ne sont pas complètes en raison de l'existence des états du *continuum* associés aux états de diffusion.

Décrivons maintenant en détail l'importante méthode de Hartree - Fock²³, dont le champ d'application est immense (atomes complexes, molécules, matière condensée, matière nucléaire, *etc.*). L'idée de base est celle de toutes les méthodes de champ moyen : on va essayer de remplacer le terme d'interaction à deux corps par un terme effectif représentant l'action moyenne de tous les électrons sur l'un d'entre eux. Autrement dit, on admet qu'il est possible de trouver une approximation où finalement le terme à deux corps est remplacé par un terme ne dépendant que des variables d'un seul électron, l'action de tous les autres étant moyennée d'une certaine façon.

Dès lors, on retrouve *formellement* une description où le Hamiltonien effectif est une somme de termes à un corps et possède donc des fonctions propres construites sur les produits de spin-orbitales. Compte tenu des nécessités imposées par le Principe de Pauli, la fonction générique de l'approximation de Hartree - Fock est donc un déterminant de Slater, construit sur *des* spin-orbitales orthonormées :

$$\Psi_A(1, 2, \dots, N) = \sqrt{N!} Y_A \prod_{i=1}^N \psi_{k_i}(i) \equiv \sqrt{N!} Y_A \Phi(1, 2, \dots, N) . \quad (3.50)$$

Par opposition avec les écritures du modèle à particules indépendantes où les spin-orbitales sont données d'avance, les $\{\psi_{k_i}\}$ sont maintenant *inconnues* et toute la question est précisément de les trouver, c'est-à-dire d'écrire les équations qu'elles satisfont (équations SCF), et de les résoudre.

En définitive, il s'agit de trouver la meilleure fonction du genre déterminant ; une méthode pour obtenir les équations SCF consiste à invoquer le Principe Variationnel. Il convient donc d'abord de calculer l'expression formelle de l'énergie moyenne E associée à un tel déterminant et d'effectuer la minimisation²⁴.

Le Hamiltonien (3.1) – comme toute observable relative à un ensemble de particules identiques –, commute avec toutes les permutations du groupe symétrique \mathcal{S}_N . En vertu des relations (3.8), et en prenant une fonction du type (3.50), on trouve :

$$E \equiv \langle \Psi_A | H | \Psi_A \rangle = N! \langle Y_A \Phi | H | Y_A \Phi \rangle = \langle \Phi | H | N! Y_A \Phi \rangle . \quad (3.51)$$

H contient deux types de termes (voir (3.1) ; $H_1(i)$ n'agit que sur une particule à la fois ; en vertu de l'orthogonalité des spin-orbitales, tout terme du ket issu d'une simple transposition (la plus "petite" permutation) donnera zéro. Il reste donc :

$$E = \sum_{i=1}^N \langle \psi_{k_i} | H_1 | \psi_{k_i} \rangle + \sum_{\lambda=1}^{N!} (-1)^{\sigma_\lambda} \langle \Phi | V | P_\lambda \Phi \rangle . \quad (3.52)$$

Le terme V agit sur deux électrons au plus ; si donc la permutation P_λ implique plus de deux particules, le terme correspondant est nul par l'orthogonalité des spin-orbitales. L'expression de l'énergie se réduit à :

$$E = \sum_{i=1}^N \langle \psi_{k_i} | H_1 | \psi_{k_i} \rangle + \frac{1}{2} \sum_{i \neq j} \langle \Phi | \frac{e'^2}{r_{ij}} [1 - \sum_{r,s} T_{rs}] | \Phi \rangle \quad (3.53)$$

²²Le calcul se fait à la main sans difficulté.

²³En anglais : *self-consistent field*, en abrégé SCF. En français, on désigne aussi cette approximation par *méthode du champ auto-cohérent*.

²⁴Le calcul effectué plus haut pour l'Hélium avec la fonction (3.46) n'est pas un calcul SCF : rien ne permet d'affirmer que les orbitales SCF sont les fonctions $1s$ hydrogénoïdes ! Le choix d'exponentielles dans (3.46) restreint l'ensemble des fonctions participant à la variation, alors que, dans le cadre SCF, on n'impose rien *a priori* (en principe du moins).

où les T_{rs} sont les transpositions échangeant r et s (le signe $-$ vient de la signature des transpositions). Parmi tous les termes issus des deux sommes doubles, seuls ceux où i et j sont aussi les variables permutées donnent une contribution non-nulle. Au total, la valeur moyenne de l'énergie a pour expression :

$$\begin{aligned} E &= \sum_{i=1}^N \langle \psi_{k_i} | H_1 | \psi_{k_i} \rangle + \frac{1}{2} \sum_{i \neq j} \left[\langle \psi_{k_i} \psi_{k_j} | \frac{e'^2}{r_{ij}} | \psi_{k_i} \psi_{k_j} \rangle - \langle \psi_{k_i} \psi_{k_j} | \frac{e'^2}{r_{ij}} | \psi_{k_j} \psi_{k_i} \rangle \right] \\ &\equiv \sum_{i=1}^N \langle \psi_{k_i} | H_1 | \psi_{k_i} \rangle + \frac{1}{2} \sum_{i \neq j} (\mathcal{J}_{k_i k_j} - \mathcal{K}_{k_i k_j}) . \end{aligned} \quad (3.54)$$

Par convention, dans les éléments de matrice à deux corps, les variables muettes de sommation sont écrites de gauche à droite dans l'ordre naturel. Les éléments de matrice se calculent en effectuant les produits scalaires comme d'habitude : intégration sur les variables orbitales et sommation sur les variables de spin. La valeur moyenne de V apparaît ainsi comme une somme de deux termes appelés respectivement intégrales directes et d'échange et définis ici en termes de spin-orbitales. Si on écrit chaque spin-orbitale sous la forme :

$$\psi_{k_i}(1) = \phi_{k_j}(\vec{r}_1) \xi_j(m_{s1}) , \quad (3.55)$$

alors :

$$\begin{aligned} \mathcal{J}_{k_i k_j} &= \sum_{m_{s1}} \sum_{m_{s2}} \int d^3 r_1 d^3 r_2 \phi_{k_i}^*(\vec{r}_1) \phi_{k_j}^*(\vec{r}_2) \frac{e'^2}{r_{12}} \phi_{k_i}(\vec{r}_1) \phi_{k_j}(\vec{r}_2) \xi_i(m_{s1}) \xi_i(m_{s1}) \xi_j(m_{s2}) \xi_j(m_{s2}) \\ &= \int d^3 r_1 d^3 r_2 \phi_{k_i}^*(\vec{r}_1) \phi_{k_j}^*(\vec{r}_2) \frac{e'^2}{r_{12}} \phi_{k_i}(\vec{r}_1) \phi_{k_j}(\vec{r}_2) \equiv J_{ij} \end{aligned} \quad (3.56)$$

$$\begin{aligned} \mathcal{K}_{k_i k_j} &= \sum_{m_{s1}} \sum_{m_{s2}} \int d^3 r_1 d^3 r_2 \phi_{k_i}^*(\vec{r}_1) \phi_{k_j}^*(\vec{r}_2) \frac{e'^2}{r_{12}} \phi_{k_j}(\vec{r}_1) \phi_{k_i}(\vec{r}_2) \xi_j(m_{s1}) \xi_i(m_{s1}) \xi_i(m_{s2}) \xi_j(m_{s2}) \\ &= \int d^3 r_1 d^3 r_2 \phi_{k_i}^*(\vec{r}_1) \phi_{k_j}^*(\vec{r}_2) \frac{e'^2}{r_{12}} \phi_{k_j}(\vec{r}_1) \phi_{k_i}(\vec{r}_2) \sum_{m_s} \xi_j(m_s) \xi_i(m_s) \equiv K_{ij} \langle \xi_j | \xi_i \rangle , \end{aligned} \quad (3.57)$$

où l'on a utilisé le fait que le produit scalaire de spin vaut 0 ou 1 et est, de ce fait, égal à son propre carré, et où ont été réintroduites les intégrales directes et d'échange relatives aux orbitales. Remarquons que dans la somme double de (3.54), on peut relâcher la contrainte $i \neq j$ puisque le terme correspondant est nul.

Les équations SCF vont s'obtenir en effectuant une petite variation de tous les $\{\psi_{k_i}\}$ et en annulant la variation correspondante de l'énergie, δE . À partir de (3.54)²⁵, la condition de stationnarité de E ($\delta E = 0$) par rapport aux variations *indépendantes* des $\{\psi_{k_i}\}$ montre que les fonctions cherchées doivent satisfaire les équations suivantes²⁶ ($v(r_{12}) = \frac{e'^2}{r_{12}}$) :

$$H_1(1) \psi_{k_i}(1) + \sum_j \int_{\vec{x}_2} \psi_{k_j}^*(2) v(r_{12}) \psi_{k_j}(2) \psi_{k_i}(1) - \sum_j \int_{\vec{x}_2} \psi_{k_j}^*(2) v(r_{12}) \psi_{k_i}(2) \psi_{k_j}(1) = \varepsilon \psi_{k_i}(1) , \quad (3.58)$$

où les ψ_{k_i} et les valeurs propres ε sont les inconnues à déterminer (bien noter que ψ_{k_i} est en facteur de la première somme, mais est à l'intérieur de la seconde).

Ces équations sont relativement aisées à interpréter ; la première somme double est finalement l'interaction électrostatique entre un électron et la densité résultant de tous les autres, distribués suivant la densité $\sum_j |\psi_{k_j}|^2$. C'est le terme dit de Hartree, qui serait seul présent si la fonction d'onde pouvait être choisie comme un simple produit (non-antisymétrisé) et qui, à lui seul, justifie la terminologie de "champ moyen". D'ailleurs, les équations (3.58) peuvent aussi s'écrire :

$$\tilde{H}_1(1) \psi_{k_i}(1) - \sum_j \int_{\vec{x}_2} \psi_{k_j}^*(2) v(r_{12}) \psi_{k_i}(2) \psi_{k_j}(1) = \varepsilon \psi_{k_i}(1) , \quad \tilde{H}_1(1) = H_1(1) + \sum_j \int_{\vec{x}_2} \psi_{k_j}^*(2) v(r_{12}) \psi_{k_j}(2) , \quad (3.59)$$

²⁵ Le facteur $\frac{1}{2}$ disparaît car les termes d'interaction sont quadratiques par rapport aux spin-orbitales.

²⁶ Le symbole $\int_{\vec{x}}$ représente une sommation sur le spin et une intégration sur la coordonnée \vec{r} ($\vec{x} \equiv (\vec{r}, m_s)$).

avec :

$$\tilde{H}_1(1) = H_1(1) + \sum_j \int_{\vec{x}_2} \psi_{k_j}^*(2) v(r_{12}) \psi_{k_j}(2) \equiv H_1(1) + \sum_j \int_{\vec{x}_2} [\sum_j |\psi_{k_j}(2)|^2] v(r_{12}) ; \quad (3.60)$$

\tilde{H}_1 est le Hamiltonien de Hartree, et décrit le mouvement d'un électron dans le champ électrostatique moyen de tous les autres. En revanche, la seconde sommation dans (3.58), qui résulte directement du Principe de Pauli, n'a pas d'interprétation classique – ce n'est pas une surprise.

En introduisant le Hamiltonien dit SCF, H_{SCF} , défini par son action sur une spin-orbitale comme reconstituant le premier membre de (3.58), cette dernière prend formellement l'allure d'une équation aux fonctions propres :

$$H_{\text{SCF}}[\{\psi_{k_j}\}_j] \psi_{k_i} = \varepsilon \psi_{k_i} . \quad (3.61)$$

Il faut cependant bien voir que l'opérateur au premier membre contient un terme intégral (non-local), et est une fonction explicite de ses fonctions propres.

La résolution des équations de Hartree - Fock nécessite un ordinateur et procède par itération ; on se donne un jeu de fonctions au départ, avec lesquelles on calcule l'opérateur du premier membre, ou plus précisément sa matrice sur une base donnée. Après diagonalisation, on obtient un jeu de fonctions propres avec lesquelles le premier membre est recalculé, et ainsi de suite. Comme toute méthode itérative, le choix du point de départ est crucial pour la rapidité de la convergence²⁷. Une fois celle-ci obtenue, on dispose donc d'un jeu de spin-orbitales auto-cohérentes, satisfaisant :

$$H_{\text{SCF}} \psi_{\lambda}^{\text{SCF}} = \varepsilon_{\lambda}^{\text{SCF}} \psi_{\lambda}^{\text{SCF}} . \quad (3.62)$$

L'énergie associée au déterminant Ψ_{SCF} construit sur ces solutions n'est pas égale à $\sum_{\lambda} \varepsilon_{\lambda}^{\text{SCF}}$ – alors on prendrait en compte deux fois l'interaction au sein d'une même paire –, mais doit être calculée en prenant la valeur moyenne du Hamiltonien de départ avec Ψ_{SCF} :

$$E_{\text{SCF}} = \langle \Psi_{\text{SCF}} | H | \Psi_{\text{SCF}} \rangle . \quad (3.63)$$

Bien évidemment, on a :

$$E_{\text{SCF}} \geq E_{\text{exacte}} ; \quad (3.64)$$

de façon (très) conventionnelle, on appelle "énergie de corrélation" E_{corr} la différence entre ces deux énergies :

$$E_{\text{corr}} = E_{\text{SCF}} - E_{\text{exacte}} \geq 0 . \quad (3.65)$$

La terminologie est peu heureuse : au sein d'un schéma de Hartree - Fock, les électrons sont *déjà* corrélés.

3.5 L'atome d'hélium

L'atome d'hélium ($Z = 2$) est l'atome le plus simple contenant déjà toute la complexité des systèmes polyélectroniques. Une description précise doit évidemment inclure les termes d'interaction spin-orbite donnant lieu à la structure fine mais, pour s'en tenir à l'approche la plus simple, on ne retiendra ici que les interactions de type électrostatique de Coulomb – étant entendu que les corrections relativistes ne peuvent être ignorées dans le cas des atomes lourds²⁸. Dans ces conditions, le Hamiltonien s'écrit avec des notations évidentes²⁹ :

$$H(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2) = \frac{\vec{p}_1^2}{2m} - \frac{2e'^2}{r_1} + \frac{\vec{p}_2^2}{2m} - \frac{2e'^2}{r_2} + \frac{e'^2}{r_{12}} , \quad (3.66)$$

ou, sous une forme plus concise :

$$H(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2) = [H_1(\vec{r}_1, \vec{p}_1) + H_1(\vec{r}_2, \vec{p}_2)] + \frac{e'^2}{r_{12}} \equiv H_0 + V . \quad (3.67)$$

²⁷La convergence est parfois problématique, et peut être sujette à des instabilités, ou à des brisures spontanées de symétrie vis-à-vis du spin total.

²⁸Il n'empêche que la structure fine de l'hélium est aisément observable.

²⁹On maintient l'approximation du noyau infiniment massif.

$H_1(\vec{r}, \vec{p}) = \frac{p^2}{2m} - \frac{2e'^2}{r}$ est le Hamiltonien hydrogénoïde pour $Z = 2$. La présence dans H du terme de répulsion $\frac{e'^2}{r_{12}}$ fait de ce problème d'apparence pourtant simple un problème non-soluble exactement, tout comme l'est le problème des trois corps en Mécanique Classique. On dispose heureusement d'un arsenal de méthodes pour obtenir des solutions approchées dont le calcul est facilité par l'outil numérique ; on devine que ce dernier joue un rôle précieux en pratique.

Un point de départ pas trop déraisonnable consiste, dans un premier temps, à oublier le terme répulsif entre les deux électrons, même si l'on se doute qu'il est de l'ordre de la dizaine d'eV. Ceci revient à considérer d'abord H_0 défini en (3.67), qui est une somme d'opérateurs à une particule dont on connaît les solutions exactes (H_0 définit l'approximation des électrons indépendants). Si ϕ_n est une fonction propre hydrogénoïde avec $Z = 2$, d'énergie :

$$E_n = -\frac{m(2e'^2)^2}{2n^2\hbar^2} \simeq -\frac{54.45}{n^2} \text{ eV} , \quad (3.68)$$

alors tout produit $\phi_n\phi_m$ est fonction propre de H_0 :

$$H_0 \phi_n(\vec{r}_1)\phi_m(\vec{r}_2) = (E_n + E_m) \phi_n(\vec{r}_1)\phi_m(\vec{r}_2) \equiv E_{nm}\phi_n(\vec{r}_1)\phi_m(\vec{r}_2) . \quad (3.69)$$

Les ϕ_n ne contiennent pas le spin ; il convient donc de les multiplier par une fonction χ ($= \alpha$ ou β) ; la fonction résultante est toujours propre de H_0 puisque H_0 (pas plus que H d'ailleurs) ne contient le spin :

$$H_0 \phi_n(\vec{r}_1)\chi(m_{s_1})\phi_m(\vec{r}_2)\chi(m_{s_2}) = E_{nm}\phi_n(\vec{r}_1)\chi(m_{s_1})\phi_m(\vec{r}_2)\chi(m_{s_2}) . \quad (3.70)$$

À partir de la fonction $\phi_n(\vec{r}_1)\chi(m_{s_1})\phi_m(\vec{r}_2)\chi(m_{s_2})$, il faut construire l'état antisymétrique Ψ_A . Pour un système de deux fermions, après antisymétrisation, Ψ_A se présente encore sous la forme factorisée, variables d'espace d'un côté, variables de spin de l'autre. Les fonctions de spin sont les X_{S, M_S} :

1. état triplet, dont les composantes sont symétriques dans l'échange :

$$X_{1;+1} = \alpha(m_{s_1})\alpha(m_{s_2}) , X_{1;-1} = \beta(m_{s_1})\beta(m_{s_2}) , X_{1;0} = \frac{1}{\sqrt{2}}[\alpha(m_{s_1})\beta(m_{s_2}) + \beta(m_{s_1})\alpha(m_{s_2})] . \quad (3.71)$$

2. état singulet, dont la seule et unique composante est antisymétrique :

$$X_{1,0} = \frac{1}{\sqrt{2}}[\alpha(m_{s_1})\beta(m_{s_2}) - \beta(m_{s_1})\alpha(m_{s_2})] . \quad (3.72)$$

À ces fonctions de spin, il convient d'associer des fonctions d'espace ayant la bonne symétrie. Pour l'état triplet, la bonne fonction d'espace est :

$$\Phi_1(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}}[\phi_n(\vec{r}_1)\phi_m(\vec{r}_2) - \phi_m(\vec{r}_1)\phi_n(\vec{r}_2)] , \quad (3.73)$$

tandis que la partie spatiale du singulet est :

$$\Phi_0(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}}[\phi_n(\vec{r}_1)\phi_m(\vec{r}_2) + \phi_m(\vec{r}_1)\phi_n(\vec{r}_2)] . \quad (3.74)$$

En définitive, les états construits sur un produit d'orbitales sont :

$$\Psi_1 = \frac{1}{\sqrt{2}}[\phi_n\phi_m - \phi_m\phi_n] X_{1, M_S} \quad (\text{triplet}) \quad (3.75)$$

$$\Psi_0 = \frac{1}{\sqrt{2}}[\phi_n\phi_m + \phi_m\phi_n] X_{0,0} \quad (\text{singulet}) . \quad (3.76)$$

Au total, les fonctions ont la bonne symétrie : elles sont le produit d'une fonction d'espace symétrique (resp. antisymétrique) par une fonction de spin antisymétrique (resp. symétrique). Il faut toutefois se souvenir que cette séparation espace-spin en un seul produit dont chaque facteur est symétrique ou antisymétrique pour

toutes les composantes des multiplets ne s'observe que dans le cas $N = 2$ (ou, pour $N > 2$, pour l'état "ferromagnétique"). Enfin, comme les parties spatiales des fonctions d'onde complètes ont des symétries bien tranchées, on utilise parfois une terminologie spécifique : pour les états triplets (partie d'espace antisymétrique), on parle d'ortho-hélium, alors que l'hélium dans un état singulet est dit para-hélium.

Les états de spin étant déterminés, il convient de préciser la partie spatiale. L'usage désigne 1s, 2s, 2p, *etc.* les états hydrogénoïdes que l'on note ici ϕ_{nlm_l} , pour rappeler qu'il s'agit d'orbitales. Avec deux électrons, les configurations atomiques se forment en choisissant deux états de ce type ; dans l'ordre des énergies croissantes et compte tenu de la dégénérescence accidentelle du champ Coulombien, les configurations sont :

$$\begin{aligned} & (1s)^2 \\ & (1s) (2s) \quad (1s) (2p) \\ & (1s) (3s) \quad (1s) (3p) \quad (1s) (3d) \quad \textit{etc.} \end{aligned} \quad (3.77)$$

Pour un atome complexe, on définit plus généralement la notion de couche, notée $(nf(l))$, où $f(l) = s, p, d, f, \dots$ selon que $l = 0, 1, 2, 3, \dots$, et pouvant contenir $2n^2$ électrons, compte tenu du spin – qui ajoute un facteur 2 à la dégénérescence n^2 des états ayant tous le même nombre quantique principal n . Pour l'hélium, à la configuration $(1s)^2$ correspond le produit $\phi_n \phi_m = \phi_{100} \phi_{100}$ et ainsi de suite :

$$\begin{aligned} (1s)^2 & : \phi_{100} \phi_{100} \\ (1s) (2s) & : \phi_{100} \phi_{200} \quad (1s) (2p) : \phi_{100} \phi_{21m_l} \\ (1s) (3s) & : \phi_{100} \phi_{300} \quad (1s) (3p) : \phi_{100} \phi_{31m_l} \quad (1s) (3d) : \phi_{100} \phi_{32m_l} \quad \textit{etc.} \end{aligned} \quad (3.78)$$

Les états propres de H_0 sont complètement déterminés une fois la partie d'espace choisie. On remarque que si ϕ_n et ϕ_m sont une seule et même fonction orbitale ϕ_{nlm_l} , seul l'état singulet existe, en conformité avec le Principe d'exclusion de Pauli (techniquement : alors la partie spatiale de l'état triplet est identiquement nulle) ; dit autrement, si $\phi_n \equiv \phi_{nlm_l}$, les trois nombres quantiques orbitaux sont les mêmes pour les deux électrons, qui doivent donc avoir un spin différent ; ceci n'est pas possible pour l'état triplet.

Dans cette description initiale, l'état fondamental est un singulet S ($L = 0$) et est noté par les spectroscopistes 1^1S :

$$\Psi_{1^1S}(1, 2) = \phi_{100}(\vec{r}_1) \phi_{100}(\vec{r}_2) \frac{1}{\sqrt{2}} [\alpha(m_{s_1})\beta(m_{s_2}) - \beta(m_{s_1})\alpha(m_{s_2})] , \quad (3.79)$$

et son énergie vaut $2E_1 \simeq -108.9$ eV. Pour l'hélium, les états excités pertinents sont ceux pour lesquels l'un des électrons reste dans l'orbitale (1s) ; en effet, on constate numériquement que l'énergie de tout état où les deux électrons sont excités est supérieure à l'énergie de l'état fondamental de l'ion³⁰ He^+ . Ceci étant, le premier niveau excité est formé sur la configuration (1s) (2s) ou (1s) (2p), et contient à chaque fois un état triplet et un état singulet, puisque les deux orbitales ϕ_n et ϕ_m sont maintenant différentes :

$$\text{triplet} : \frac{1}{\sqrt{2}} [\phi_{100} \phi_{2lm_l} - \phi_{2lm_l} \phi_{100}] X_{1, M_S} \quad (l = 0, 1, -l \leq m_l \leq l) \quad (3.80)$$

$$\text{singulet} : \frac{1}{\sqrt{2}} [\phi_{100} \phi_{2lm_l} + \phi_{2lm_l} \phi_{100}] X_{0, 0} \quad (l = 0, 1, -l \leq m_l \leq l) . \quad (3.81)$$

Ce niveau d'ordre zéro a une énergie égale à $E_1 + E_2 \simeq -54.45 - \frac{54.45}{4} \simeq -68.06$ eV et a une dégénérescence d'ordre 16 : il y a 4 composantes de spin ($3 + 1$) et 4 possibilités pour le couple (l, m_l) . Le trou de Fermi est visible pour l'état triplet.

Si on choisit (1s) (2s), on a encore un état de type S ($L = 0$) ; d'où un singulet S et un triplet S, notés spectroscopiquement 2^1S et 3^3S , épuisant 4 directions dans le sous-espace dégénéré. De même, en prenant (1s) (2p), on forme deux états de type P ($L = 1$), notés 1^1P et 3^3P dont les dégénérescences sont 3 et 9, et ainsi de suite.

La description précédente, à particules indépendantes, n'est qu'un point de départ. L'inclusion la plus simple de la répulsion entre les électrons consiste à appliquer la théorie des perturbations stationnaires en traitant

³⁰Un tel état est dégénéré avec des états de continuum et se dissocie spontanément ; de ce fait, il n'a aucune pertinence pour la description des états liés.

$V = \frac{e'^2}{r_{12}}$ comme une perturbation. On verra après coup que la répulsion est loin d'être une petite correction (ce qui n'est guère surprenant physiquement), mais la prise en compte, même grossière, de la répulsion de Coulomb permet de revenir sur le phénomène surprenant (et crucial) déjà rencontré : en dépit de l'absence d'interactions magnétiques, les états de l'atome ont, grâce au Principe de Pauli, une énergie qui dépend explicitement de la valeur du spin total de l'atome ("magnétisme sans magnétisme"). De toute façon, en pratique et quand il le faut, on peut toujours raffiner le calcul afin d'obtenir une description numériquement plus convenable.

En ce qui concerne l'état fondamental (non-dégénéré), la correction d'énergie au premier ordre est simplement :

$$\Delta E_{1^1S} = \langle \Psi_{1^1S} | \frac{e'^2}{r_{12}} | \Psi_{1^1S} \rangle . \quad (3.82)$$

L'opérateur de perturbation ne dépend pas du spin ; après sommation sur les variables de spin, on obtient une intégrale portant sur les seules variables orbitales :

$$\Delta E_{1^1S} = \langle \phi_{100}\phi_{100} | \frac{e'^2}{r_{12}} | \phi_{100}\phi_{100} \rangle , \quad (3.83)$$

où l'élément de matrice est l'intégrale :

$$\int d^3r_1 \int d^3r_2 \phi_{100}(r_1)\phi_{100}(r_2) \frac{e'^2}{r_{12}} \phi_{100}(r_1)\phi_{100}(r_2) \equiv J_{1s1s} . \quad (3.84)$$

L'intégrale J_{1s1s} s'appelle intégrale directe³¹ et représente la répulsion électrostatique des deux distributions de charges $\phi_{100}(\vec{r}_1)$ et $\phi_{100}(\vec{r}_2)$; elle vaut³² $\frac{5}{4} \frac{e'^2}{a_0} \simeq 34.01$ eV [7] ; c'est visiblement une quantité positive – c'est également une correction très importante ! À cet ordre, l'énergie du fondamental est :

$$E_{1^1S} = 2E_1 + J_{1s1s} = -74.89 \text{ eV} . \quad (3.85)$$

La répulsion de Coulomb entre électrons déstabilise considérablement le fondamental, ce qui n'est pas surprenant.

Le premier niveau excité d'énergie $E_1 + E_2$ est dégénéré 16 fois ; il convient donc d'écrire la matrice de la perturbation dans le sous-espace dégénéré et de la diagonaliser. En fait, la matrice est déjà diagonale car $V = \frac{e'^2}{r_{12}}$ est invariant par rotation des variables d'espace et ne dépend pas du spin ; il en résulte :

$$[V, \vec{L}] = 0 \quad [V, \vec{S}] = 0 , \quad (3.86)$$

de sorte que V est un opérateur scalaire vis-à-vis de l'espace et du spin séparément. Pour deux états $\Psi_{LSM_L M_S}$ on a donc nécessairement :

$$\langle \Psi_{LSM_L M_S} | \frac{e'^2}{r_{12}} | \Psi_{L'S'M'_L M'_S} \rangle \propto \delta_{LL'} \delta_{SS'} \delta_{M_L M'_L} \delta_{M_S M'_S} , \quad (3.87)$$

et ceci montre que la matrice de la perturbation est d'emblée diagonale ; les corrections à l'énergie au premier ordre sont simplement données par ces éléments diagonaux, qui vont différer suivant les valeurs de L et S : la répulsion de Coulomb va lever partiellement la dégénérescence présente dans l'approximation grossière des électrons indépendants.

En ce qui concerne l'état singulet 2^1S , la correction $\Delta E_{2^1S} = \langle \Psi_{2^1S} | \frac{e'^2}{r_{12}} | \Psi_{2^1S} \rangle$, après sommation sur les variables de spin, prend la forme :

$$\Delta E_{2^1S} = \frac{1}{2} \langle \phi_{100}\phi_{200} + \phi_{200}\phi_{100} | \frac{e'^2}{r_{12}} | \phi_{100}\phi_{200} + \phi_{200}\phi_{100} \rangle , \quad (3.88)$$

³¹ Attention ! Ceci est la terminologie usuelle en physique atomique et moléculaire (voir par exemple [14]). En revanche, Cohen-Tannoudji *et al.* [8] notent K l'intégrale directe et J l'intégrale d'échange (traditionnellement, une intégrale d'échange est notée J en physique de la matière condensée).

³² Les intégrales du genre J et K se calculent en développant $\frac{1}{r_{12}}$ en harmoniques sphériques, ce qui sépare les variables ; voir [1], Appendice B.

soit :

$$\Delta E_{2^1S} = \langle \phi_{100}\phi_{200} | \frac{e'^2}{r_{12}} | \phi_{100}\phi_{200} \rangle + \langle \phi_{100}\phi_{200} | \frac{e'^2}{r_{12}} | \phi_{200}\phi_{100} \rangle \equiv J_{1s2s} + K_{1s2s} . \quad (3.89)$$

On retrouve une intégrale du type direct, J_{1s2s} (répulsion électrostatique entre les deux distributions $|\phi_{100}|^2$ et $|\phi_{200}|^2$, qui est égale à $\frac{34}{81}(\frac{e'^2}{a_0})$, soit environ 11.42 eV), mais apparaît maintenant en plus une autre intégrale, K_{1s2s} , appelée intégrale d'échange³³, égale³⁴ ici à $\frac{32}{729}\frac{e'^2}{a_0} \simeq 1.19$ eV . Cela ne saute pas aux yeux mais il est possible de montrer qu'une intégrale d'échange³⁵ est toujours positive [15] :

$$K_{nm} = \int d^3r_1 \int d^3r_2 \phi_n^*(\vec{r}_1)\phi_m^*(\vec{r}_2) \frac{e'^2}{r_{12}} \phi_m(\vec{r}_1)\phi_n(\vec{r}_2) > 0 \quad \forall \phi_n, \phi_m . \quad (3.90)$$

Noter qu'une telle expression ne peut recevoir d'interprétation classique en termes d'interaction entre deux distributions continues de charge : les produits $\phi_n^*(\vec{r}_1)\phi_m(\vec{r}_1)$ et $\phi_n^*(\vec{r}_2)\phi_m(\vec{r}_2)$ n'ont aucune raison d'être positifs.

À cet ordre, on a maintenant :

$$E_{2^1S} = E_1 + E_2 + J_{1s2s} + K_{1s2s} \simeq -55.45 \text{ eV} . \quad (3.91)$$

Le calcul se fait de la même façon pour l'état triplet 1^3S ; compte tenu du changement de signe dans la partie d'espace des fonctions d'onde entre 2^1S et 1^3S , on trouve maintenant :

$$E_{1^3S} = E_1 + E_2 + J_{1s2s} - K_{1s2s} \simeq -57.84 \text{ eV} , \quad (3.92)$$

et, compte tenu du signe de l'intégrale d'échange, (3.90), on obtient l'inégalité importante :

$$E_{1^3S} < E_{2^1S} , \quad E_{2^1S} - E_{1^3S} = 2K_{1s2s} > 0 . \quad (3.93)$$

qui est la traduction énergétique du trou de Fermi. Bien qu'il n'y ait aucune interaction magnétique, l'énergie dépend donc de la valeur du spin total ; notamment, l'énergie du triplet est inférieure à l'énergie du singulet, toutes choses égales par ailleurs. La dépendance de l'énergie par rapport au spin est un fait majeur, autant sur le plan qualitatif que quantitatif : elle rend compte fondamentalement des propriétés magnétiques de la matière. Le recours à des interactions magnétiques "classiques" non seulement ne rendrait pas compte de tous les phénomènes observés mais de plus serait bien incapable d'expliquer les ordres de grandeur.

Pour les états P on trouve de façon analogue³⁶ :

$$E_{1^1P} = E_1 + E_2 + J_{1s2p} + K_{1s2p} \simeq -53.91 \text{ eV} , \quad (3.94)$$

$$E_{1^3P} = E_1 + E_2 + J_{1s2p} - K_{1s2p} \simeq -55.77 \text{ eV} , \quad (3.95)$$

avec à nouveau :

$$E_{1^3P} < E_{1^1P} , \quad E_{1^1P} - E_{1^3P} = K_{1s2p} > 0 . \quad (3.96)$$

En définitive, la succession des niveaux classés par énergie croissante est, dans l'approximation en cours :

$$E_{1^1S} < E_{1^3S} < E_{1^3P} < E_{2^1S} < E_{1^1P} . \quad (3.97)$$

Au total, l'approximation assez grossière³⁷ donne des résultats moyennement satisfaisants. S'il est heureux de voir levée la dégénérescence d'échange et le bon ordonnancement des états singulets et triplets, en revanche

³³L'apparition du terme d'échange est clairement une conséquence du postulat de symétrisation et n'a pas d'interprétation classique.

³⁴On note un bon facteur 10 entre les intégrales directe et d'échange.

³⁵L'origine de la terminologie est la suivante : en imaginant que l'on peut préparer l'atome initialement dans l'état $\phi_n(\vec{r}_1)\phi_m(\vec{r}_2)$, alors celui-ci oscille entre cet état et $\phi_m(\vec{r}_1)\phi_n(\vec{r}_2)$ à la fréquence $\hbar^{-1}K_{nm}$. On peut aussi dire que $\hbar^{-1}K_{nm}$ est la fréquence à laquelle "les électrons échangent leurs nombres quantiques". Pour la configuration (1s)(2p), l'échange a lieu toutes les 10^{-14} s ; en revanche, pour la configuration (1s)(n=10, l=9), le temps d'échange est égal à ... 1600 ans, alors que le diamètre de l'orbite excitée est seulement de l'ordre de $100a_0$ [14]. K provient du postulat de symétrisation : la symétrisation est donc inutile si les fonctions d'onde ne se recouvrent pas [11].

³⁶ $J_{1s2p} \simeq 13.22$ eV, $K_{1s2p} \simeq 0.93$ eV : il y a à nouveau un gros facteur 10 entre les termes direct et d'échange.

³⁷On traite par perturbation et au premier ordre les effets d'un opérateur dont les valeurs moyennes ne sont pas très petites vis-à-vis des écarts de niveaux du problème non perturbé.

l'ordre des niveaux n'est pas conforme à l'expérience ; du spectre observé, on peut en effet conclure que l'ordre réel est le suivant :

$$E_{1^1S} < E_{1^3S} < E_{2^1S} < E_{1^3P} < E_{1^1P} . \quad (3.98)$$

Le *splitting* entre 1^3S et 2^1S vaut environ 0.8 eV, celui entre 1^3P et 1^1P vaut 0.25 eV. L'écart entre E_{2^1S} et E_{1^3P} est égal à 0.34 eV.

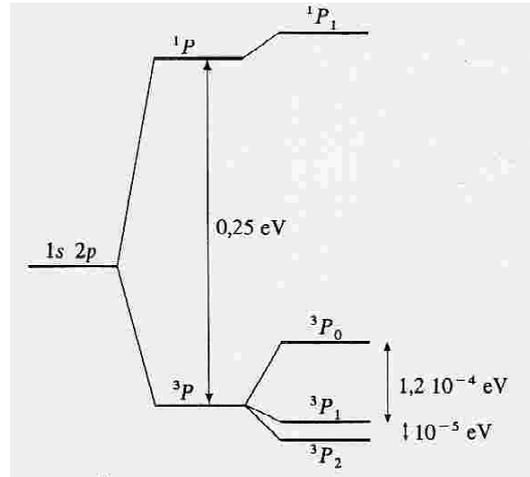


Figure 3.1: Structure fine de He pour les multiplets P issus de (1s)(2p) (figure extraite de [8]).

La prise en compte de la structure fine se fait suivant la méthode décrite dans la section 3.2, voir notamment (3.27). Numériquement, l'accord n'est pas très bon, tout particulièrement pour les termes 3P_2 et 3P_1 : l'interaction dipolaire magnétique spin - spin (négligée) y joue un rôle important. L'étude détaillée de la structure fine du multiplet 3P présente un grand intérêt sur le plan fondamental : elle est un cas d'école commode pour la détermination précise de la constante de structure fine α .

Remarque

Pour un système à $N = 2$ électrons, le Principe de Pauli produit une levée de dégénérescence entre les états triplet et singulet. Ceci peut être représenté par un opérateur effectif³⁸ du genre $V_{\text{eff}} = -A\vec{S}_1 \cdot \vec{S}_2$, où A est une constante positive et $S_i = \frac{1}{2}$. En effet, on a :

$$\vec{S}_1 \cdot \vec{S}_2 = \frac{1}{2}(\vec{S}^2 - \vec{S}_1^2 - \vec{S}_2^2) = \frac{1}{2}(\vec{S}^2 - \frac{3}{4}\hbar^2) , \quad (3.99)$$

d'où résulte :

$$\langle S = 0 | V_{\text{eff}} | S = 0 \rangle - \langle S = 1 | V_{\text{eff}} | S = 1 \rangle = A\hbar^2 \quad (3.100)$$

La constante A est donc $2\hbar^{-2}K$. Cette idée est à la base des Hamiltoniens-modèles souvent rencontrés en Physique de la matière condensée (modèle de Heisenberg, par exemple).

³⁸Bien évidemment, une telle formulation ne doit pas laisser croire qu'il s'agit d'interactions magnétiques au sens classique du terme : à nouveau, l'ordre de grandeur énergétique n'y serait pas. Il s'agit tout juste d'une représentation effective, en terme d'opérateurs, du *splitting* singulet - triplet obtenu plus haut. Dans le même ordre d'idée, rien n'interdit de représenter un atome à deux niveaux dans un langage de *spin fictif*, $S = \frac{1}{2}$.

Chapitre 4

Illustration des postulats de la Mécanique Quantique

Il s'agit de présenter quelques illustrations simples des postulats en décrivant notamment quelques expériences récentes¹. Trois sujets seront abordés ; deux d'entre eux (effet Zenon et Cryptographie) discutent d'idées récemment introduites. Au contraire, le troisième (sauts quantiques) repose sur un concept énoncé par Bohr (1913) qui n'a pu être observé que récemment en raison de la technologie expérimentale qu'il requiert pour sa mise en évidence.

4.1 L'effet Zenon quantique

L'effet Zenon² quantique a été précisément formulé en 1977 [17], mais certains travaux antérieurs avaient déjà posé ce qui paraît à première vue constituer un paradoxe : on va voir qu'un système quantique soumis à une succession de mesures très rapprochées n'évolue pas dans le temps³. L'idée est facile à comprendre. Toute opération de mesure projette dans l'état propre normalisé de l'observable mesurée, associé à la valeur propre obtenue en tant que résultat de la mesure effectuée. La vitesse d'évolution du vecteur d'état étant finie, la probabilité d'obtenir la même valeur lors d'une deuxième mesure rapprochée dans le temps est relativement grande – en quelque sorte, la deuxième mesure “freine” l'évolution. Si on procède à une série de mesures très voisines en temps, on conçoit que se produise un ralentissement de la dynamique. L'argument qui suit montre, dans un cas d'école, que dans la limite d'un nombre infini de mesures effectuées entre deux temps fixés t_i et t_f , la probabilité pour le système d'être à l'instant t_f dans le *même* état qu'à l'instant $t_i \dots$ est égale à 1.

De façon plus précise, cet effet “paradoxal” peut se formuler comme suit. Soit un système quantique, de Hamiltonien H , initialement ($t_i = 0$) dans l'état $|\psi_0\rangle$, état supposé instable pour fixer les idées⁴. Le projecteur :

$$P_0 = |\psi_0\rangle\langle\psi_0| \quad (4.1)$$

¹D'autres questions mériteraient d'être discutées, comme par exemple le paradoxe EPR ou l'ordinateur quantique. Des expériences assez récentes ont permis de vérifier que les prévisions quantiques concernant les corrélations à distance – le point litigieux selon Einstein, Podolsky et Rosen – existent réellement ; quant à l'ordinateur quantique, il semble, selon les experts en la matière, que si les idées de base sont éminemment séduisantes, il y a loin de la théorie à la pratique. Aujourd'hui, l'ordinateur quantique se situe plutôt dans l'utopie.

²Le paradoxe de Zenon “classique” repose sur la négation *a priori* de la possibilité de diviser à l'infini l'espace et le temps. Il en résulte des affirmations paradoxales tendant toutes à montrer que le mouvement est impossible (voir par exemple [16]). Il est intéressant de noter que Schrödinger revint dans les années 30 sur la question de pouvoir représenter l'espace et le temps par des grandeurs continues.

³“*It is common knowledge that a watched kettle never boils.*” [18].

⁴Cette hypothèse n'est pas essentielle, mais elle en rajoute au spectaculaire : une particule instable que l'on observe en continu est stable !

est un opérateur hermitique, dont les valeurs propres ϖ_λ sont 0 et 1 :

$$\varpi_1 = 1 : P_0 |\psi_0\rangle = |\psi_0\rangle ; \quad \varpi_0 = 0 : P_0 |\phi\rangle = 0 \quad \forall |\phi\rangle \perp |\psi_0\rangle . \quad (4.2)$$

La valeur propre 0 est en général fortement dégénérée (sauf si l'espace des états est de dimension égale à 2). En tout cas, P_0 est un opérateur dont toutes les valeurs propres sont réelles et dont l'ensemble des vecteurs propres constitue une base complète de \mathcal{E} : c'est donc une observable au sens de la Mécanique Quantique. De ce fait, il est licite d'appliquer les postulats concernant les résultats d'une opération de mesure effectuée sur un système dans l'état $|\Psi\rangle$ donné :

- le résultat d'une mesure est l'une des valeurs propres ϖ_λ de P_0
- la probabilité d'obtenir le résultat ϖ_λ est le module au carré de la composante de $|\Psi\rangle$ sur le vecteur propre correspondant, $|\varpi_\lambda\rangle$
- juste après la mesure dont le résultat est $\varpi_1 = 1$, le système se trouve dans l'état $|\varpi_1 = 1\rangle \equiv |\psi_0\rangle$ (réduction du paquet d'ondes en l'absence de dégénérescence) ; si le résultat est $\varpi_0 = 0$, le système est, juste après cette mesure, dans l'état (normalisé) qui est la projection de l'état juste avant la mesure sur le sous-espace $\{|\varpi_\lambda = 0\rangle\}$ (réduction du paquet d'ondes en présence de dégénérescence).

Ce postulat a d'étranges conséquences (le *Chat de Schrödinger* !). Il est toutefois une nécessité logique pour que la théorie quantique ait un sens physique. En effet, si la mesure à l'instant t_0 d'une certaine observable A sur le système dans l'état $|\Psi\rangle$ a produit la valeur non-dégénérée a_0 , une deuxième mesure effectuée immédiatement après doit évidemment redonner la *même* valeur (avec probabilité un). Ceci n'est possible que si à l'issue de la première mesure le système est dans l'état $|a_0\rangle$, propre de A avec la valeur propre a_0 ; la première mesure a donc bien effectué la projection :

$$|\Psi(t_0)\rangle \rightarrow |a_0\rangle \equiv |\Psi(t_0 + 0)\rangle . \quad (4.3)$$

Effectuons une mesure unique de P_0 à l'instant t , le système étant parti à $t_i = 0$ de l'état $|\psi_0\rangle$; le descendant de l'état de départ est $U(t)|\psi_0\rangle$ où $U(t) = e^{\frac{i}{\hbar}Ht}$. Le résultat de cette mesure est donc :

$$1 \text{ avec la probabilité } p_1 = |\langle \varpi_1 | U(t) | \psi_0 \rangle|^2 \quad (4.4)$$

$$0 \text{ avec la probabilité } p_0 = 1 - p_1 . \quad (4.5)$$

Quand on a trouvé 1, on convient de dire que le résultat de la mesure est positif. $|\psi_0\rangle$ étant l'unique vecteur propre associé à la valeur propre 1, $|\psi_0\rangle \equiv |\varpi_1\rangle$, et (4.4) s'écrit aussi :

$$p_1(t) = |\langle \psi_0 | U(t) | \psi_0 \rangle|^2 . \quad (4.6)$$

Supposons maintenant que l'instant t de mesure soit très petit comparé aux inverses de toutes les fréquences de Bohr ; alors, toutes les exponentielles apparaissant dans le développement de $U(t)|\psi_0\rangle$ sur les états propres de H peuvent être développées, et l'amplitude de transition est approximativement donnée par :

$$A_{0 \rightarrow 0}(t) \equiv \langle \psi_0 | U(t) | \psi_0 \rangle \simeq \langle \psi_0 | 1 + \frac{Ht}{i\hbar} + \frac{1}{2!} \left(\frac{Ht}{i\hbar} \right)^2 | \psi_0 \rangle + \mathcal{O}(t^3) . \quad (4.7)$$

Dans ces conditions :

$$p_1(t) = |A_{0 \rightarrow 0}(t)|^2 \simeq \left(1 - \frac{t^2}{2\hbar^2} \langle H^2 \rangle \right)^2 + \frac{t^2}{\hbar^2} \langle H \rangle^2 \simeq 1 - \hbar^{-2} \Delta H^2 t^2 + \mathcal{O}(t^4) . \quad (4.8)$$

Toutes les moyennes $\langle \dots \rangle$ sont prises avec l'état $|\psi_0\rangle$, ΔH^2 est l'écart quadratique⁵ de l'énergie dans l'état $|\psi_0\rangle$. À la réflexion, $p_1(t)$ est en fait une probabilité *conditionnelle* : dire que l'état de départ est $|\psi_0\rangle$ est équivalent

⁵Tout ce qui suit suppose évidemment que la variance ΔH de l'énergie est *finie*.

à affirmer que l'on a fait juste avant une mesure qui a produit le résultat $\varpi_\lambda = 1$; avec la notation habituelle pour une probabilité conditionnelle, on peut donc écrire :

$$p_1(t) \equiv p(t|0) , \quad (4.9)$$

et $p(t)$ est bien la probabilité de trouver 1 à l'instant t sachant que l'on a trouvé 1 à $t = 0$. En disant les choses autrement : supposons que l'on fasse deux mesures successives de l'observable P_0 aux deux instants t_1 et $t_2 > t_1$. Si on trouve la valeur 1 en t_1 (mesure positive), alors, l'état du système à $t_1 + 0$ est encore $|\psi_0\rangle$; l'état à l'instant t_2 est donc $U(t_2 - t_1)|\psi_0\rangle$. Dans ces conditions, la probabilité de trouver à nouveau la valeur propre 1 à l'instant t_2 sachant que l'on a trouvé 1 à l'instant t_1 est bien :

$$p(t_2|t_1) = p_1(t_2 - t_1) \quad (4.10)$$

soit approximativement, d'après (4.8) :

$$p(t_2|t_1) \simeq 1 - \hbar^{-2} \Delta H^2 (t_2 - t_1)^2 . \quad (4.11)$$

Cette probabilité ne dépend que de la différence des temps, puisque H est depuis le début supposé indépendant du temps.

Effectuons maintenant une série de N mesures de P_0 aux instants successifs $t_n = n\Delta t$, $n = 1, 2, \dots, N$, où $\Delta t = \frac{t}{N}$, t et N étant fixés. À l'issue de cette série, la probabilité d'avoir trouvé à chaque fois la valeur propre 1 (toutes les mesures donnent un résultat positif) est⁶ :

$$p_{\text{pos}}(t_1, t_2, \dots, t_N) = p(t_N|t_{N-1})p(t_{N-1}|t_{N-2}) \dots p_1(t_2|t_1) \simeq \left[1 - \hbar^{-2} \Delta H^2 \frac{t^2}{N^2} \right]^N . \quad (4.12)$$

En utilisant maintenant :

$$\left(1 - \frac{C}{N^2} \right)^N = e^{N \ln(1 - \frac{C}{N^2})} \simeq e^{-\frac{C}{N}} \simeq 1 \quad (N \gg 1) \quad (4.13)$$

on voit que la probabilité p_{pos} que toutes les mesures donnent un résultat positif tend vers 1 quand le nombre de mesures tend vers l'infini :

$$\lim_{N \rightarrow \infty} p_{\text{pos}}(t_1, t_2, \dots, t_N) = 1 \quad (4.14)$$

Autrement dit, un système observé (mesuré) en "continu" n'évolue pas.

Remarques

1. La probabilité de trouver toujours la même valeur (1 en l'occurrence) lors de la mesure de la "variable binaire" associée à P_0 tend donc vers 1. Ceci est complètement différent de la probabilité de trouver toujours pile lors du jet d'une pièce, effectué un grand nombre de fois, N : cette dernière vaut 2^{-N} et tend vers zéro exponentiellement. Fondamentalement, la différence tient à la corrélation statistique des événements, forte pour l'effet Zenon, nulle (ou supposée négligeable) pour le jet de la pièce. Pour l'effet Zenon quantique, plus on raccourcit l'intervalle de temps entre deux mesures successives, plus grande est la probabilité de retrouver la même valeur que lors de la mesure immédiatement précédente. Pour le jet de la pièce, tout nouveau lancement est supposé complètement indépendant de tous les précédents.
2. L'instabilité supposée de l'état $|\psi_0\rangle$ ne joue aucun rôle et n'a été adoptée que pour fixer les idées. La même conclusion peut être tirée pour un système à deux niveaux séparés par l'énergie $\hbar\omega_0$, dont la dynamique libre est périodique de période $\frac{2\pi}{\omega_0}$; en la circonstance, la seule condition à retenir est $\Delta t \ll \omega_0^{-1}$.
3. La conclusion exprimée par (4.14) repose crucialement sur le fait que la probabilité $p_1(t)$ s'écarte quadratiquement de la valeur 1 (voir (4.8)). Si la variation était linéaire (par exemple, le début d'une exponentielle décroissante), $p_1(t) \simeq (1 - \Gamma t)$, on aurait :

$$p_{\text{pos}}(t_1, t_2, \dots, t_N) \simeq \left(1 - \Gamma \frac{t}{N} \right)^N \rightarrow e^{-\Gamma t} . \quad (4.15)$$

⁶La probabilité des N résultats positifs est le produit des probabilités de transition $p(t_{n+1}|t_n)$: il s'agit donc d'une *chaîne de Markov*.

Plus généralement, avec un comportement du type :

$$p_1(t) \simeq (1 - Ct^\alpha) , \quad (4.16)$$

on obtient :

$$p_{\text{Pos}}(t_1, t_2, \dots, t_N) \simeq \left[1 - C \left(\frac{t}{N} \right)^\alpha \right]^N \simeq e^{-Ct^\alpha N^{1-\alpha}} . \quad (4.17)$$

Dès que $\alpha > 1$, il y a effet Zenon : c'est bien quand le système s'éloigne "lentement" de son point de départ que celui-là peut survenir. Si on considère un atome à deux niveaux $|e\rangle$ (excité, radiatif) et $|g\rangle$ (fondamental), la première phase de déclin de l'état excité est exponentielle⁷ sauf sur un très petit intervalle, où le comportement – parabolique comme en (4.8) – permet l'effet Zenon ; ce régime étant de courte durée, il semble donc très difficile d'observer l'effet Zenon dans une configuration aussi simple.

L'argument élémentaire ci-dessus, conduisant à la mise en évidence de l'effet Zenon, est fondé sur la notion idéale que l'on se fait de l'opération de mesure en Mécanique Quantique, caractérisée notamment par la réduction instantanée du paquet d'ondes, et sans se risquer à introduire "l'appareil de mesure" dans le formalisme (le peut-on, d'ailleurs ?). La difficulté d'observation de l'effet est liée à la définition d'un scénario permettant de réaliser expérimentalement une suite de processus singeant aussi bien que possible l'opération de mesure au sens de la Mécanique Quantique.

Le ralentissement de la dynamique quantique peut aussi être mis en évidence dans un petit modèle que l'on peut énoncer innocemment, sans référence initiale au concept de mesure.

Considérons un système à petit nombre de niveaux, couplé à un grand système, ce dernier étant caractérisé par le fait qu'il possède un très grand nombre de degrés de liberté ; cette propriété se traduit usuellement par le fait que le spectre d'énergie du grand système forme un quasi-continuum. Dès lors, il existe un grand nombre de couples de fréquences dont la différence est très petite, associés à des périodes très longues ; au total, les temps de récurrence (ou de pseudo-récurrences) deviennent gigantesques au point de perdre tout sens pour le physicien : l'irréversibilité de la dynamique du petit système couplé au grand devient un état de fait.

Il est tout à fait possible de bâtir un formalisme montrant ceci en détail. Il apparaît alors que très souvent, notamment lorsque le couplage entre les deux systèmes est faible en un sens à préciser, des approximations légitimes permettent d'établir que sur un grand intervalle de temps, la dynamique oscillante est amortie suivant un simple régime exponentiel. Cette conclusion repose sur le fait que le couplage avec le grand système décale les pôles de la résolvante⁸ de l'axe réel vers le demi-plan complexe inférieur (plus précisément dans le deuxième feuillet de Riemann), fabriquant ce que l'on appelle dans d'autres contextes une *résonance*, d'autant plus fine que le temps de relaxation du petit système est long. Si $\hbar\Gamma$ est largeur de la résonance, le temps de relaxation (ou durée de vie de l'état instable) est $\sim \Gamma^{-1}$. Outre l'apparition de la largeur, la partie réelle du pôle est (un peu) différente de l'énergie "nue" : on parle parfois de *renormalisation* de l'énergie, et c'est ainsi que l'on peut rendre compte du *Lamb shift* d'un atome (dans ce dernier cas, le grand système n'est autre que le champ électromagnétique quantifié dans une grande boîte – c'est pourquoi on parle alors de *déplacement radiatif*).

D'un autre point de vue, ce comportement exponentiel peut être d'emblée fabriqué de façon purement phénoménologique en incorporant des parties imaginaires (négatives) dans les énergies, ce qui revient peu ou prou à introduire un Hamiltonien *ad hoc* non hermitique. Cette recette, dont on il est certes intéressant de discuter les fondements, permet de faire l'économie de la description quantique du grand système, forcément un peu laborieuse, et c'est elle que l'on applique maintenant.

Imaginons un atome à trois niveaux $|g\rangle$, $|m\rangle$ et $|e\rangle$. $|g\rangle$ est l'état fondamental, $|m\rangle$ un état excité métastable dont la désexcitation radiative vers le fondamental a une probabilité très faible, que l'on néglige dans la suite. Enfin, $|e\rangle$ est un état fortement couplé au fondamental par échange de photons avec l'extérieur. L'ordre des

⁷Toutefois, aux temps très longs, le déclin suit une loi puissance $\sim t^{-\lambda}$, qui est *relativement* dominante une fois l'exponentielle éteinte.

⁸Essentiellement, la résolvante est la transformée de Laplace (ou de Fourier) de l'opérateur d'évolution $U(t)$. C'est un avatar de fonction de Green, dont les singularités dans \mathbb{C} donnent le spectre de H et déterminent donc la dynamique (pour plus de détails, voir [11]). En particulier, quand le spectre de H contient une composante continue, la résolvante possède des points de branchement ; la (ou les) coupure correspondante définit la frontière commune des différents feuillets constituant la surface de Riemann.

énergies est $E_g < E_m < E_e$. La fuite radiative à partir de $|e\rangle$ vers le fondamental va être pragmatiquement représentée par une partie imaginaire $-i\hbar\frac{\Gamma_e}{2}$ ajoutée à l'énergie de l'état $|e\rangle$ ($\Gamma_e > 0$) et, dans la suite, on considère exclusivement les deux niveaux $|m\rangle$ et $|e\rangle$, supposés satisfaire les équations suivantes :

$$H_{at}|m\rangle = \hbar\omega_m|m\rangle \quad H_{at}|e\rangle = \hbar(\omega_e - i\frac{\Gamma_e}{2})|e\rangle . \quad (4.18)$$

Supposons maintenant cet atome perturbé par un champ externe (semi-classique) harmonique résonnant représenté par l'opérateur⁹ $V(t)$:

$$V(t) = \hbar\Omega \left[|e\rangle\langle m| e^{i(\omega_m - \omega_e)t} + |m\rangle\langle e| e^{i(\omega_e - \omega_m)t} \right] . \quad (4.19)$$

Ceci étant posé, on écrit le vecteur d'état sous la forme :

$$|\Psi(t)\rangle = a_m(t)e^{-i\omega_m t}|m\rangle + a_e(t)e^{-i\omega_e t}|e\rangle \quad (4.20)$$

et l'équation de Schrödinger (effective) pour les coefficients s'écrit :

$$\dot{a}_m(t) = -i\Omega a_e , \quad \dot{a}_e(t) = -i\Omega a_m - \frac{\Gamma_e}{2} a_e(t) . \quad (4.21)$$

L'intégration est immédiate¹⁰ ; supposant l'atome initialement dans l'état $|m\rangle$, on trouve :

$$a_m(t) = e^{-\frac{1}{4}\Gamma_e t} \left(\cos \tilde{\Omega} t + \frac{\Gamma_e}{4\tilde{\Omega}} \sin \tilde{\Omega} t \right) , \quad a_e(t) = -i \frac{\Omega}{\tilde{\Omega}} e^{-\frac{1}{4}\Gamma_e t} \sin \tilde{\Omega} t , \quad (4.22)$$

avec :

$$\tilde{\Omega} = \sqrt{\Omega^2 - \frac{\Gamma_e^2}{16}} . \quad (4.23)$$

L'oscillation de Rabi correspond à $\frac{\Gamma_e}{\Omega} \ll 1$; l'atome effectue alors un va-et-vient sous-amorti de l'état $|m\rangle$ à l'état $|e\rangle$ à la pulsation $\simeq \Omega$:

$$P_m(t) \equiv |a_m(t)|^2 \simeq e^{-\frac{1}{2}\Gamma_e t} \cos^2 \Omega t \quad (\Gamma_e \ll \Omega) . \quad (4.24)$$

Au contraire, si $\frac{\Gamma_e}{\Omega} \gg 1$:

$$P_m(t) \simeq e^{-4(\Omega/\Gamma_e)^2 \Gamma_e t} \quad (\Omega \ll \Gamma_e) \quad (4.25)$$

et l'atome est alors bloqué très longtemps (relativement à l'échelle de temps Γ_e^{-1}) dans l'état métastable, son état initial. Dans la limite $\frac{\Omega}{\Gamma_e} \rightarrow 0$, $P_m(t) \rightarrow 1$, $\forall t$. En définitive, ce modèle très simple montre que la fuite rapide d'énergie vers l'extérieur (*via* l'état excité $|e\rangle$) inhibe la transition induite par le champ résonnant provoquant l'oscillation de Rabi.

La parenté avec l'effet Zenon se situe au niveau de la dynamique, qui est ralentie de façon spectaculaire. En outre, l'évacuation rapide d'énergie, quasi-instantanée quand on la rapporte à l'échelle propre du système Ω^{-1} , est un processus physique de durée très courte à destination du milieu extérieur que l'on peut considérer comme l'un de ces mécanismes qui se produisent quoi qu'il arrive quand un système microscopique interagit avec un appareil de mesure. Le modèle ci-dessus ne fait à aucun moment référence à une réduction de paquet d'ondes – la dynamique est hamiltonienne (quoique le Hamiltonien effectif est non-hermitique) – mais on sent néanmoins que la parenté ne se réduit pas au simple ralentissement, même s'il semble difficile de pousser davantage la comparaison tout en conservant la définition première de l'effet Zenon.

Cook [19] a proposé une expérience destinée à mettre en évidence cet effet, par une inhibition de l'absorption induite dans un atome à trois niveaux, du type de celui qui vient d'être introduit. Par opposition avec ce qui précède, c'est maintenant le couple ($|g\rangle$, $|m\rangle$) qui est sollicité par une excitation résonnante de pulsation ω et participe à l'oscillation de Rabi (voir fig. 4.1) ; la transition $|g\rangle \longleftrightarrow |m\rangle$, d'énergie $\hbar\omega_0$, est dans le domaine des radiofréquences (quelques centaines de MHz), l'oscillation de Rabi se produit à une fréquence de l'ordre de quelques dizaines de Hertz ou même moins. Si on peut effectuer une mesure de l'état du système durant cette oscillation, on trouve soit $|g\rangle$, soit $|m\rangle$ – l'état $|e\rangle$ étant pour l'instant hors-circuit – avec des probabilités connues, dépendant de la position de l'instant de mesure par rapport au cycle de Rabi (et aussi de l'écart¹¹ à la

⁹Le champ donnant lieu au couplage $V(t)$ avec l'atome est bien *résonnant*, puisqu'il oscille précisément à la pulsation de Bohr $\omega_e - \omega_m$ de l'atome.

¹⁰Utiliser par exemple une transformation de Laplace.

¹¹C'est seulement quand cet écart est nul que les probabilités P_g et P_m varient harmoniquement (à la pulsation Ω) entre 0 et 1.

résonance $\omega - \omega_0$). En définitive, l'opération de mesure – à définir plus précisément – révèle soit l'état $|g\rangle$, soit l'état $|m\rangle$.

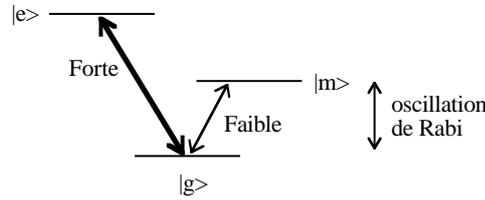


Figure 4.1: Atome à trois niveaux pour l'effet Zenon.

L'état $|e\rangle$ est supposé être couplé exclusivement au fondamental $|g\rangle$ par l'intermédiaire du champ électromagnétique : un *pulse* optique peut donc provoquer des transitions entre $|g\rangle$ et $|e\rangle$. $|e\rangle$ est un état excité "ordinaire", avec une durée de vie très courte, 10^{-10} s pour fixer les idées.

L'atome est donc soumis d'une part à l'excitation résonnante provoquant l'oscillation de Rabi, d'autre part à un *pulse* optique de durée très courte, ou plusieurs *pulses* successifs, et chacun de ceux-ci provoque le *collapse* de la fonction d'onde. En effet, si l'atome est dans l'état $|g\rangle$ quand arrive un tel *pulse*, des cycles d'absorption-émission spontanée de photons se produisent, se manifestant par une diffusion de photons dans une direction aléatoire par rapport à la direction incidente¹² ; si au contraire l'atome est dans l'état $|m\rangle$, aucune diffusion de photon ne se produit. Ainsi, l'absence de photons diffusés révèle que l'état de l'atome est l'état $|m\rangle$ et produit donc le *collapse* de la fonction d'onde en $|m\rangle$; au contraire, la présence de photons diffusés produit le *collapse* en $|g\rangle$. Pour parler de façon imagée, l'observation ou non de photons diffusés joue le rôle de l'aiguille de l'ampèremètre.

En outre, si on "trouve" $|m\rangle$ au sens ci-dessus, l'atome reste en $|m\rangle$ durant tout le *pulse* et un deuxième *pulse* rapproché en temps révèle à nouveau le même état. Si au contraire on "trouve" $|g\rangle$, l'atome retourne très vite dans le même état $|g\rangle$, puisque la durée de vie Γ_e^{-1} est très petite.

En définitive, le *pulse* optique constitue bien une mesure de l'état au sens où l'entend la Mécanique Quantique, à un petit détail près : la durée finie du *pulse* et le séjour éphémère de l'atome dans l'état $|e\rangle$ – alors que le processus de mesure, tel qu'il est ordinairement décrit, affirme l'instantanéité de la réduction du paquet d'ondes. Comme toujours, l'essentiel est que l'instantanéité supposée repose sur la considération de deux échelles de temps, la plus petite étant très petite – mais évidemment jamais strictement nulle. En tout cas, ce dispositif réalise presque parfaitement ce que la Mécanique Quantique appelle une mesure idéale, avec toutes ses conséquences y compris la réduction du paquet d'ondes, dans le cas où il y a deux résultats possibles : $|g\rangle$ ou $|m\rangle$. Un tel atome se prête donc *a priori* à l'observation de l'effet Zenon.

Itano *et al.* [20] ont d'une part réalisé cette expérience, d'autre part en ont donné la description théorique ; l'accord entre théorie et expérience est remarquable, justifiant que leur travail soit considéré comme la première mise en évidence de l'effet Zenon quantique.

L'expérience a été menée sur une petite population d'ions ${}^9\text{Be}^+$ (environ 5000), en tout premier lieu soumise à un champ magnétique statique assurant l'éclatement Zeeman des états pertinents. Les états $|g\rangle$ et $|m\rangle$ sont deux sous-niveaux¹³ hyperfins du fondamental ${}^2\text{S}_{\frac{1}{2}}$ ($(m_I, m_J) = (\frac{3}{2}, \frac{1}{2})$ et $(\frac{1}{2}, \frac{1}{2})$). Un champ de radiofréquence¹⁴ (RF) provoque la résonance de Rabi entre ces deux niveaux. Le rôle de $|e\rangle$ est tenu par une composante appartenant au multiplet excité ${}^2\text{P}_{\frac{3}{2}}$ ($(m_I, m_J) = (\frac{3}{2}, \frac{3}{2})$), couplée exclusivement au fondamental (l'émission spontanée de ${}^2\text{S}_{\frac{1}{2}, (\frac{1}{2}, \frac{1}{2})}$ est totalement négligeable).

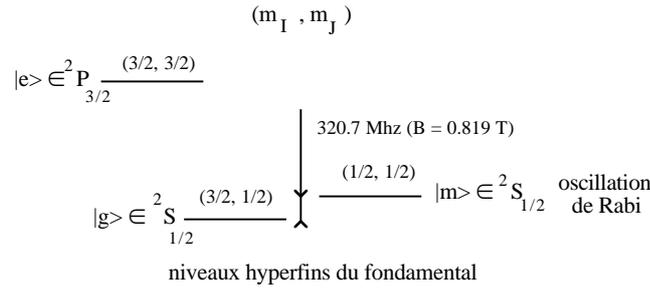
Soit $\hbar\Omega$ l'élément de matrice couplant les deux niveaux hyperfins pertinents, entre lesquels se produit l'oscillation de Rabi. On va voir que si le champ RF a une durée¹⁵ T égale à $\frac{\pi}{\Omega}$ – on parle alors de *π -pulse* –, l'ion

¹² C'est précisément l'une des caractéristiques de l'émission *spontanée*.

¹³ I se réfère au spin nucléaire, J au spin électronique.

¹⁴ Dans leur expérience, le champ magnétique vaut 0.819 T et la séparation de deux niveaux hyperfins est égale à 320.7 MHz ; l'amplitude du champ RF est ajustée de sorte que $\Omega T = \pi$, avec $T = 256$ ms, soit $\Omega \simeq 4\pi \simeq 12$ rd/s.

¹⁵ Comme la pulsation RF ω est très grande devant Ω (fréquence de Rabi), cela a un sens de parler d'un champ RF de fréquence (presque) parfaitement déterminée de durée $T \sim \Omega^{-1} \gg \omega^{-1}$.

Figure 4.2: Atome à trois niveaux de Itano et *al.*

partant du fondamental $|g\rangle$ se retrouverait avec la probabilité $P_m = 1$ dans l'état $|m\rangle$ à l'issue de l'excitation RF s'il évoluait librement, en dehors de toute mesure. En fait, si une série de *pulses* optiques de durée très courte est simultanément envoyée sur l'ion, à des intervalles de temps $p\frac{T}{N}$ ($p = 1, 2, \dots, N$), ces mesures successives vont ralentir la dynamique au point de donner $P_m = 0$ (soit $P_g = 1$) dans la limite d'un nombre infini de *pulses* : c'est très exactement l'effet Zénon quantique.

La description théorique repose fondamentalement sur l'idée suivante : entre deux *pulses*, le système évolue selon l'équation de Schrödinger (qui, de fait, n'implique que les deux niveaux hyperfins, ceux qui participent à l'oscillation de Rabi). Au contraire, l'effet d'une *pulse* (une mesure) ne relève pas de cette équation et est décrit conformément au postulat de réduction du paquet d'ondes ; comme on va le voir, ceci revient à annuler "à la main" des éléments non-diagonaux de l'opérateur densité. Avant de continuer la discussion de l'expérience, et pour en comprendre le traitement théorique, quelques éléments concernant l'opérateur densité sont donc maintenant nécessaires.

■ L'opérateur densité

L'opérateur densité est la quantité centrale de la statistique quantique, tout comme l'est la fonction d'onde pour la mécanique quantique. Pour un système quantique d'états $|E_n\rangle$ à l'équilibre thermique avec un thermostat à température T , les probabilités de Boltzmann sont proportionnelles à $e^{-\beta E_n}$. En pareil cas, on définit l'opérateur densité (d'équilibre) ρ :

$$\rho(\beta) = e^{-\beta H} = \sum_n Z^{-1} e^{-\beta E_n} |E_n\rangle\langle E_n| \equiv \sum_n \rho_{nn} |E_n\rangle\langle E_n|, \quad Z = \sum_n e^{-\beta E_n}, \quad E_n \leq E_{n+1}; \quad (4.26)$$

sur la base des états propres du Hamiltonien du système, ρ est représenté par la matrice diagonale dont les éléments de matrice sont les probabilités de Boltzmann, ce qui justifie en retour l'expression formelle $\rho = e^{-\beta H}$. L'entropie statistique s'obtient comme :

$$S = -k_B T \sum_n \rho_{nn} \ln \rho_{nn} = -k_B T \text{Tr}(\rho \ln \rho), \quad (4.27)$$

et s'annule si $\rho_{nn} = \delta_{nn_0}$.

La limite température nulle de ρ est :

$$\lim_{\beta \rightarrow +\infty} \rho(\beta) = |E_1\rangle\langle E_1| \quad (4.28)$$

où $|E_1\rangle$ est l'état fondamental ; dans cette limite, ρ est idempotent :

$$\rho^2(+\infty) = \rho(+\infty) \quad (4.29)$$

et on dit alors qu'il s'agit d'un *cas pur* (l'entropie est nulle dans tout cas pur, qu'il s'agisse du fondamental ou non). Au contraire, à toute température finie, on a :

$$\rho^2(\beta) \neq \rho(\beta) \quad (4.30)$$

et on parle alors de *cas mixte* ($S > 0$). D'une façon générale, pour un cas pur, il existe un vecteur d'état $|\Psi\rangle$ tel que :

$$\rho = |\Psi\rangle\langle\Psi| \quad (\text{cas pur}) , \quad (4.31)$$

dont la donnée est strictement équivalente à celle de $|\Psi\rangle$. L'idempotence de ρ est évidente sur l'écriture (4.31).

Les considérations précédentes, énoncées à propos de l'équilibre thermodynamique, peuvent être étendues à une situation variable dans le temps. Décrire la dynamique passe de prime abord par l'écriture de l'équation fondamentale pour $\rho(t)$ qui, en un sens, généralise au champ statistique l'équation de Schrödinger pour $|\Psi(t)\rangle$. Dans un contexte classique, le passage de la Mécanique à la Mécanique statistique s'opère par l'introduction de la fonction de densité ρ_{cl} qui donne la densité de probabilité dans l'espace des phases du système. La manipulation de l'équation fondamentale de la dynamique – ou ses variantes plus élaborées, les équations canoniques de Hamilton – fournit alors l'équation :

$$\frac{\partial\rho_{\text{cl}}}{\partial t} = \{H, \rho_{\text{cl}}\} , \quad (4.32)$$

où $\{H, \rho_{\text{cl}}\}$ est le crochet de Poisson : c'est l'équation de Liouville. Comme toujours, la version quantique de l'équation classique s'obtient en remplaçant le crochet de Poisson par $(i\hbar)^{-1}[\ , \]$, et l'on obtient ainsi l'équation de Liouville quantique :

$$i\hbar\frac{\partial\rho}{\partial t} = [H, \rho] . \quad (4.33)$$

C'est cette équation que l'on postule en Mécanique statistique quantique – tout comme on postule l'équation de Schrödinger $i\hbar\frac{\partial\Psi}{\partial t} = H\Psi$; quand H ne dépend pas du temps, son intégration formelle donne :

$$\rho(t) = e^{\frac{1}{i\hbar}Ht} \rho(0) e^{-\frac{1}{i\hbar}Ht} . \quad (4.34)$$

Si l'état de départ est un cas pur, alors il existe un $|\Psi(t=0)\rangle$ tel que $\rho(0) = |\Psi(0)\rangle\langle\Psi(0)|$ et alors :

$$\rho(t) = e^{\frac{1}{i\hbar}Ht} |\Psi(0)\rangle\langle\Psi(0)| e^{-\frac{1}{i\hbar}Ht} \equiv |\Psi(t)\rangle\langle\Psi(t)| . \quad (4.35)$$

D'où le résultat important : un cas pur reste un cas pur tant que l'évolution se fait suivant l'équation de Schrödinger¹⁶. Aucune évolution hamiltonienne (unitaire) ne peut transformer un cas pur en cas mixte ou inversement – tant qu'il est question de l'opérateur densité total.

Par ailleurs, (4.34) montre que $\rho(t)$ est un opérateur hermitique si $\rho(t=0)$ l'est, et que la trace de $\rho(t)$ est une constante du mouvement ; la convention habituelle de normalisation est $\text{Tr } \rho = 1$ de sorte que finalement :

$$\text{Tr } \rho(t) = \text{Tr } \rho(0) = 1 . \quad (4.36)$$

Le sens physique de $\rho(t)$ est le même que celui de l'opérateur d'équilibre. Sur la base propre du Hamiltonien, $\{|E_n\rangle\}_n$, $\rho(t)$ est représenté par une matrice (en général non-diagonale). Les éléments diagonaux ρ_{nn} sont les probabilités pour qu'une mesure effectuée à l'instant t donne la valeur E_n pour l'énergie :

$$\rho_{nn}(t) \equiv \langle E_n|\rho(t)|E_n\rangle = \text{Prob}[E = E_n, t] . \quad (4.37)$$

Autrement dit, les éléments diagonaux de ρ sont les *populations* (relatives) des différents états au sein de l'ensemble statistique¹⁷. De surcroît, on vérifie immédiatement que la valeur moyenne d'une observable A dans un cas pur est :

$$\langle A \rangle \equiv \langle \Psi(t)|A|\Psi(t) \rangle = \text{Tr} (A\rho(t)) . \quad (4.38)$$

La formule $\langle A \rangle(t) = \text{Tr} (A\rho(t))$ reste vraie pour un cas mixte, compte tenu du sens physique des éléments diagonaux de ρ et de la propriété d'invariance de la trace.

En tant qu'opérateur représenté par une matrice, ρ possède aussi des éléments non-diagonaux, que Claude Cohen-Tannoudji a proposé d'appeler *cohérences*, puisqu'elles jouent un rôle essentiel dans les interférences quantiques. En effet, soit un cas pur construit avec l'état non-stationnaire¹⁸ ($E_n = \hbar\omega_n$) :

$$|\Psi(t)\rangle = c_1 e^{-i\omega_1 t} |E_1\rangle + c_2 e^{-i\omega_2 t} |E_2\rangle . \quad (4.39)$$

¹⁶On devine que, précisément, cette hypothèse exclut radicalement tout processus de mesure quantique.

¹⁷(4.36) ne traduit rien d'autre que la conservation de la somme des probabilités.

¹⁸On a choisi un système à deux niveaux pour la simplicité. Bien évidemment, tout se généralise à un nombre quelconque d'états.

L'opérateur densité correspondant est $\rho(t) = |\Psi(t)\rangle\langle\Psi(t)|$ soit, après distribution :

$$\rho(t) = |c_1|^2|E_1\rangle\langle E_1| + |c_2|^2|E_2\rangle\langle E_2| + c_1c_2^*e^{-i(\omega_1-\omega_2)t}|E_1\rangle\langle E_2| + c_2c_1^*e^{-i(\omega_2-\omega_1)t}|E_2\rangle\langle E_1|, \quad (4.40)$$

ou, sous forme matricielle :

$$\rho(t) = \begin{bmatrix} |c_1|^2 & c_1c_2^*e^{-i(\omega_1-\omega_2)t} \\ c_2c_1^*e^{-i(\omega_2-\omega_1)t} & |c_2|^2 \end{bmatrix} \equiv \begin{bmatrix} \rho_{11} & \rho_{12}(t) \\ \rho_{21}(t) & \rho_{22} \end{bmatrix}. \quad (4.41)$$

Les cohérences $\rho_{nm}(t)$, $n \neq m$, oscillent à la pulsation de Bohr, laquelle caractérise l'oscillation des valeurs moyennes d'une observable ne commutant pas avec le Hamiltonien. Elles représentent globalement la cohérence de phase entre les différentes composantes d'un état non-stationnaire. En outre, leur existence – et leur forme en *produit* de deux facteurs – est essentiellement liée à l'idempotence de ρ , c'est-à-dire à la propriété de ρ de représenter un cas pur.

L'équation de Liouville (4.33) engendre une évolution Hamiltonienne au sens où un cas pur le reste au cours du temps. La transition cas pur \rightarrow cas mixte peut s'opérer essentiellement de deux façons : soit en couplant explicitement le système quantique à un grand système (ayant un nombre infini de degrés de liberté à la limite thermodynamique), souvent appelé *réservoir*, et en intégrant sur les degrés de liberté de ce dernier (ce qui conduit à un opérateur densité *réduit*), soit en appliquant strictement le scénario de la mesure idéale telle que l'impose la Mécanique Quantique.

Quand on couple un système à un réservoir, le “petit” système reste l'objet principal d'étude et on admet même qu'il n'a aucune influence en retour sur la dynamique du réservoir. Il en résulte que seuls les degrés de liberté du système présentent un intérêt, ceux du réservoir étant connus (d'une façon ou d'une autre) une fois pour toutes. L'idée est donc d'écrire l'équation de Liouville pour le “supersystème”, puis de prendre la trace sur tous les degrés de liberté du réservoir. On obtient alors une équation pour l'opérateur densité réduit, laquelle (sauf exception) ne conserve pas le caractère pur au cours du temps¹⁹. Toute la difficulté est la modélisation astucieuse du réservoir et de son interaction avec le petit système, permettant au total de définir un modèle raisonnable physiquement et tractable techniquement (pour quelques exemples, voir [21]).

L'opérateur densité n'est pas seulement requis pour décrire une évolution statistique à partir d'une situation hors d'équilibre, il est aussi l'outil idéal pour la description de l'état juste avant et juste après une mesure, effectuée simultanément (et indépendamment) sur les (micro)systèmes d'un ensemble au sens statistique ; ρ est donc très bien adapté pour transcrire algébriquement l'inévitable processus de réduction du paquet d'ondes. En effet, la mesure à un instant t de l'énergie sur un ensemble décrit par ρ donne les différentes valeurs propres E_n avec des probabilités P_n , lesquelles sont précisément les éléments diagonaux ρ_{nn} de l'opérateur ρ , représenté sur la base propre de H . Juste après la mesure, une fraction $|c_1|^2 \equiv P_1$ de systèmes se retrouve dans l'état $|E_1\rangle$, une fraction $|c_2|^2 \equiv P_2 = 1 - P_1$ se retrouve dans $|E_2\rangle$. L'ensemble statistique est alors décrit par un opérateur densité que la mesure a rendu *diagonal* ; en d'autres termes, la réduction du paquet d'ondes se traduit par *l'annulation des cohérences* :

$$\rho_{\text{juste avant la mesure à l'instant } t} = \begin{bmatrix} \rho_{11} & \rho_{12}(t) \\ \rho_{21}(t) & \rho_{22} \end{bmatrix} \longrightarrow \begin{bmatrix} \rho_{11} & 0 \\ 0 & \rho_{22} \end{bmatrix} = \rho_{\text{juste après la mesure à l'instant } t}. \quad (4.42)$$

■

Retournons maintenant à l'effet Zénon observé pour l'atome à trois niveaux et à l'expérience de Cook. La dynamique s'effectue principalement dans le sous-espace ($|g\rangle, |m\rangle$), la réduction effective du paquet d'ondes conduisant soit vers $|m\rangle$, soit vers $|g\rangle$, suivant l'instant auquel arrive le *pulse* optique dans le cycle de Rabi.

¹⁹Un phénomène analogue se produit déjà en Physique classique dans des cas très simples, même lorsque peu de degrés de liberté sont impliqués. La fonction de densité $\rho(q, p, t)$ d'un système à un degré de liberté obéit à une équation réversible par renversement du temps, puisqu'elle n'est qu'une réécriture des équations de la Mécanique incorporant la conservation de la probabilité. Toutefois, la fonction réduite (marginale) $\rho_q(q, t)$, obtenue par intégration sur l'impulsion et donnant la densité probabilité de présence au point d'abscisse q quelle que soit l'impulsion, a une évolution irréversible en général. On pourra retenir l'idée suivante : prendre une trace partielle c'est perdre de l'information, une perte qui *peut* se traduire par une brisure de la symétrie de renversement du temps.

C'est aussi pour cette raison que, tout en restant strictement dans un cadre classique – et en excluant évidemment toute sensibilité aux conditions initiales –, il existe des systèmes très simples (une particule libre !) pour lesquels toute imprécision (inévitable !) sur les conditions initiales ruine toute prévision à long terme.

Le Hamiltonien $H(t)$ de l'atome couplé (semi-classiquement) au champ RF est ainsi représenté par une matrice 2×2 , que l'on peut toujours prendre de trace nulle en choisissant le zéro d'énergie au milieu des énergies E_g et E_m . En désignant toujours par $\hbar\omega_0$ la différence $E_m - E_g$, par ω la pulsation du champ RF, la matrice de $H(t)$ sur la base ordonnée $(|m\rangle, |g\rangle)$ est de la forme :

$$H(t) = \frac{\hbar}{2} \begin{bmatrix} \omega_0 & \Omega e^{-i\omega t} \\ \Omega e^{i\omega t} & -\omega_0 \end{bmatrix} . \quad (4.43)$$

Dans le cas d'un système à deux niveaux, on peut toujours représenter les matrices de trace nulle en combinaison linéaire des trois matrices de Pauli²⁰ définissant l'opérateur vectoriel $\vec{\sigma}$:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} , \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} , \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} . \quad (4.44)$$

C'est ainsi que, s'agissant de H , on peut écrire :

$$H(t) = \frac{\hbar}{2} \vec{\mathcal{H}} \cdot \vec{\sigma} , \quad (4.45)$$

où $\vec{\mathcal{H}}$ a pour composantes les quantités $\hbar^{-1} \text{Tr}(H\sigma_u)$, ici égales à :

$$\vec{\mathcal{H}} = (\Omega \cos \omega t, \Omega \sin \omega t, \omega_0) . \quad (4.46)$$

Pour un système à deux niveaux, l'opérateur densité est aussi (représenté par) une matrice 2×2 , d'éléments ρ_{gg} , ρ_{mm} , ρ_{mg} et ρ_{gm} , matrice hermitique si la base est orthonormée (ce qui est le cas). Comme de plus la trace de ρ est égale à 1, sa matrice possède au total trois degrés de liberté réels. Il existe de ce fait une représentation vectorielle²¹ commode de ρ à l'aide d'un vecteur \vec{R} de \mathbb{R}^3 , dont les trois composantes réelles sont traditionnellement définies comme :

$$R_x = \rho_{gm} + \rho_{mg} , \quad R_y = i(\rho_{mg} - \rho_{gm}) , \quad R_z = \rho_{mm} - \rho_{gg} . \quad (4.47)$$

Les cohérences de $\rho(t)$ sont donc associées aux composantes transverses R_x et R_y , alors que les populations sont contenues dans la composante longitudinale R_z . On vérifie sans peine que la relation précise entre ρ et \vec{R} est²² :

$$\rho = \frac{1}{2} (\mathbf{1} + \vec{R} \cdot \vec{\sigma}) , \quad R_u = \text{Tr}(\rho \sigma_u) ; \quad (4.48)$$

en particulier, on a :

$$\rho_{mm} = \frac{1}{2}(1 + R_z) , \quad \rho_{gg} = \frac{1}{2}(1 - R_z) . \quad (4.49)$$

Avec cette paramétrisation, l'équation de Liouville (4.33) pour ρ s'écrit :

$$i\hbar \frac{d\rho}{dt} = \frac{\hbar}{4} [\vec{\mathcal{H}} \cdot \vec{\sigma}, \vec{R} \cdot \vec{\sigma}] , \quad (4.50)$$

et, après calcul²³, prend la forme gyroscopique :

$$\frac{d\vec{R}}{dt} = \vec{\mathcal{H}}(t) \times \vec{R} . \quad (4.51)$$

En passant dans le repère $Ox'y'z$ tournant à ωt autour de Oz , c'est-à-dire en posant²⁴ :

$$\rho'(t) = U^\dagger(t) \rho(t) U(t) , \quad U(t) = e^{-i \frac{\omega t}{2} \sigma_z} , \quad (4.52)$$

²⁰Pour des raisons évidentes, cette façon d'écrire s'appelle parfois *représentation en spin fictif*.

²¹Souvent appelée *représentation de Bloch*. Pour un état pur, \vec{R} est unitaire d'où la notion de *sphère de Bloch* : en pareil cas, l'évolution de $\rho(t)$ se traduit par un mouvement de l'extrémité de $\vec{R}(t)$ sur la sphère unité.

²²Si ρ dépend du temps, il en va de même du vecteur \vec{R} .

²³On utilise au passage : $(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) = \vec{A} \cdot \vec{B} \mathbf{1} + i\vec{\sigma} \cdot (\vec{A} \times \vec{B})$, relation vraie pour des opérateurs tels que $[A_u, \sigma_{u'}] = [B_v, \sigma_{v'}] = 0$.

²⁴ $U(t)$ est bien la rotation d'angle ωt autour de Oz , qui ne change pas la composante R_z : $R'_z = R_z$.

on obtient pour \vec{R}' associé à ρ' :

$$\frac{d\vec{R}'}{dt} = \vec{\mathcal{H}}' \times \vec{R}' , \quad \vec{\mathcal{H}}' = (\Omega, 0, \omega_0 - \omega) ; \quad (4.53)$$

le point important techniquement est que, maintenant, $\vec{\mathcal{H}}'$ est indépendant du temps ; de ce fait l'intégration de (4.53) est immédiate. À la résonance stricte ($\omega = \omega_0$), \vec{R}' précesse autour de Ox' à la pulsation Ω ; si à l'instant $t = 0$, l'ion est dans le fondamental $|g\rangle$, alors $\vec{R}' = (0, 0, -1)$ – et en l'absence de toute autre excitation – se retrouve²⁵ dans l'état $|m\rangle$ au temps $T = \frac{\pi}{\Omega}$, et alors $\vec{R}' = (0, 0, +1)$: c'est bien parce que le vecteur \vec{R}' a fait un demi-tour que l'on parle de π -pulse (pour le champ RF). À un instant quelconque, \vec{R}' est quelque part dans le plan $y'Oz$ du repère tournant, les deux populations sont différentes de zéro, tout comme les cohérences : c'est l'oscillation de Rabi.

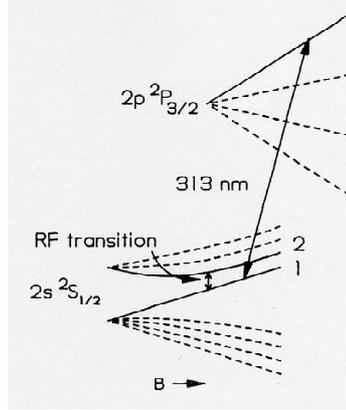


Figure 4.3: Diagramme des niveaux d'énergie de l'ion Be^+ dans un champ magnétique B (figure extraite de [20]).

Introduisons maintenant le train régulier de *pulses*, comme dans l'expérience d'Itano *et al.*, chacun d'entre eux jouant le rôle d'une opération de mesure et survenant à des instants t_n calés sur la période T , $t_n = n \frac{T}{N}$. Juste *avant* la première mesure, survenant à $t_1 = \frac{T}{N} = \frac{\pi}{N\Omega}$, on a :

$$\vec{R}'(t = \frac{\pi}{N\Omega}) = (0, \sin \Omega t, -\cos \Omega t)_{t=\frac{\pi}{N\Omega}} = (0, \sin \frac{\pi}{N}, -\cos \frac{\pi}{N}) . \quad (4.54)$$

Survient alors le premier *pulse* optique ; traité comme une mesure idéale, il projette l'ion soit dans l'état $|m\rangle$, soit dans l'état de $|g\rangle$. Sur un ensemble d'ions représentés par l'opérateur densité, l'effet est d'annuler les éléments non-diagonaux (cohérences, composantes R_x et R_y , combinaisons linéaires de R'_x et R'_y) et de laisser intacts les éléments diagonaux (populations, composante $R_z = R'_z$). À l'issue du 1^{er} *pulse* optique, l'opérateur densité est donc associé au vecteur :

$$\vec{R}'(t = \frac{\pi}{N\Omega} + 0) = (0, 0, -\cos \frac{\pi}{N}) . \quad (4.55)$$

Au total, après la première opération de mesure effectuée par le premier *pulse* optique, \vec{R}' se retrouve dans la direction de départ mais a une longueur contractée d'un facteur $\cos \frac{\pi}{N}$. Chaque *pulse* successif donne lieu à une telle contraction : en vertu de la linéarité de l'équation gyroscopique (4.53), au bout de N mesures, le vecteur \vec{R}' pointe dans la même direction qu'au départ, mais sa longueur est réduite d'un facteur $(\cos \frac{\pi}{N})^N$. Ainsi, à l'issue de la $N^{\text{ème}}$ mesure, on a :

$$\vec{R}'(t = T + 0) = (0, 0, -\cos^N \frac{\pi}{N}) . \quad (4.56)$$

Compte tenu de (4.47) et de $R'_z = R_z$, la probabilité pour un ion d'être dans $|m\rangle$ à $T + 0$ est donc :

$$P_m(T + 0) \equiv \rho_{mm}(T + 0) = \frac{1}{2} [1 + R_z(T + 0)] = \frac{1}{2} \left(1 - \cos^N \frac{\pi}{N} \right) . \quad (4.57)$$

²⁵En moyenne ! Toute mesure révèle qu'un ion donné est à tout instant *soit* en $|g\rangle$, *soit* en $|m\rangle$.

En développant le cosinus dans l'hypothèse $N \gg 1$, on a :

$$P_m(T+0) \simeq \frac{1}{2} \left[1 - e^{+N \ln(1 - \frac{\pi^2}{2N^2})} \right] \simeq \frac{1}{2} \left(1 - e^{-\frac{\pi^2}{2N}} \right) \simeq \frac{\pi^2}{4N} \ll 1 . \quad (4.58)$$

En l'absence de mesures répétées, on trouverait $P_m(T+0) = 1$, puisqu'il s'agit d'un π -pulse : (4.58) montre que la succession des mesures très rapprochées ruine la possibilité pour l'ion de filer dans l'état $|m\rangle$ (plus précisément : d'y être trouvé). Inversement, la probabilité d'être toujours dans l'état de départ $|g\rangle$ est $P_g(T+0) \simeq 1 - \frac{\pi^2}{4N} \sim 1$: au total, le train de mesures successives empêche l'ion d'évoluer.

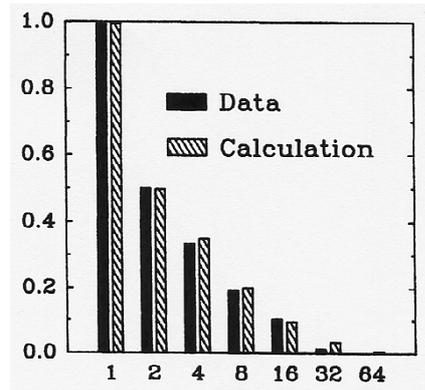


Figure 4.4: Variation de la probabilité d'être en $|m\rangle$ en fonction du nombre N de pulses (figure extraite de [20]).

La décroissance de P_m avec le nombre de mesures (voir fig. 4.4) décrit quantitativement l'effet Zenon, et se trouve en plein accord avec les mesures de Itano *et al*, constituant ainsi la mise en évidence expérimentale de cet effet (voir [20] pour plus de détails).

4.2 Sauts quantiques

La notion de saut quantique remonte à la préhistoire de la Mécanique Quantique. À la suite de l'introduction du photon par Einstein (effet photo-électrique, 1905), Bohr suggéra en 1913 que l'interaction entre un atome et le rayonnement procède par absorption ou émission de photons, l'électron atomique effectuant des sauts d'une orbite à une autre. À titre d'exemple, le processus d'émission suit la règle des fréquences de Bohr, qui n'est rien d'autre que l'expression de la conservation de l'énergie²⁶ :

$$E_i = E_f + h\nu . \quad (4.59)$$

E_i et E_f sont les énergies initiale et finale de l'atome, ν est la fréquence du photon reçu par le milieu extérieur à l'atome. En passant de l'état i à l'état f , l'atome effectue une transition, que l'on peut convenir d'appeler "saut quantique" ("*Quantum jumps*").

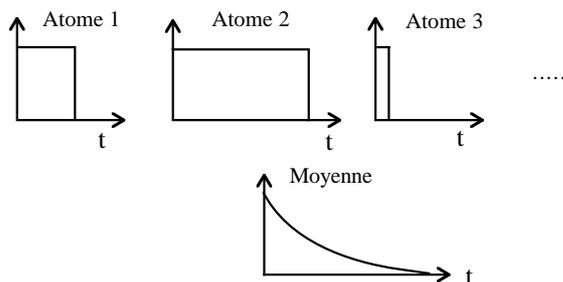
Pendant des décennies, ce mécanisme primaire a été inobservable en tant que tel. En effet, une raie d'émission étant choisie pour un atome donné, son observation s'est faite usuellement sur une vapeur atomique contenant un très grand nombre d'atomes²⁷. De ce fait, les transitions individuelles, noyées dans la masse et survenant à des instants aléatoires décorrélés, construisent au total une intensité de fluorescence macroscopique dont les fluctuations sont imperceptibles au point de confondre l'intensité aléatoire avec sa valeur *moyenne* (voir fig. 4.5). La mesure de cette intensité dans une vapeur où une fraction notable d'atomes est portée à l'état

²⁶On suppose que l'atome est infiniment massif. Dans le cas contraire, il faut prendre en compte l'énergie de recul de l'atome.

²⁷En 1952, Schrödinger écrivait :

"... we never experiment with just one electron or atom. In thought experiments we sometimes assume we do ; this invariably entails ridiculous consequences."

(cité en [22]).

Figure 4.5: Émissions individuelles *vs* intensité moyenne de fluorescence.

excité révèle un déclin essentiellement exponentiel, permettant de définir un temps caractéristique $\tau_e = \Gamma_e^{-1}$; τ_e n'est autre que la durée de vie *moyenne* d'un atome à l'état excité. En réalité, certains atomes se désexcitent bien avant τ_e , d'autres bien après et c'est seulement quand on fait la moyenne des temps de mort individuels que l'on retrouve la durée de vie moyenne²⁸. Pour espérer voir les sauts quantiques, il faut donc travailler avec un petit nombre de sources (atomes, ions), mais encore faut-il que le signal attestant de la présence d'un atome dans un état ou un autre soit assez intense pour pouvoir être observé et mesuré.

Assez récemment (au début des années 80), il est devenu possible de piéger quelques ions et de les refroidir à l'aide de lasers, de façon à les observer dans des conditions quasi-idéales. C'est alors qu'une proposition d'expérience de double résonance optique dûe à Dehmelt [23] a pu être mise en œuvre ; en essence, l'idée est la suivante.

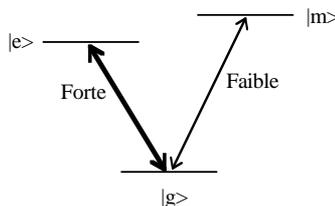


Figure 4.6: Atome à trois niveaux pour l'expérience de double résonance optique.

On considère un atome à trois niveaux, comme celui introduit dans la section 4.1 (voir figure 4.6). Les deux niveaux $|m\rangle$ et $|e\rangle$ sont supposés avoir des énergies voisines, les deux différences $E_m - E_g$ et $E_e - E_g$ sont dans le domaine optique ou UV – ce qui distingue l'état métastable considéré ici, de l'état métastable utilisé²⁹ pour l'effet Zenon (section 4.1). Typiquement, la durée de vie de $|e\rangle$ est de l'ordre de 10^{-8} s, celle de $|m\rangle$ est de l'ordre de la seconde. Ces deux ordres de grandeurs, très différents, indiquent que les probabilités des transitions sont elles-mêmes très différentes : la transition $|g\rangle \leftrightarrow |m\rangle$ est très rare, comparée à $|g\rangle \leftrightarrow |e\rangle$. Par ailleurs, l'atome est soumis à deux lasers ; le premier, très intense, sature la transition $|g\rangle \leftrightarrow |e\rangle$, tandis que le second est calé sur la transition rare $|g\rangle \leftrightarrow |m\rangle$.

Si l'état de l'ion est le fondamental $|g\rangle$ ou l'état excité $|e\rangle$, le laser intense induit un mouvement de

²⁸Il en va exactement de même pour la décroissance radioactive. Au début (beaucoup de noyaux actifs), l'énormité du nombre de noyaux présents dans une source macroscopique engendre un déclin exponentiel ne présentant aucune anomalie, aucune marche n'étant visible à la sensibilité affichée (pendant l'intervalle élémentaire de comptage, δt , il se produit un nombre faramineux de désintégrations) ; pendant cette première phase, il est licite de confondre la valeur de la variable aléatoire N (nombre de noyaux actifs à l'instant considéré, observée) avec sa valeur *moyenne*. Au contraire, au bout d'un temps très long devant la période radioactive, des fluctuations toujours présentes deviennent visibles (puisqu'il reste alors un petit nombre de noyaux encore excités) et l'exponentielle, mesurée pour une source donnée, devient "bruitée". Analysant le déclin radioactif comme il doit l'être (un processus *aléatoire*), on trouve d'ailleurs que les fluctuations relatives $\frac{\Delta N}{\langle N \rangle}$ *divergent* quand $t \rightarrow +\infty$.

²⁹En particulier, l'aller-retour entre $|g\rangle$ et $|m\rangle$ n'est pas une oscillation de Rabi : c'est simplement le mécanisme ordinaire d'absorption et d'émission (essentiellement spontanée) entre deux états couplés par un champ électromagnétique. Dans le domaine des micro-ondes, l'émission induite l'emporte sur l'émission spontanée, c'est le contraire dans le domaine optique ou UV. Cette règle générale possède des exceptions : même dans le domaine optique, l'émission spontanée peut être faible en comparaison de l'émission induite pourvu que l'on utilise un laser très intense (alors, on peut avoir $A \ll BW$, dans les notations introduites peu après).

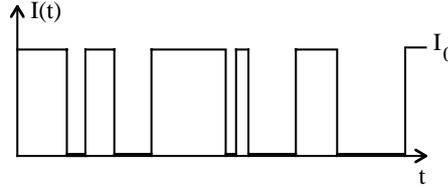


Figure 4.7: Fluorescence intermittente d'un atome unique.

va-et-vient rapide entre ces deux états ; cette “oscillation” est rapide puisque la transition est permise et que la durée de vie de $|e\rangle$ est courte ; elle produit une fluorescence intense, facilement observable. Cependant, l'autre laser peut à un certain instant porter l'atome à l'état $|m\rangle$; il faut du temps pour cela, puisque la transition est rare – mais elle finit bien par se produire. Dès lors, la fluorescence s'arrête et reste “noire” tant que l'atome n'est pas redescendu en $|g\rangle$ à partir de $|m\rangle$. Quand ceci se produit, la fluorescence recommence, et ainsi de suite. Au total, dans l'expérience proposée par Dehmelt, on doit assister à une fluorescence intermittente³⁰ illustrée schématiquement sur la figure 4.7. Par ailleurs, on peut affirmer que quand la fluorescence est absente, l'ion est dans l'état $|m\rangle$ et que quand elle se produit, il est dans le sous-espace (secteur) $(|g\rangle, |e\rangle)$. En un sens, l'intermittence de la fluorescence est une succession de réductions de paquets d'ondes, ou, dit autrement, l'observation de l'état de fluorescence constitue une observation directe des sauts quantiques³¹ impliquant l'état $|m\rangle$: l'extinction de la fluorescence signifie que l'ion vient d'arriver en $|m\rangle$, sa réapparition que l'ion vient de regagner le secteur $(|g\rangle, |e\rangle)$.

Le premier traitement théorique a été donné par Cook et Kimble [24] et repose sur des équations *cinétiques* pour les populations des états atomiques, qui ne sont autres que des équations de bilan. En introduisant les coefficients d'Einstein A_a et B_{bc} , et les densités spectrales W_i des lasers, ces équations s'écrivent :

$$\dot{P}_g = -(B_{eg}W_1 + B_{mg}W_2)P_g + (A_e + B_{eg}W_1)P_e + (A_m + B_{mg}W_2)P_m, \quad (4.60)$$

$$\dot{P}_e = -(A_e + B_{eg}W_1)P_e + B_{eg}W_1P_g, \quad (4.61)$$

$$\dot{P}_m = -(A_m + B_{mg}W_2)P_m + B_{mg}W_2P_g, \quad (4.62)$$

ce que l'on peut traduire matriciellement :

$$|\dot{P}\rangle = -\mathbf{\Gamma}|P\rangle. \quad (4.63)$$

On vérifie que $\sum \dot{P}_a = 0$, relation qui exprime la conservation de la somme $P_g + P_e + P_m = 1$ (la somme des éléments d'une colonne de $\mathbf{\Gamma}$ est nulle, d'où une valeur propre nulle pour $\mathbf{\Gamma}$) ; par hypothèse, $A_m \ll A_e$. Le système (4.60) - (4.63) a une solution d'équilibre $|P^{eq}\rangle$ obtenue en annulant tous les premiers membres ; on trouve ainsi³², l'émission spontanée :

$$\frac{P_e^{eq}}{P_g^{eq}} = \frac{B_{eg}W_1}{A_e + B_{eg}W_1}, \quad \frac{P_m^{eq}}{P_g^{eq}} = \frac{B_{mg}W_2}{A_m + B_{mg}W_2}; \quad (4.64)$$

si on peut négliger le terme d'émission spontanée dans chaque dénominateur, les populations d'équilibre sont égales entre elles et valent chacune $\frac{1}{3}$. Cette situation est atteinte longtemps après la plus longue échelle de temps, usuellement A_m^{-1} , qui est gigantesque par rapport à toutes les autres.

Avec W_2 strictement nul (deuxième laser coupé) et $P_g(t=0) = 1$, on obtient la quasi-égalité³³ des deux probabilités P_g et P_e ($P_g \simeq P_e = \frac{1}{2}$) après un temps de l'ordre de $(A_e + 2B_{eg}W_1)^{-1}$: on dit alors que

³⁰La transition faible survient à des instants aléatoires ; si ces derniers sont décorrélés, ce processus aléatoire binaire (prenant deux valeurs) est connu sous le nom de processus des télégraphistes. Le mot *intermittence* n'est pas à prendre dans une acception récente qui qualifie la dynamique “par bouffées” de certains systèmes chaotiques.

³¹Le signal observé est la fluorescence induite par le premier laser, dont la puissance spectrale détermine l'intensité de la fluorescence ; la question n'est pas de pouvoir observer le *photon unique* émis lors de la transition $|m\rangle \rightarrow |g\rangle$.

³²Les expressions (4.64) montrent bien la nécessité de l'émission spontanée pour obtenir des populations décroissant quand l'énergie croît. Dans un tout autre contexte (équilibre thermodynamique du rayonnement en présence d'un peu de matière), c'est bien l'émission spontanée qui permet de retrouver les probabilités de Boltzmann $P_E = C e^{-\beta E}$, décroissantes avec l'énergie.

³³Ceci provient de $A_e \ll B_{eg}W_1$, réalisable avec un W_1 suffisamment grand. Quand $W_2 = 0$ et avec $P_g(t=0) = 1$, la solution exacte de (4.60) - (4.61) est $P_g(t) = \frac{1}{A_e + 2B_{eg}W_1}(A_e + B_{eg}W_1 + B_{eg}W_1 e^{-(A_e + 2B_{eg}W_1)t})$, $P_e(t) = 1 - P_g(t)$.

la transition (forte) $e \leftrightarrow g$ est *saturée*. Avec W_2 non nul (mais $B_{mg}W_2 \ll B_{eg}W_1$) la transition forte reste saturée³⁴ – ce que l'on obtient quand W_1 est très grand (assurant que $A_e \ll B_{eg}W_1$) – les populations P_e et P_g sont à peu près égales³⁵ ; on se retrouve de fait avec une sorte de système à deux niveaux : l'atome est soit dans le secteur ($|g\rangle, |e\rangle$) avec des populations égales, soit dans l'état $|m\rangle$; il suffit dès lors de déterminer les deux populations \mathcal{P}_\pm :

$$\mathcal{P}_+ = \mathcal{P}_m \quad \mathcal{P}_- = \mathcal{P}_g + \mathcal{P}_e , \quad (4.65)$$

sachant que, dans ce régime, on a :

$$P_g \simeq P_e \simeq \frac{1}{2}\mathcal{P}_- . \quad (4.66)$$

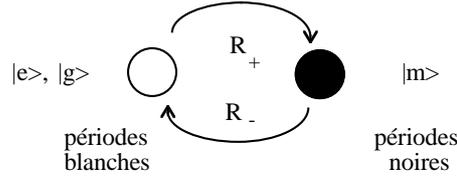


Figure 4.8: R_+ éteint la fluorescence, R_- la rallume.

En reportant les nouvelles variables \mathcal{P}_\pm dans (4.62), il vient compte tenu de (4.66) :

$$\dot{\mathcal{P}}_+ = -(A_m + B_{mg}W_2)\mathcal{P}_+ + \frac{1}{2}B_{mg}W_2\mathcal{P}_- , \quad (4.67)$$

soit :

$$\dot{\mathcal{P}}_+ = -R_- \mathcal{P}_+ + R_+ \mathcal{P}_- , \quad (4.68)$$

où $R_+ = \frac{1}{2}B_{mg}W_2$, $R_- = A_m + B_{mg}W_2$. Comme $\mathcal{P}_+ + \mathcal{P}_- = 1$, on en déduit :

$$\dot{\mathcal{P}}_- = -R_+ \mathcal{P}_- + R_- \mathcal{P}_+ . \quad (4.69)$$

L'inspection de ces équations révèle que R_- allume la fluorescence, alors que R_+ l'éteint (voir fig. 4.8).

L'état d'équilibre des équations (4.68) et (4.69) est :

$$\mathcal{P}_{+eq} = \frac{R_+}{R_+ + R_-} \quad \mathcal{P}_{-eq} = \frac{R_-}{R_+ + R_-} . \quad (4.70)$$

Ces probabilités donnent la *moyenne* temporelle du signal de fluorescence, une fois les transitoires éteints ; si I_0 désigne l'intensité (constante) d'une période allumée (blanche) :

$$\bar{I} = \mathcal{P}_{+eq} \times 0 + \mathcal{P}_{-eq} \times I_0 = I_0 \mathcal{P}_{-eq} = \frac{R_-}{R_+ + R_-} I_0 \quad (4.71)$$

En revanche, les probabilités stationnaires (4.70) ne disent évidemment rien sur les *fluctuations* de ce signal. Celles-ci ne peuvent être obtenues que comme sous-produits d'une description plus fine du processus stochastique, que l'on suppose stationnaire (toutes les probabilités à deux temps ne dépendent que de la différence des temps, toutes les probabilités à un temps sont indépendantes du temps).

Désignons par $P_{n\pm}(t)$ la probabilité qu'il y ait entre les instants t_0 et $t_0 + t$, n transitions d'un secteur à l'autre et que l'état final (à l'instant t) soit \pm ; par construction, les valeurs de ces probabilités à $t = 0$ sont 0 et 1, suivant l'état à $t = 0$ (aucune transition ne peut se produire sur un intervalle de temps nul). Par l'hypothèse

³⁴Les deux populations P_g et P_e sont presque égales, mais différent de $\frac{1}{2}$. Sur une échelle de temps courte par rapport à $(B_{mg}W_2)^{-1}$, tout se passe comme si le niveau métastable était absent et ces deux populations valent à peu près $\frac{1}{2}$. En revanche, aux temps longs, bien après que le niveau $|m\rangle$ a eu le temps de jouer son rôle et toujours dans l'hypothèse où la transition forte reste saturée, toutes les populations tendent vers $\frac{1}{3}$ quand l'état $|m\rangle$ a une très longue durée de vie (pour la discussion détaillée, voir [22]). En définitive, avec les hypothèses en vigueur, le système se comporte d'abord comme un atome à deux niveaux, puis comme un atome à trois niveaux.

³⁵Ce que confirme la résolution du système ((4.60) - (4.62)) dans ce régime de saturation.

de stationnarité (invariance par translation dans le temps, indifférence de l'instant appelé t_0), ces probabilités ne dépendent que de t . Pour obtenir leur équation d'évolution, il suffit de faire un bilan classique de probabilité entre deux instants très proches t et $t + \Delta t$, et de prendre la limite $\Delta t \rightarrow 0$. Par exemple, entre t et $t + \Delta t$, P_{n+} peut, par un seul *flip* au plus, évoluer de deux façons :

1. il ne se passe rien, probabilité $1 - R_- \Delta t$
2. il y a eu $n - 1$ transitions jusqu'à t , qui ont conduit le système dans l'état $-$ et l'une se produit entre t et $t + \Delta t$, probabilité $R_+ \Delta t$

Évidemment, bien d'autres processus sont possibles, mais comme ils impliquent plus d'un *flip*, leurs probabilités sont de ce fait en $\Delta t^{k \geq 2}$: ils ne contribuent donc pas, dans la limite $\Delta t \rightarrow 0$, à l'histoire de P_{n+} entre t et $t + \Delta t$. Au total, pour P_{n+} , le bilan de probabilité s'exprime comme :

$$P_{n+}(t + \Delta t) = (1 - R_- \Delta t)P_{n+}(t) + R_+ \Delta t P_{n-1-}(t) + \mathcal{O}(\Delta t^{k \geq 2}) ; \quad (4.72)$$

de même, pour P_{n-} , on obtient :

$$P_{n-}(t + \Delta t) = (1 - R_+ \Delta t)P_{n-}(t) + R_- \Delta t P_{n-1+}(t) + \mathcal{O}(\Delta t^{k \geq 2}) . \quad (4.73)$$

Dans la limite $\Delta t \rightarrow 0$, ces équations deviennent ($n = 0, 1, 2, \dots$) :

$$\dot{P}_{n+}(t) = -R_- P_{n+} + R_+ P_{n-1-} , \quad (4.74)$$

$$\dot{P}_{n-}(t) = -R_+ P_{n-} + R_- P_{n-1+} . \quad (4.75)$$

En particulier, comme $P_{-1\pm} = 0$:

$$\dot{P}_{0\pm}(t) = -R_{\mp} P_{0\pm}(t) \iff P_{0\pm}(t) = C e^{-R_{\mp} t} . \quad (4.76)$$

Supposons que le signal de fluorescence s'éteigne à l'instant t_0 ; ceci signifie que l'atome vient d'arriver dans l'état métastable ; la probabilité qu'il n'y ait eu aucune transition entre $t = t_0$ et $t = t_0$ et que l'atome soit dans l'état $+$ est donc égale à 1, d'où $P_{0+}(0) = 1$ et :

$$P_{0+}(t) = e^{-R_- t} . \quad (4.77)$$

Ce dernier résultat peut d'ailleurs se retrouver immédiatement comme suit. $P_{0+}(t)$ est finalement la probabilité pour que le système étant parti de $|+\rangle$ ($= |m\rangle$) à l'instant zéro, soit en t dans le même état *sans avoir jamais transité*, autrement dit la probabilité pour que l'atome ait été en permanence dans $|+\rangle$ entre 0 et t . La probabilité de transition par unité de temps pour le passage $|m\rangle \rightarrow |g\rangle$ est R_- ; on peut donc écrire de façon évidente :

$$P_{0+}(t + \Delta t) = (1 - R_- \Delta t)P_{0+}(t) + \mathcal{O}(\Delta t^{k \geq 2}) , \quad (4.78)$$

d'où l'on déduit par passage à la limite :

$$\dot{P}_{0+}(t) = -R_- P_{0+}(t) \iff P_{0+}(t) = e^{-R_- t} . \quad (4.79)$$

Cette expression donne donc la probabilité pour que le signal de fluorescence soit encore éteint à l'instant $t + t_0$ quand il l'était à $t = t_0$ – quel que soit l'instant t_0 puisque l'on est en régime permanent – sans s'être jamais allumé dans l'intervalle ; évidemment, plus le temps passe, plus cette probabilité est petite. Si T désigne la variable aléatoire : durée de la période noire, on a donc :

$$\text{Prob}[T > t] = e^{-R_- t} . \quad (4.80)$$

De cette fonction de répartition, on déduit la densité de probabilité de la durée des périodes noires :

$$\lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \text{Prob}[t < T < t + \delta t] = R_- e^{-R_- t} . \quad (4.81)$$

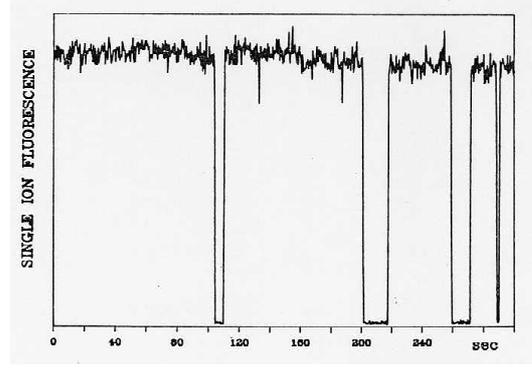


Figure 4.9: Signal de fluorescence intermittente d'un ion Ba^+ (figure extraite de [26]).

Les temps de résidence dans l'état métastable³⁶ sont donc théoriquement distribués suivant une loi exponentielle, qui est bien vérifiée expérimentalement [25] moyennant quelques corrections venant affiner après coup le modèle théorique de Cook et Kimble. La connaissance de la loi (4.81) permet notamment d'obtenir la durée moyenne des périodes sans signal :

$$\langle T \rangle = \int_0^{+\infty} dt t R_- e^{-R_- t} = \frac{1}{R_-} . \quad (4.82)$$

Des arguments analogues montrent que la durée des périodes "blanches" est distribuée selon la densité $R_+ e^{-R_+ t}$ et que sa moyenne est $\frac{1}{R_+}$. La durée des périodes noires est donc distribuée selon une loi exponentielle donnée par (4.81), à nouveau assez bien vérifié expérimentalement [25].

Il est également possible de calculer (et de mesurer) la fonction d'autocorrélation (centrée) de l'intensité, qui donne des renseignements plus fins sur le processus aléatoire. Elle est définie comme :

$$C(t) = \langle I(t_0)I(t_0+t) \rangle - \langle I(t_0) \rangle \langle I(t_0+t) \rangle = \langle I(t_0)I(t_0+t) \rangle - I_0^2 ; \quad (4.83)$$

comme la variable aléatoire I vaut 0 ou I_0 , le produit $I(t)I(t')$ peut prendre deux valeurs, 0 ou I_0^2 – cette dernière valeur étant obtenue si le signal existe à l'instant t et à l'instant t' . L'espérance mathématique du produit contient donc un seul terme :

$$\langle I(t)I(t') \rangle = I_0^2 P_{\text{on on}}(t, t') ; \quad (4.84)$$

où $P_{\text{on on}}(t, t')$ est la probabilité pour que le signal soit "on" à l'instant t et à l'instant t' . Grâce à la stationnarité, il suffit de trouver $P_{\text{on on}}(t_0, t_0+t) \equiv P_{\text{on on}}(t)$, probabilité pour que l'on soit dans une période blanche à l'instant t_0 et à l'instant t_0+t , quel que soit le nombre de transitions intervenues entre t_0 et t_0+t (on peut donc définitivement poser $t_0 = 0$).

Pour trouver cette probabilité, il suffit de faire à nouveau le bilan des probabilités entre t et $t + \Delta t$. C'est R_+ qui éteint la fluorescence, donc $1 - R_+ \Delta t$ donne la probabilité la fluorescence reste allumée Δt plus tard. Il y a un autre terme dans le bilan, introduisant la probabilité $P_{\text{on off}}$ pour que le système soit "on" à t et "off" à 0. En effet, *via* R_- – qui allume la fluorescence – il y a un autre chemin contribuant à $P_{\text{on on}}(t)$: c'est celui où le système est allumé en 0, éteint à t et bascule entre t et $t + \Delta t$. Le bilan est donc :

$$P_{\text{on on}}(t + \Delta t) = (1 - R_+ \Delta t) P_{\text{on on}}(t) + R_- \Delta t P_{\text{on off}}(t) . \quad (4.85)$$

Il faut donc aussi connaître $P_{\text{on off}}(t)$; des arguments analogues montrent que cette probabilité satisfait :

$$P_{\text{on off}}(t + \Delta t) = (1 - R_- \Delta t) P_{\text{on off}}(t) + R_+ \Delta t P_{\text{on on}}(t) . \quad (4.86)$$

Dans la limite $\Delta t \rightarrow 0$, on obtient :

$$\dot{P}_{\text{on on}}(t) = -R_+ P_{\text{on on}} + R_- P_{\text{on off}} , \quad (4.87)$$

³⁶On appelle *shelving* le phénomène de piégeage dans cet état.

$$\dot{P}_{\text{on off}}(t) = -R_- P_{\text{on off}} + R_+ P_{\text{on on}} . \quad (4.88)$$

Lorsque le processus a atteint sa phase stationnaire³⁷, $P_{\text{on on}}(t=0) = \mathcal{P}_{-\text{eq}} = \frac{R_-}{R_+ + R_-}$, d'après (4.70) ; par ailleurs, $P_{\text{on off}}(t=0) = 0$. La solution du système (4.87, 4.88) avec ces conditions initiales est :

$$P_{\text{on on}}(t) = \frac{R_-^2}{(R_+ + R_-)^2} + \frac{R_+ R_-}{(R_+ + R_-)^2} e^{-(R_+ + R_-)t} . \quad (4.89)$$

Le report de cette expression dans (4.83) donne :

$$C(t) = \sigma_I^2 e^{-(R_+ + R_-)t} \quad (t > 0) , \quad (4.90)$$

avec :

$$\langle I \rangle = \frac{R_-}{R_+ + R_-} I_0 , \quad \sigma_I^2 \equiv \langle I^2 \rangle - \langle I \rangle^2 = \frac{R_+ R_-}{(R_+ + R_-)^2} I_0^2 . \quad (4.91)$$

Lorsqu'un processus aléatoire est stationnaire, toutes les fonctions d'autocorrélation sont paires ; l'expression (4.90) est donc vraie $\forall t$ à condition d'y remplacer t par $|t|$. Dans ces conditions, le spectre $S(\omega)$ des fluctuations d'intensité est une fonction à valeurs réelles :

$$S(\omega) \equiv \frac{1}{2} \int_{-\infty}^{+\infty} dt C(t) e^{i\omega t} = \int_0^{+\infty} dt C(t) \cos \omega t = \frac{\sigma_I^2 (R_+ + R_-)}{\omega^2 + (R_+ + R_-)^2} . \quad (4.92)$$

C'est une lorentzienne de largeur $\Delta\omega = R_+ + R_-$.

La possibilité d'observer un très petit nombre d'atomes ou ions a ainsi permis de mettre à jour les mécanismes élémentaires imaginés par Bohr. En outre, notamment pour des raisons méthodologiques, elle a ouvert une nouvelle voie théorique construite sur la notion de fonction d'onde stochastique [27]. En effet, la description usuelle des processus optiques dissipatifs consiste à écrire une équation maîtresse (équations de Bloch optiques) pour haïlt l'opérateur-densité relatif à un ensemble statistique d'atomes, bien approché en pratique par une vapeur contenant un très grand nombre d'entités élémentaires *indépendantes*. En-dehors des difficultés techniques à surmonter pour obtenir la solution d'une telle équation³⁸, il est bien clair que ce formalisme est très mal adapté au cas des petits amas et, à la limite, au cas de l'atome unique – on commence par calculer des propriétés d'ensemble pour essayer d'en déduire des comportements individuels. Un traitement plus direct est donc visiblement souhaitable.

Le choix de se cantonner à un formalisme “purent mécanique” pose le problème de la description de l'irréversibilité (dissipation). Comme déjà mentionné, celle-ci apparaît naturellement quand on couple un petit système à un grand système, et est *a contrario* exclue dans un formalisme hamiltonien standard concernant un petit nombre de degrés de liberté et en l'absence de toute trace partielle. S'en tenir à une description en terme de fonction d'onde passe donc par la réconciliation de deux concepts antagonistes, en forçant la mécanique à être irréversible. Une façon d'induire l'irréversibilité est de ponctuer la dynamique par des opérations de réduction de paquet d'ondes, c'est-à-dire d'effectuer par la pensée des mesures idéales, entrant dans la catégorie des *gedanken experiments*.

Ces idées vont maintenant être illustrées dans le cas [27] d'un atome à deux niveaux couplé au champ électromagnétique quantifié, l'atome étant par ailleurs piloté par une onde monochromatique de pulsation ω , traitée classiquement et représentée par $\mathcal{E}_0 \cos \omega_L t$. Pour simplifier, on supposera que le champ quantifié peut être soit dans l'état fondamental $|0\rangle$ (vide de photons), soit dans un état à un photon d'impulsion $\hbar\vec{k}$ et de polarisation ε , état noté $|\hbar\vec{k} \varepsilon\rangle$. Les deux états atomiques “nus” sont notés $|g\rangle$ et $|e\rangle$ et ont pour énergies E_g et $E_e = E_g + \hbar\omega_0$. Le champ classique fait osciller l'atome à *la Rabi* ; en définissant le zéro d'énergie à la valeur E_g , le Hamiltonien effectif décrivant l'oscillation de Rabi est³⁹ :

$$H_{\text{Rabi}} = \hbar(\omega_0 - \omega) |e\rangle\langle e| + \frac{\hbar\Omega}{2} (|e\rangle\langle g| + |g\rangle\langle e|) , \quad (4.93)$$

³⁷où, notamment, toutes les fonctions de corrélation $\langle A(t)A(t') \rangle$ ne dépendent que de la *différence* des temps.

³⁸Pour un système à N niveaux, l'opérateur densité est représenté par une matrice $N \times N$; il faut donc calculer et engranger environ N^2 quantités réelles.

³⁹On s'en tient à l'approximation de l'onde tournante (*Rotating Wave Approximation*, acronyme RWA), où les termes anti-résonnants en $e^{\pm 2i\omega_L t}$ sont négligés.

où $\Omega = -d\mathcal{E}_0$, d étant le moment dipolaire électrique de l'atome (Ω est la fréquence de Rabi) ; $\delta = \omega - \omega_0$ est le désaccord entre le laser et la pulsation atomique. Quand $\delta = 0$, le Hamiltonien (4.93) induit bien un mouvement où les probabilités pour l'atome d'être dans ses deux états alternent en prenant successivement dans le temps les deux valeurs maximales 0 et 1.

Supposons qu'à l'instant t , le système complet (atome + champ quantifié) est dans l'état factorisé :

$$|\Psi(t)\rangle = (a_g|g\rangle + a_e|e\rangle) \otimes |0\rangle \equiv |\Phi(t)\rangle \otimes |0\rangle . \quad (4.94)$$

Maintenant, entre t et $t + \delta t$ ⁴⁰, il y a possibilité d'émission spontanée de photons. La fonction d'onde à l'instant $t + \delta t$ peut se décomposer en la somme de deux termes, chacun d'entre eux correspondant à un nombre donné de photons du champ quantifié :

$$|\Psi(t + \delta t)\rangle = |\Psi^{(0)}(t + \delta t)\rangle + |\Psi^{(1)}(t + \delta t)\rangle \quad (4.95)$$

avec⁴¹ :

$$|\Psi^{(0)}(t + \delta t)\rangle = (a'_g|g\rangle + a'_e|e\rangle) \otimes |0\rangle \quad |\Psi^{(1)}(t + \delta t)\rangle = |g\rangle \otimes \sum_{\vec{k}, \varepsilon} \beta_{\vec{k}, \varepsilon} |\vec{k}, \varepsilon\rangle . \quad (4.96)$$

Les expressions précises de ces états peuvent être obtenues de diverses façons (à la Wigner - Weisskopf, résolvante dans l'approximation du pôle, Hamiltonien effectif non-hermitique, etc.). Le carré de la norme de $|\Psi^{(1)}\rangle$ est la probabilité δp d'émission d'un photon entre t et $t + \delta t$, soit $\delta p = \Gamma_e |a_e|^2 \delta t$; l'intervalle de temps δt est supposé court au sens où $\delta p \ll 1$. Par la conservation de la probabilité totale, il en résulte que $|\langle \Psi^{(0)} | \Psi^{(0)} \rangle|^2 = 1 - \delta p$. Les amplitudes a'_g et a'_e s'obtiennent en faisant évoluer l'atome avec $H \equiv H_{\text{Rabi}} - i\hbar \frac{\Gamma_e}{2} |e\rangle\langle e|$; enfin, $\beta_{\vec{k}, \varepsilon}$ est l'amplitude de probabilité d'émission d'un photon d'impulsion $\hbar\vec{k}$ et de polarisation ε . Au total, l'évolution du système total est unitaire (la norme de $|\Psi(t + \delta t)\rangle$ est la même, par construction, que celle de $|\Psi(t)\rangle$).

C'est maintenant que s'introduit l'irréversibilité induite par l'"observation". On suppose que tous les photons émis spontanément sont détectés avec un compteur parfait, qui mesure le nombre de photons dans le champ à l'instant $t + \delta t$; tout photon ainsi mesuré est absorbé par le compteur et disparaît donc du champ électromagnétique. Conformément au postulat de réduction du paquet d'onde, si le résultat est 0, on doit projeter l'état juste avant la mesure, $|\Psi(t + \delta t)\rangle$, sur $|\Psi^{(0)}(t + \delta t)\rangle$, si au contraire on trouve 1, on projette sur $|\Psi^{(1)}(t + \delta t)\rangle$; dans tous les cas, il faut renormaliser l'état ainsi obtenu puisque qu'une deuxième mesure effectuée immédiatement après la première doit donner le même résultat avec probabilité 1.

Dans un schéma numérique, le caractère aléatoire de l'émission spontanée pendant le petit intervalle δt est très facile à singer : on se fixe un seuil $\eta \ll 1$ et on tire une suite de nombres au hasard r_i répartis uniformément, entre 0 et 1 par exemple. $0 < r_i \leq \eta$ correspond à l'émission d'un photon, alors que $\eta < r_i < 1$ correspond à la non-émission. Au total, après cette mesure, l'alternative est la suivante :

- soit $r_i < \eta$ auquel cas :

$$|\Psi(t + \delta t)\rangle = |g\rangle \otimes \sum_{\vec{k}, \varepsilon} \beta_{\vec{k}, \varepsilon} |\vec{k}, \varepsilon\rangle , \quad (4.97)$$

- soit $r_i > \eta$ auquel cas⁴² :

$$|\Psi(t + \delta t)\rangle = \mu (a'_g|g\rangle + a'_e|e\rangle) \otimes |0\rangle = \mu (\mathbf{1} - i\hbar H \delta t) |\Phi(t)\rangle \otimes |0\rangle , \quad (4.98)$$

avec $\mu = (1 - \delta p)^{-\frac{1}{2}}$. En tout cas, à l'issue de chaque mesure, on en revient à un état factorisé de la forme (4.94), où le champ est dans l'état vide, et le cycle peut être recommencé. Dans ce schéma, l'évolution de la fonction d'onde est une marche au hasard dans l'espace des états atomiques, d'où le nom de fonction d'onde stochastique.

⁴⁰ δt est supposé très petit devant Γ_e^{-1} , durée de vie de l'état excité, δ^{-1} et Ω^{-1} .

⁴¹ Dans le secteur à zéro photon, l'atome peut être soit en $|g\rangle$, soit en $|e\rangle$ (transitions virtuelles sans évacuation du photon). Dans le secteur à 1 photon, l'atome est dans l'état fondamental, puisqu'un photon a été émis vers l'extérieur. Par ailleurs, la déséxcitation suivie d'une réexcitation pendant δt est un processus à deux étapes qui est d'ordre δt^2 .

⁴² Après mesure, le photon émis a été extrait du système par le compteur et le champ est bien à nouveau dans l'état vide.

La question se pose de l'équivalence entre l'approche stochastique et la description conventionnelle à l'aide des équations de Bloch optiques. Dalibard *et al.* [27] ont également montré que la moyenne d'ensemble de l'opérateur densité construit sur la fonction d'onde stochastique satisfait une équation d'évolution identique à celle de Bloch. D'ailleurs, les simulations numériques [27] montrent qu'avec une centaine de tirages (d'histoires), les résultats obtenus par l'approche stochastique sont quasiment indiscernables de ceux fournis par les équations de Bloch. Même sans rentrer dans une comparaison détaillée des deux procédures, il est clair que la description stochastique est bien moins coûteuse. La situation est analogue à celle que l'on rencontre dans la théorie conventionnelle du mouvement Brownien ; alors l'alternative est la suivante : ou bien on résout une équation aux dérivées partielles (équation de Fokker - Planck) pour la fonction de densité⁴³ – ce qui permet d'avoir accès à toutes les moyennes en tant qu'espérances mathématiques –, ou bien on simule numériquement la trajectoire brownienne, ce qui est en principe très facile, et on moyenne sur un (grand) nombre de trajectoires (histoires). Dans tous les cas, la dynamique moyennée est foncièrement irréversible : qu'elle se produise dans l'espace physique ou dans un espace plus abstrait, une marche au hasard présente toujours un comportement où la symétrie de renversement du temps est brisée⁴⁴.

4.3 Cryptographie quantique

La cryptographie est l'art – vieux comme le monde⁴⁵ – de communiquer au moyen de conventions assurant que le message transmis est à l'abri des indiscrétions. L'idée la plus naturelle est d'utiliser un *codage* des messages confidentiels : toute lettre de l'alphabet est remplacée biunivoquement par un autre signe⁴⁶ (nombre – d'où l'expression message chiffré –, image, symbole exotique, *etc.*). Le lexique donnant la correspondance biunivoque s'appelle la clé, et c'est elle qui doit rester absolument confidentielle. Ce procédé porte précisément le nom de *cryptographie à clé secrète*⁴⁷.

L'élément crucial est la confidentialité de la clé permettant de coder et de décoder le message transmis, étant entendu que celui-ci peut à la limite être rendu public. Tout le problème est donc d'assurer la sécurité du message initial définissant la clé.

Toute transmission classique d'information peut être interceptée passivement, c'est-à-dire sans modification de la teneur du message ; bien sûr, il peut y avoir d'énormes difficultés techniques à surmonter pour y parvenir, mais c'est en principe possible : l'écoute d'une ligne téléphonique perturbe en un sens la ligne – donc ouvre la voie à la révélation de l'écoute – mais on peut réduire cet effet révélateur pour rendre l'écoute quasiment indétectable. Au contraire, parce qu'une mesure quantique perturbe en général⁴⁸ le système objet de la mesure,

⁴³définie dans l'espace des phases, en général.

⁴⁴Essentiellement, la brisure de symétrie provient du fait que l'on s'intéresse à un petit nombre de degrés de liberté, relatifs à un sous-système du système considéré. Dans le cas de la particule brownienne, on regarde sa position, sa vitesse, *etc.*, étant entendu que le "bain" – qui provoque la marche au hasard – est complètement évacué d'emblée et n'est présent que par son effet erratique sur la marche. Un autre exemple est fourni par l'amortissement de Landau dans un plasma : l'équation de départ (de Vlasov, où le terme de collisions est absent, au contraire de l'équation de Boltzmann) est réversible et produit une densité $f(\vec{r}, \vec{p}, t)$ permettant de calculer les valeurs moyennes. Si l'on s'intéresse seulement à la densité de particules, il convient d'intégrer sur les impulsions, ce qui fait "perdre" une grande quantité d'information ; l'irréversibilité qui en découle traduit une évidence : une fois effectuée la sommation sur les vitesses, il n'est plus possible de reconstruire le mouvement, il y a une infinité de conditions initiales susceptibles de donner la même densité en position. Sommer, intégrer ou laisser dans l'ombre certains degrés de liberté en les confiant au hasard produit ici la flèche du temps.

⁴⁵On raconte que Jules César y a recouru : était-ce pour correspondre avec Cléopâtre ?

⁴⁶Ce procédé de chiffrement est visiblement le plus simple qui soit. On sait faire beaucoup mieux, par exemple : on code une lettre en langage binaire, lui associant une suite $\{x_i\}$ de 0 ou de 1, laquelle n'a rien de confidentiel (tout le monde peut prendre la même) ; puis on tire au hasard une autre suite $\{y_i\}$ de 0 ou de 1, qui constitue la clé à proprement parler et qui, elle, est confidentielle. Le chiffrement de la lettre choisie est constitué par la suite des sommes $\{x_i + y_i \text{ modulo } 2\}$.

⁴⁷Il existe aussi des procédés dits à *clé publique*, l'idée étant que certaines opérations mathématiques s'effectuent trivialement dans un sens et exigent un travail colossal dans l'autre (pour cette raison, on parle de *fonctions à sens unique*) : par exemple, il est facile de multiplier deux nombres premiers, alors que la factorisation d'un (grand) nombre N en ses premiers entiers exige un temps de calcul croissant en gros exponentiellement avec N . Typiquement, factoriser un nombre de mille chiffres en ses facteurs premiers est, avec les ordinateurs actuels, totalement inconcevable.

C'est d'ailleurs au sujet de ce type de problème que l'ordinateur quantique repointe le bout du nez, compte tenu des capacités phénoménales que l'on en attend...

⁴⁸La mesure ne provoque pas *toujours* une "perturbation incontrôlable", comme on le voit parfois écrit. Si l'état du système est, avant la mesure, propre de l'observable mesurée, il l'est encore après : en pareil cas, la mesure est "non-destructive".

l'observation indésirable peut être en principe facilement détectée. La Mécanique Quantique permet ainsi de construire des procédures permettant de savoir si, lors d'une transmission, on est écouté ou non. Il s'agit ici de se borner à donner quelques idées de base, étant entendu qu'il existe une littérature⁴⁹ énorme sur le sujet, où divers scénarios sont envisagés et décrits⁵⁰.

Dans la suite, on considère un premier exemple, formel, posant les idées essentielles montrant comment la Mécanique quantique permet de concevoir des protocoles révélant *à coup sûr* la présence d'une écoute indésirable. Deux exemples utilisant soit un atome à trois niveaux, soit des spins $\frac{1}{2}$ sont ensuite exposés en détail. Dans un dernier temps, on examine le point le plus essentiel, à savoir comment transmettre la clé en toute confidentialité, suivant la proposition d'Eckert [29].

4.3.1 Principes de la détection d'une écoute indésirable

L'idée de base est simple et met en jeu des séquences de mesures successives d'observables qui ne commutent pas. De façon un peu abstraite, les choses se présentent comme suit. Soit A et B deux observables incompatibles :

$$[A, B] \neq 0, \quad (4.99)$$

dont l'une, A pour fixer les idées, est une constante du mouvement :

$$[A, H] = 0 \quad (4.100)$$

alors que B n'en est pas une ($[B, H] \neq 0$). À un certain instant, $t = 0$, on mesure A , et on trouve a_m , l'une des valeurs propres de A ; on sait d'avance que la probabilité d'obtenir cette valeur est égale à $|\langle a_m | \Psi(t=0) \rangle|^2$. Juste après cette mesure, l'état du système est $|\Psi(t=0+0)\rangle = |a_m\rangle$. Si aucune autre mesure n'est effectuée⁵¹ jusqu'à l'instant t , une nouvelle mesure de A à cet instant donne le résultat a_m , cette fois avec certitude. En effet, l'état à l'instant t est $|\Psi(t)\rangle = U(t)|\Psi(t=0+0)\rangle = U(t)|a_m\rangle$, mais comme A et H commutent, l'état $|a_m\rangle$ produit par la réduction du paquet d'onde consécutive à la première mesure de A , est également⁵² propre de H ; il en résulte que $|\Psi(t)\rangle = e^{-\frac{i}{\hbar}E_m t} |a_m\rangle$ et que la deuxième mesure de A produit la même valeur a_m avec probabilité 1. Si les deux mesures sont effectuées par deux observateurs distincts (Alice puis Bob), ces derniers doivent donc constater que leurs résultats sont toujours identiques, quand la même séquence est répétée un grand nombre de fois, à la limite infini.

Imaginons maintenant au contraire que, entre la première mesure de A et la seconde, à l'instant t_1 ($0 < t_1 < t$), une tierce personne mesure B , qui n'est pas une constante du mouvement. Cet "espion" trouve l'une des valeurs propres de B , b_n , et l'état issu de cette mesure est⁵³ $|\Psi(t=t_1+0)\rangle = |b_n\rangle$. Il ne s'agit pas d'un état stationnaire, puisque H et B ne commutent pas ; à l'instant ultérieur t , cet état est devenu :

$$|\Psi(t)\rangle = \sum_{m'} c_{m'} e^{-\frac{i}{\hbar}E_{m'}(t-t_1)} |a_{m'}\rangle, \quad c_{m'} = \langle a_{m'} | b_n \rangle, \quad t_1 < t. \quad (4.101)$$

Dans ces conditions, la mesure de A à l'instant t postérieur à t_1 peut alors produire *l'une quelconque* des valeurs propres $a_{m'}$ – la probabilité de trouver $a_{m'}$ est $P_{m'} = |c_{m'}|^2$; le résultat brut est que les résultats obtenus par Alice et Bob effectuant les deux mesures de A ne sont plus nécessairement identiques. En fait, lorsque la séquence est répétée un grand nombre de fois, il suffit que *l'un* de leurs résultats soit différent pour qu'ils puissent affirmer avec certitude qu'ils sont victimes d'une écoute et que la sécurité de leur transmission n'est pas assurée.

⁴⁹La littérature a pris l'habitude d'appeler Alice et Bob les deux personnes communiquant entre elles et souhaitant se mettre à l'abri des malversations d'un "espion". Ce dernier est souvent appelé Eve (vient de *to eavesdrop* = *écouter aux portes* (!)).

⁵⁰Pour une bibliographie assez récente, voir [28].

⁵¹Dans ces conditions, le système évolue librement par l'équation de Schrödinger.

⁵²Pour simplifier la discussion, on suppose pour l'instant qu'il n'existe aucune dégénérescence.

⁵³On continue à supposer qu'il n'existe pas de dégénérescence. L'exemple 1 ci-dessous présente un cas dégénéré.

4.3.2 Exemples

Exemple 1 : trois observables

Précisons les idées précédentes dans un cas concret, où d'ailleurs le Hamiltonien présente de la dégénérescence. Soit un atome à trois niveaux $\{|p\rangle\}$ (un fondamental $p = 0$ et deux niveaux excités dégénérés, $p = 1, 2$) :

$$H = \hbar\omega(-|0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2|) . \quad (4.102)$$

A est une observable donnée par :

$$A = a(|0\rangle\langle 0| + i|1\rangle\langle 2| - i|2\rangle\langle 1|) \quad (a \in \mathbb{R}) . \quad (4.103)$$

A commute visiblement avec H , c'est donc une constante du mouvement. Ses valeurs propres sont $+a$ (dégénérescence 2) et $-a$ (non-dégénérée). Les vecteurs propres de A (qui sont aussi propres de H sont les $|\psi_{\alpha\beta}\rangle$ ($\alpha, \beta = \pm$) :

$$|+a, 1\rangle = |0\rangle \equiv |\psi_{-+}\rangle , \quad |+a, 2\rangle = \frac{1}{\sqrt{2}}(|1\rangle - i|2\rangle) \equiv |\psi_{++}\rangle . \quad (4.104)$$

$$|-a\rangle = \frac{1}{\sqrt{2}}(-i|1\rangle + |2\rangle) \equiv |\psi_{+-}\rangle . \quad (4.105)$$

Le premier indice de $|\psi_{\alpha\beta}\rangle$ renvoie à la valeur propre $\pm\hbar\omega$ de H , le second au signe de la valeur propre $\pm a$ de A .

L'observable B est prise sous la forme :

$$B = b(|0\rangle\langle 1| + |1\rangle\langle 0| + |2\rangle\langle 2|) \quad (b \in \mathbb{R}) ; \quad (4.106)$$

clairement :

$$[A, B] \neq 0 , \quad [H, B] \neq 0 . \quad (4.107)$$

Les valeurs propres de B sont $+b$ (dégénérescence 2) et $-b$ (non-dégénérée) ; les vecteurs propres sont :

$$|+b, 1\rangle = |2\rangle , \quad |+b, 2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) , \quad |-b\rangle = \frac{1}{\sqrt{2}}(-|0\rangle + |1\rangle) . \quad (4.108)$$

Maintenant, Alice et Bob effectuent des mesures de A aux instants 0 et t . Compte tenu de la dégénérescence en énergie, l'opération de mesure peut porter sur un ensemble complet d'observables qui commutent, ce que constituent A et H ; à l'issue d'une telle opération, l'état du système est parfaitement défini. Si aucune mesure de B n'est faite entre-temps, les résultats des deux observateurs coïncident dans tous les cas.

En revanche, soit une mesure de B effectuée à l'instant t_1 compris entre 0 et t . Pour fixer les idées, on suppose que le premier observateur a trouvé $+\hbar\omega$ et $+a$; l'état de départ est donc $|\Psi(0+)\rangle = |\psi_{++}\rangle$. Compte tenu des décompositions ci-dessus, l'état à l'instant t_1 est :

$$|\Psi(t_1)\rangle = e^{-i\omega t_1} \frac{1}{\sqrt{2}} \left[\frac{1}{\sqrt{2}}(|+b, 2\rangle + |-b\rangle) - i|+b, 1\rangle \right] . \quad (4.109)$$

Il en résulte que la mesure à t_1 donne $+b$ avec la probabilité $\frac{3}{4}$, $-b$ avec la probabilité $\frac{1}{4}$. Dans le premier cas, l'état à $t_1 + 0$ est :

$$|\Psi(t_1 + 0)\rangle = \frac{\sqrt{2}}{\sqrt{3}} \left[\frac{1}{\sqrt{2}}|+b, 2\rangle - i|+b, 1\rangle \right] . \quad (4.110)$$

Un calcul simple montre que cet état devient, à l'instant t :

$$|\Psi(t)\rangle = \frac{\sqrt{2}}{\sqrt{3}} \left[\frac{1}{2} e^{+i\omega(t-t_1)} |\psi_{-+}\rangle + \frac{3}{2\sqrt{2}} e^{-i\omega(t-t_1)} |\psi_{++}\rangle - \frac{i}{2\sqrt{2}} e^{-i\omega(t-t_1)} |\psi_{+-}\rangle \right] . \quad (4.111)$$

Pour cette “histoire”, le deuxième observateur va trouver les valeurs $(-\hbar\omega, +a)$, $(+\hbar\omega, +a)$ et $(+\hbar\omega, -a)$ avec les probabilités respectives $\frac{1}{6}$, $\frac{3}{4}$ et $\frac{1}{12}$, au lieu de trouver à coup sûr le même résultat que le premier observateur.

Si la mesure intermédiaire donne $-b$, on a successivement :

$$|\Psi(t_1 + 0)\rangle = |-b\rangle, \quad (4.112)$$

qui devient :

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}} e^{+i\omega(t-t_1)} |\psi_{-+}\rangle + \frac{1}{2} e^{-i\omega(t-t_1)} |\psi_{++}\rangle + \frac{i}{2\sqrt{2}} e^{-i\omega(t-t_1)} |\psi_{+-}\rangle. \quad (4.113)$$

Le deuxième observateur va trouver les valeurs $(-\hbar\omega, +a)$, $(+\hbar\omega, +a)$ et $(+\hbar\omega, -a)$ avec les probabilités respectives $\frac{1}{2}$, $\frac{1}{4}$ et $\frac{1}{4}$.

Ainsi, au coup par coup, Alice et Bob peuvent trouver des résultats différents. La probabilité qu'ils trouvent des résultats identiques est une probabilité conditionnelle :

$$P_{\text{résultats identiques}} \equiv P(++; t | ++; t = 0), \quad (4.114)$$

qui peut s'exprimer suivant la relation de chaîne habituelle⁵⁴ :

$$P(++; t | ++; 0) = \sum_{\varepsilon=\pm 1} P(++; t | \varepsilon b; t_1) P(\varepsilon b; t_1 | ++; 0). \quad (4.115)$$

Cette dernière expression est la somme⁵⁵ sur toutes les “trajectoires” possibles, une mesure intermédiaire ayant été faite⁵⁶. On a donc ici :

$$P(++; t | ++; 0) = \frac{3}{4} \frac{3}{4} + \frac{1}{4} \frac{1}{4} = \frac{5}{8} \quad (4.116)$$

Le choix d'un autre résultat pour la première mesure du couple (H, A) conduit au même type de conclusion : en raison de la mesure effectuée par un tiers, la probabilité que les résultats soient identiques est loin d'être égale à 1 ; le calcul montre qu'elle vaut $\frac{1}{2}$ si la première mesure a donné $(-\hbar\omega, +a)$ et $\frac{5}{8}$ si on a d'abord trouvé $(+\hbar\omega, -a)$. Dans tous les cas, au bout de quelques expériences, Alice et Bob peuvent avec une quasi-certitude affirmer qu'ils sont écoutés⁵⁷.

Exemple 2 : Alice, Bob et des spins $\frac{1}{2}$

Une autre procédure utilise l'envoi de spins $\frac{1}{2}$ par Alice à Bob⁵⁸. Le spin $\frac{1}{2}$ est une variable binaire qui se prête donc bien au codage : on peut mettre en correspondance biunivoque l'écriture binaire d'un nombre avec une configuration de spins $+-+-+-\dots+$. Ici, pour un Hamiltonien à symétrie sphérique, toutes les composantes S_u des spins sont des constantes du mouvement ; les observables incompatibles en jeu sont bien évidemment les différentes composantes cartésiennes de \vec{S} , satisfaisant les relations caractéristiques d'un moment cinétique :

$$\vec{S} \times \vec{S} = i\hbar \vec{S} \iff [S_x, S_y] = i\hbar S_z, \text{ etc.} \quad (4.117)$$

⁵⁴relation dite de Bachelier - Chapman - Kolmogorov.

⁵⁵la somme représente l'addition des probabilités d'événements mutuellement exclusifs.

⁵⁶Lorsqu'aucune mesure intermédiaire n'est faite, la probabilité est le module au carré d'une *somme d'amplitudes*. Ici, compte tenu de la réduction liée à la mesure de B à t_1 , la probabilité est une *somme de probabilités*, chacune de celles-ci étant le module carré d'une amplitude.

⁵⁷Toutes les probabilités exhibées ne dépendent pas du temps. En effet :

- quand il s'agit de mesures de A , toutes les réductions de paquet d'onde produisent des états stationnaires
- en ce qui concerne la mesure de B : l'état sur lequel on mesure est propre de A et H , c'est encore un état stationnaire et donc les probabilités de trouver les valeurs $\pm b$ sont indépendantes du temps. Enfin, l'état issu de la mesure de B à t_1 est $|b, i = 1, 2\rangle$ ou $|-b\rangle$. Lors de la mesure de A et H à $t > t_1$, les probabilités sont de la forme $|\langle \psi_{\alpha\beta} | U(t - t_1) | b, i = 1, 2 \rangle|^2$ ou $|\langle \psi_{\alpha\beta} | U(t - t_1) | -b \rangle|^2$ (voir (4.104) et (4.105)) ; U agissant sur le bra produit une simple phase temporelle qui disparaît du module carré.

⁵⁸D'autres versions sont possibles ; par exemple, on peut disposer d'une source fabriquant des paires de spins d'état bien défini, l'un des spins étant envoyé vers Alice, l'autre vers Bob ; un espion mesurant le spin allant vers Bob va induire des différences entre les résultats des mesures des deux observateurs.

En vertu de la relation $\vec{S} = \frac{\hbar}{2}\vec{\sigma}$, on peut raisonner avec les trois composantes de $\vec{\sigma}$. Par ailleurs, puisque les composantes du spin sont des constantes du mouvement, seule une mesure intermédiaire d'une composante incompatible est susceptible de modifier, lors d'une mesure ultérieure, la valeur trouvée au départ.

Alice choisit de mesurer la composante suivant Oz , trouve donc à chaque fois ± 1 et envoie le spin mesuré à Bob. Celui-ci mesure la composante de \vec{S} le long d'un axe OZ faisant l'angle θ avec l'axe Oz d'Alice et situé dans le plan xOz (Alice et Bob sont d'avance convenus d'une même orientation de leurs repères locaux). L'observable que mesure Bob est donc :

$$\sigma_Z = \cos \theta \sigma_z + \sin \theta \sigma_x, \quad [\sigma_z, \sigma_Z] = 2i \sin \theta \sigma_y \quad (4.118)$$

Bien évidemment, Alice et Bob ne vont pas trouver les mêmes résultats, puisqu'ils mesurent deux observables qui ne commutent pas – sauf si $\theta = 0$. Les états propres de σ_Z , qui a aussi pour valeurs propres ± 1 , sont :

$$|+\rangle_Z = \cos \frac{\theta}{2} |+\rangle_z + \sin \frac{\theta}{2} |-\rangle_z, \quad |-\rangle_Z = -\sin \frac{\theta}{2} |+\rangle_z + \cos \frac{\theta}{2} |-\rangle_z, \quad (4.119)$$

comme on le vérifie facilement. Ceci étant donné, on voit que si Alice a trouvé $+1$ – et envoie donc à Bob un spin dans l'état $|+\rangle_z$, celui-ci peut trouver $+1$ ou -1 , avec les probabilités respectives $\cos^2 \frac{\theta}{2}$ et $\sin^2 \frac{\theta}{2}$. Bien sûr, si $\theta = 0$, Alice et Bob trouvent *toujours* les mêmes résultats. Si Bob mesure suivant un axe tourné de $\frac{\pi}{2}$ (appelons-le Ox , $OZ_{\theta=\pi/2} = Ox$), alors :

$$|+\rangle_x = \frac{1}{\sqrt{2}}(|+\rangle_z + |-\rangle_z), \quad |-\rangle_x = \frac{1}{\sqrt{2}}(-|+\rangle_z + |-\rangle_z); \quad (4.120)$$

dans ces conditions, si Alice a trouvé $+$, Bob a une chance sur deux de trouver $+$, *etc.* : Bob ne trouve comme Alice *que* dans 50% des cas.

Maintenant, un espion situé sur le trajet du spin procède à des mesures suivant un axe faisant l'angle ϕ avec Oz ; l'observable que mesure l'espion est donc :

$$\sigma_e = \cos \phi \sigma_z + \sin \phi \sigma_x, \quad (4.121)$$

observable dont les vecteurs propres sont :

$$|+\rangle_e = \cos \frac{\phi}{2} |+\rangle_z + \sin \frac{\phi}{2} |-\rangle_z, \quad |-\rangle_e = -\sin \frac{\phi}{2} |+\rangle_z + \cos \frac{\phi}{2} |-\rangle_z. \quad (4.122)$$

Pour fixer les idées, on raisonne après une mesure par Alice ayant donné $+1$, de sorte que l'espion effectue une mesure sur l'état $|+\rangle_z$: il trouve donc $+1$ avec la probabilité $\cos^2 \frac{\phi}{2}$ et -1 avec la probabilité $\sin^2 \frac{\phi}{2}$. Puis Bob procède à la mesure de σ_Z ; si l'espion a trouvé $+1$, Bob trouve $+1$ ou -1 avec les probabilités respectives $\cos^2 \frac{\theta-\phi}{2}$ et $\sin^2 \frac{\theta-\phi}{2}$. Si l'espion a trouvé -1 , Bob trouve -1 ou $+1$ avec les probabilités respectives $\cos^2 \frac{\theta-\phi}{2}$ et $\sin^2 \frac{\theta-\phi}{2}$. En définitive, la probabilité pour que Alice et Bob aient le même résultat $+1$ est donc :

$$P_\phi(+, \theta|+, 0) = \cos^2 \frac{\phi}{2} \cos^2 \frac{\theta-\phi}{2} + \sin^2 \frac{\phi}{2} \sin^2 \frac{\theta-\phi}{2}. \quad (4.123)$$

En particulier, si Alice et Bob sont convenus de mesurer le long du même axe Oz ($\theta = 0$), la probabilité de coïncidence de leurs résultats est :

$$P_\phi(+, 0|+, 0) = \cos^4 \frac{\phi}{2} + \sin^4 \frac{\phi}{2} = \frac{1}{2}(1 + \cos^2 \phi). \quad (4.124)$$

Au total, pour faire parvenir un message de n bits, Alice doit envoyer Nn spins, N étant pris très grand de sorte que les fréquences statistiques puissent être confondues avec les lois limites de probabilités. Bob peut retourner à Alice tous ses spins, par une voie dont la confidentialité a été établie antérieurement ; le produit Nn étant forcément très grand – et la probabilité (4.124) étant évidemment plus petite que 1 (sauf si l'espion connaît l'angle choisi par Alice et Bob) –, les deux observateurs vont très vite s'apercevoir s'ils sont écoutés ou non.

L'espion, ne connaissant pas la direction commune d'observation d'Alice et Bob, peut décider de tirer au hasard la sienne, en prenant l'angle ϕ uniformément distribué entre 0 et 2π . La probabilité de coïncidence moyennée est :

$$P(+, 0|+, 0) = \int_0^{2\pi} \frac{d\phi}{2\pi} P_\phi(+, 0|+, 0) = \frac{3}{4} . \quad (4.125)$$

Cette modification ne change pas grand chose au fait que sur un grand nombre d'envois, l'espion va être démasqué à coup presque sûr.

4.3.3 Communication de la clé secrète entre Alice et Bob

Les exemples traités ci-dessus illustrent la possibilité de savoir si la ligne est écoutée ou non, mais n'en définissent pas pour autant un mode de transmission sûr de la clé de cryptage : si être en mesure d'affirmer avec certitude que la ligne n'est pas écoutée est déjà un immense progrès, encore faut-il à un moment ou à un autre trouver les moyens d'échanger la clé elle-même. Ekert ([29], [30]) a proposé le scénario qui suit, exposé d'abord dans sa version simple, puis dans une version raffinée utilisant des paires de spin couplées *à la* EPR.

Dans sa version la plus simple, Alice envoie à Bob des spins polarisés. L'état de polarisation de chaque spin résulte d'une mesure faite par Alice et ayant produit un certain résultat $\pm \frac{\hbar}{2}$, noté simplement \pm dans la suite. Celle-ci tire au hasard sa direction de mesure ; pour simplifier, on suppose que cette direction est soit Ox (notée \leftrightarrow), soit Oz (notée \updownarrow). Donc, pour chaque spin, Alice

1. tire au hasard sa direction de mesure, Ox ou Oz
2. note son résultat, $+$ ou $-$
3. envoie le spin à Bob (sans commentaire !)

Une séquence possible est la suivante :

$$\begin{array}{cccccccccccccccccccc} \updownarrow & \updownarrow & \leftrightarrow & \updownarrow & \leftrightarrow & \leftrightarrow & \leftrightarrow & \updownarrow & \leftrightarrow & \leftrightarrow & \updownarrow & \updownarrow & \updownarrow & \leftrightarrow & \updownarrow & \leftrightarrow & \leftrightarrow & \updownarrow & \updownarrow & \leftrightarrow \\ + & - & - & - & + & + & - & + & - & - & + & + & - & + & + & - & + & - & - & + \end{array} . \quad (4.126)$$

Pour chaque spin reçu, Bob fait une mesure du spin en tirant lui aussi au hasard sa direction d'observation (il se trompe donc dans un cas sur deux) :

- quand il choisit le même axe qu'Alice, il trouve le même résultat qu'elle :

$$P(\text{Bob}, Oz, + | \text{Alice}, Oz, +) = 1 , \quad P(\text{Bob}, Oz, - | \text{Alice}, Oz, +) = 0 , \quad (4.127)$$

$$P(\text{Bob}, Oz, + | \text{Alice}, Oz, -) = 0 , \quad P(\text{Bob}, Oz, - | \text{Alice}, Oz, -) = 1 \quad (4.128)$$

- quand Bob "se trompe" d'axe, il ne trouve (statistiquement) le même résultat qu'Alice que dans la moitié des cas (voir (4.120)) :

$$P(\text{Bob}, Ox, + | \text{Alice}, Oz, +) = \frac{1}{2} , \quad P(\text{Bob}, Ox, - | \text{Alice}, Oz, +) = \frac{1}{2} . \quad (4.129)$$

$$P(\text{Bob}, Ox, + | \text{Alice}, Oz, -) = \frac{1}{2} , \quad P(\text{Bob}, Ox, - | \text{Alice}, Oz, -) = \frac{1}{2} . \quad (4.130)$$

et de même en permutant x et z . Compte tenu de la séquence d’Alice (4.126), un exemple de séquence pour Bob, *en l’absence d’espion*, est le suivant :

$$\begin{array}{cccccccccccccccccccc} \updownarrow & \leftrightarrow & \updownarrow & \updownarrow & \leftrightarrow & \leftrightarrow & \updownarrow & \leftrightarrow & \leftrightarrow & \updownarrow & \updownarrow & \leftrightarrow & \updownarrow & \leftrightarrow & \updownarrow & \updownarrow & \leftrightarrow & \updownarrow & \leftrightarrow & \updownarrow \\ + & + & - & - & + & + & + & + & - & - & + & - & - & + & + & - & + & - & - & + \end{array} . \quad (4.131)$$

Toutes les mesures étant faites, Alice fait part publiquement⁵⁹ à Bob de tous ses *choix* d’axes, mais, pour l’instant, ne dit évidemment rien de ses *résultats* ; réciproquement, Bob lui communique ses axes. Ceci permet d’éliminer tous les cas où le hasard a fait que les axes n’étaient pas les mêmes pour Alice et pour Bob :

Alice	\updownarrow	\updownarrow	\leftrightarrow	\updownarrow	\leftrightarrow	\leftrightarrow	\leftrightarrow	\updownarrow	\leftrightarrow	\leftrightarrow	\updownarrow	\updownarrow	\updownarrow	\leftrightarrow	\updownarrow	\leftrightarrow	\leftrightarrow	\updownarrow	\updownarrow	\leftrightarrow
Bob	\updownarrow	\leftrightarrow	\updownarrow	\updownarrow	\leftrightarrow	\leftrightarrow	\updownarrow	\leftrightarrow	\leftrightarrow	\updownarrow	\updownarrow	\updownarrow	\leftrightarrow	\updownarrow	\updownarrow	\leftrightarrow	\leftrightarrow	\updownarrow	\updownarrow	\leftrightarrow
Retenus	+			-	+	+				-	+			-	+	+			+	-

(4.132)

À ce stade (et toujours en l’absence d’espion), chacun sait que pour tous les *bits* ainsi retenus, il a dû trouver les mêmes résultats que l’autre⁶⁰, mais ne sait pas ce qu’il en est au juste (si c’est faux, un espion écoute !).

Pour le savoir, Alice communique maintenant à Bob *une partie* de ses résultats (elle les “sacrifie”) ; connaissant les choix d’axes et les résultats d’Alice pour un assez grand nombre de cas, un test statistique permet alors à Bob de savoir avec une faible probabilité d’erreur si la communication a été écoutée ou non. Si la confidentialité est alors établie, la séquence des *autres* résultats constituent la clé codant une lettre convenue d’avance : *sans rien communiquer d’autre*, Alice et Bob savent qu’ils ont trouvé la même chose à chaque fois puisqu’ils ont utilisé les mêmes axes. Finalement, comme la clé est constituée sans l’échange de rien d’autre, aucun moyen n’existe d’altérer la confidentialité de celle-ci !

Une autre version consiste à utiliser une source de particules produisant des paires (a, b) de spin $\frac{1}{2}$ dans l’état singulet :

$$|0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) , \quad (4.133)$$

développé sur les états propres de S_{z_a} et S_{z_b} . Après séparation des particules, l’une allant vers Alice, l’autre vers Bob, les deux observateurs procèdent à des mesures de la composante du spin le long de Ox ou de Oz , directions toutes deux perpendiculaires à la direction de propagation des particules. Dans la suite on note $|\rightarrow\rangle$ et $|\leftarrow\rangle$ respectivement les états propres $\pm\frac{\hbar}{2}$ de S_x . Compte tenu de :

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle) , \quad |\leftarrow\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle - |\downarrow\rangle) , \quad (4.134)$$

l’état singulet (4.133) s’écrit aussi :

$$|0\rangle = \frac{1}{\sqrt{2}} (|\leftarrow\rightarrow\rangle - |\rightarrow\leftarrow\rangle) . \quad (4.135)$$

Soit d’abord la séquence où Alice et Bob mesurent le long de la même direction, S_z par exemple. La première, Alice effectue une mesure, de S_z par exemple ; si elle trouve $+\frac{\hbar}{2}$, l’état des deux particules après cette mesure est la projection de (4.133), soit $|\uparrow\downarrow\rangle$. Il en résulte que quand Bob va mesurer S_z par la suite, il trouvera $-\frac{\hbar}{2}$ avec certitude⁶¹. Si Alice trouve $-\frac{\hbar}{2}$, Bob trouvera $+\frac{\hbar}{2}$: leurs résultats sont donc à coup sûr opposés. Il en va de même si ce sont des mesures de S_x qui sont effectuées : le même raisonnement vaut avec (4.135). En définitive, quand Alice et Bob mesurent la même composante, leurs résultats sont toujours “anticorrélés”. Physiquement et en résumé, sachant que l’état est singulet et qu’Alice a trouvé le spin dans un sens, Bob ne peut trouver que le spin opposé.

⁵⁹On entend par là que la communication ne nécessite aucune confidentialité.

⁶⁰Chacun sait ce qu’a trouvé l’autre sans y être allé voir et sans communication explicite !

⁶¹La mesure permet ainsi d’obtenir instantanément des informations sur ce qui se passe ailleurs : c’est ici que réside le paradoxe EPR (à ce sujet, on pourra consulter l’article de Jammer [31]).

Si Alice mesure S_z et Bob S_x , les choses vont autrement. Si Alice a trouvé $+\frac{\hbar}{2}$, l'état après cette mesure est $|\uparrow\downarrow\rangle$ et se décompose suivant :

$$\frac{1}{2}(|\rightarrow\rightarrow\rangle - |\rightarrow\leftarrow\rangle + |\leftarrow\rightarrow\rangle - |\leftarrow\leftarrow\rangle) . \quad (4.136)$$

Bob trouvera donc pour S_x les deux valeurs $\pm\frac{\hbar}{2}$ avec équiprobabilité. Il en va de même si Alice trouve $-\frac{\hbar}{2}$.

Ainsi, dans tous les cas, chacun sait (*a priori* sans communication des résultats : c'est le miracle EPR) ce que l'autre a trouvé. Le même scénario que ci-dessus peut alors être repris (communication des axes et d'une partie des résultats pour tester la confidentialité de la ligne).

Remarque

Il faut bien voir que l'espion éventuel ne peut pas extraire quoi que ce soit d'utile : dans la phase d'envoi, les spins ne véhiculent aucune information en eux-mêmes. En fait, l'information ne devient réalité qu'après la mise en œuvre du scénario décrit ci-dessus.

Partie II

Physique moléculaire

Chapitre 5

L'approximation de Born et Oppenheimer

Il s'agit de présenter l'approximation fondamentale sur laquelle repose la physique moléculaire et la physique de la matière condensée.

5.1 Ordres de grandeur

Une molécule (ou un solide) est un assemblage complexe d'électrons (particules peu massives, donc légères) et de noyaux (particules lourdes) en interaction mutuelle, édifice compliqué dont on ne sait évidemment pas trouver les états propres exacts. Le bien-fondé de l'approximation fondamentale décrite ci-dessous repose sur la grande différence d'inertie entre électrons et noyaux¹ et conduit à l'approximation dite de Born et Oppenheimer.

Le point central est en effet l'énorme différence des masses entre la masse électronique m et une masse nucléaire typique, M :

$$\frac{m}{M} \ll 1 . \quad (5.1)$$

D'un autre côté, toutes ces particules sont en interaction mutuelle par des forces qui sont toutes du même ordre de grandeur $\sim \frac{e^2}{a_0^2}$. Des forces comparables mais des inerties très différentes donnent classiquement des mouvements caractérisés par des échelles de temps dont les unes sont courtes et les autres longues : les particules massives se déplacent lentement, les particules légères vont vite. Intuitivement (et de façon imagée), on peut donc affirmer que les électrons ont un mouvement rapide autour des centres attracteurs que constituent les noyaux, lesquels se déplacent lentement.

Corrélativement, l'étude des ordres de grandeur de l'énergie d'une molécule permet de dégager simplement trois échelles typiques qui, par le jeu de $E = h\nu$, redonnent tout naturellement plusieurs échelles de temps $\sim \nu^{-1}$ très différentes. Une molécule possède deux sortes de degrés de liberté : les coordonnées électroniques $\{q_i\}$ et les coordonnées des noyaux $\{Q_j\}$. Tous les mouvements sont confinés et se produisent à l'intérieur d'un domaine de dimension atomique, de l'ordre de $2a_0 \simeq 1 \text{ \AA}$. Avec la seule masse de l'électron, l'énergie la plus simple que l'on peut former est $\frac{\hbar^2}{m(2a_0)^2}$ et c'est forcément l'ordre de grandeur de l'énergie associée au mouvement des électrons, E_{el} :

$$E_{\text{el}} \sim \frac{\hbar^2}{4ma_0^2} = -\frac{1}{2} E_{n=1} \sim 7 \text{ eV} \quad (5.2)$$

¹Il est souvent physiquement plus utile de faire une distinction entre certains électrons et ce que l'on appelle les cœurs ; ces derniers sont une sorte de souvenir tenace des constituants séparés (atomes, ions, ...) et sont eux-même un assemblage robuste d'électrons et de noyaux, peu sensibles en 1^{ère} approximation aux processus physico-chimiques ordinaires.

En ce qui concerne les noyaux, deux types de mouvement sont possibles : les mouvements de vibration et les mouvement de rotation en bloc².

Pour estimer l'énergie liée à la vibration, considérons une molécule diatomique *stable* AB, dont les noyaux sont séparés par la distance R . Physiquement, il doit être possible de définir une certaine fonction $E(R)$ représentant la variation d'énergie lorsque, toutes choses égales par ailleurs, on fait varier la longueur de la liaison AB. Elle a forcément l'allure donnée sur la figure 5.1 : fortement répulsive à courte distance (chacun des partenaires A et B résiste à l'invasion par l'autre de son domaine réservé), présentant un minimum en une certaine valeur R_0 et tendant vers une constante quand R tend vers l'infini – d'où l'identification d'une énergie de dissociation, E_{diss} . Expérimentalement, cette énergie est typiquement de l'ordre de l'eV, soit $E_{\text{diss}} \sim E_{\text{el}}$.

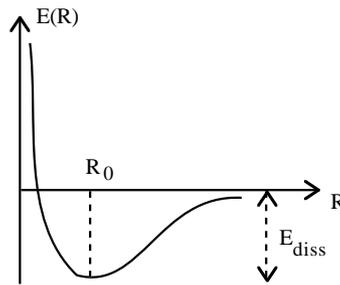


Figure 5.1: Variation typique de la fonction $E(R)$.

S'agissant toujours de dégager des ordres de grandeur, on fait l'hypothèse de petites amplitudes autour de R_0 , ce qui permet de confondre le creux de potentiel avec sa parabole osculatrice, en écrivant :

$$E(R) \simeq E(R_0) + \frac{1}{2} k(R - R_0)^2 \equiv E_{\text{harm}}(R) , \quad k = \left(\frac{d^2 E}{dR^2} \right)_{R=R_0} , \quad \frac{|R - R_0|}{R_0} \ll 1 . \quad (5.3)$$

Quand $R = 2R_0$, on a étiré la liaison de deux fois sa longueur et on peut raisonnablement admettre que l'augmentation d'énergie est du même ordre que l'énergie de dissociation :

$$E(2R_0) \sim E(R_0) + E_{\text{diss}} \iff \frac{1}{2} k R_0^2 \sim \frac{\hbar^2}{4ma_0^2} , \quad (5.4)$$

et comme $R_0 \sim 2a_0$, la constante de raideur est à peu près³ :

$$k \sim \frac{\hbar^2}{8ma_0^4} . \quad (5.5)$$

Par ailleurs, k est de la forme $M\omega_{\text{vibr}}^2$ où M est une masse nucléaire (par exemple, la masse réduite de deux noyaux), d'où :

$$\omega_{\text{vibr}} \sim \frac{\hbar}{(2a_0)^2 \sqrt{mM}} . \quad (5.6)$$

Finalement, l'énergie de vibration, $E_{\text{vibr}} = \hbar\omega_{\text{vibr}}$ est environ :

$$E_{\text{vibr}} \sim \frac{\hbar^2}{(2a_0)^2 \sqrt{mM}} . \quad (5.7)$$

Par comparaison avec (5.2), on voit que :

$$\frac{E_{\text{vibr}}}{E_{\text{el}}} \sim \frac{\hbar^2}{(2a_0)^2 \sqrt{mM}} \frac{(2a_0)^2}{\hbar^2} = \sqrt{\frac{m}{M}} \ll 1 ; \quad (5.8)$$

²On suppose évacué d'emblée le mouvement de translation d'ensemble de la molécule (translation uniforme par rapport à un repère galiléen), sans intérêt.

³Plus directement, on peut aussi écrire que $k = E''(R_0) \sim$ énergie électronique/carré d'une longueur $\sim \frac{\hbar^2}{m(2a_0)^2(2a_0)^2} = \frac{\hbar^2}{m(2a_0)^4}$.

il convient de noter au passage que si le rapport $\frac{m}{M}$ est en effet très petit, c'est sa racine carrée qui intervient ; elle peut ne pas être si petite que cela.

Enfin, trouvons l'ordre de grandeur de l'énergie de rotation, E_{rot} . Elle est du genre $\frac{J^2}{2I}$, où J est un moment cinétique et où I est un moment d'inertie typique, soit $I \sim M(2a_0)^2$. \hbar est l'ordre de grandeur fondamental pour un moment cinétique en Mécanique Quantique, donc $J \sim \hbar$, de sorte que :

$$E_{\text{rot}} \sim \frac{\hbar^2}{M(2a_0)^2} . \quad (5.9)$$

Par comparaison avec (5.7), il vient :

$$\frac{E_{\text{rot}}}{E_{\text{vibr}}} \sim \frac{\hbar^2}{M(2a_0)^2} \frac{(2a_0)^2}{\hbar^2} \sqrt{mM} = \sqrt{\frac{m}{M}} \ll 1 . \quad (5.10)$$

Ainsi, à chaque fois, c'est le petit paramètre $\varepsilon = \sqrt{\frac{m}{M}} \ll 1$ qui ressort et :

$$E_{\text{rot}} \sim \varepsilon E_{\text{vibr}} \sim \varepsilon^2 E_{\text{el}} \implies E_{\text{rot}} \ll E_{\text{vibr}} \ll E_{\text{el}} . \quad (5.11)$$

En associant sa pulsation à chaque énergie, on voit que la rotation est lente comparée à la vibration, laquelle est également lente par rapport à la rotation des électrons. En terme de transitions spectrales, si on accepte que les transitions électroniques sont dans l'UV ($\lambda \sim 1000 \text{ \AA}$) et en prenant $m/M \sim 10^{-4}$, on voit que la longueur d'onde d'une transition vibrationnelle est $\lambda_{\text{vibr}} \sim 10^{-7}/10^{-2} = 10^{-9} \text{ m}$, soit 10μ , donnant un spectre dans l'infrarouge. De la même façon, une longueur d'onde rotationnelle sera $\lambda_{\text{rot}} \sim 10 \mu/10^{-2} = 1000 \mu = 1 \text{ mm}$. En termes de fréquences ($\omega = c/\lambda$) :

$$1000 \text{ \AA} \longrightarrow \omega_{\text{el}} = \frac{3 \times 10^8}{10^{-7}} = 3 \times 10^{15} \text{ rd/s} \iff \nu_{\text{el}} \simeq 5 \times 10^{14} \text{ Hz} \longleftrightarrow 2 \text{ eV} , \quad (5.12)$$

et :

$$\omega_{\text{vibr}} = 3 \times 10^{13} \text{ rd/s} \iff \nu_{\text{vibr}} \simeq 5 \times 10^{12} \text{ Hz} \longleftrightarrow 0.02 \text{ eV} , \quad (5.13)$$

$$\omega_{\text{rot}} = 3 \times 10^{11} \text{ rd/s} \iff \nu_{\text{rot}} \simeq 5 \times 10^{10} \text{ Hz} = 50 \text{ Ghz} \longleftrightarrow 0.2 \text{ meV} . \quad (5.14)$$

5.2 Approximation de Born et Oppenheimer

Compte tenu de l'analyse précédente des ordres de grandeurs, la différence de rapidité des mouvements des électrons et des noyaux suggère de traiter dans une première étape le seul mouvement des électrons, les noyaux étant pris carrément immobiles, quitte à introduire ultérieurement le mouvement lent de ces derniers.

Dans l'approximation électrostatique, le Hamiltonien d'une molécule s'écrit :

$$H = \sum_j \frac{\vec{P}_j^2}{2M_j} + \sum_i \frac{\vec{p}_i^2}{2m} + V(\{q_i\}, \{Q_j\}) \equiv T_N + H_0(p_i, q_i; Q_j) . \quad (5.15)$$

V contient toutes les interactions de Coulomb (électrons - électrons, électrons - noyaux, noyaux - noyaux). La première étape consiste à radicaliser la situation en posant $M_j = +\infty$ dans (5.15), ce qui revient à dire que les noyaux sont fixes. Notons que H_0 ne contient pas d'opérateurs différentiels par rapport aux Q_j , de sorte que :

$$[H_0, Q_j] = 0 \quad \forall j . \quad (5.16)$$

5.2.1 Étape 1 : mouvement des électrons dans le champ des noyaux fixes

Fixer les noyaux, c'est finalement mettre de côté le terme T_N , et en être réduit à ne devoir chercher que les états propres de H_0 . Dans cet opérateur, les p_i et les q_i sont des opérateurs au sens usuel de la Mécanique Quantique ; au contraire, les coordonnées nucléaires Q_i sont de simples paramètres. L'équation aux valeurs propres de H_0 en représentation- q s'écrit précisément :

$$H_0 \Phi(q_i; Q_j) = E(Q_j) \Phi(q_i; Q_j) . \quad (5.17)$$

Les coordonnées nucléaires Q_j paramétrisent tout autant les fonctions propres que les valeurs propres. H_0 contient notamment le potentiel attractif des noyaux pour les électrons et, tout comme pour un atome, l'application de conditions aux limites (normalisabilité des fonctions propres des états liés) produit la quantification spontanée des énergies propres. Chaque mode propre est donc repérable par un jeu de nombres quantiques, collectivement notés n . L'équation (5.17) peut donc être écrite plus précisément :

$$H_0 \Phi_n(q_i; Q_j) = E_n(Q_j) \Phi_n(q_i; Q_j) . \quad (5.18)$$

Résoudre effectivement ce type d'équation est évidemment toute une histoire – admettons cette étape franchie, formellement. En tout état de cause, les Φ_n sont orthogonales, et supposées normalisées à l'unité :

$$\int \prod_i dq_i \Phi_n^*(q_i; Q_j) \Phi_{n'}(q_i; Q_j) = \delta_{nn'} , \forall Q_j , \quad (5.19)$$

où les Q_j sont les *mêmes*⁴ dans Φ_n et $\Phi_{n'}$.

5.2.2 Étape 2 : inclusion du mouvement des noyaux

En admettant que les fonctions $\{\Phi_n\}$ ainsi définies constituent une base complète, on peut *a priori* développer la fonction d'onde totale $\Psi(q_i, Q_j)$ représentant la molécule dans sa globalité sous la forme :

$$\Psi(q_i, Q_j) = \sum_n \chi_n(Q_j) \Phi_n(q_i; Q_j) . \quad (5.20)$$

Les $\chi_n(Q_j)$, qui constituent les coefficients du développement, ne dépendent que des coordonnées nucléaires. Le report d'une telle expression dans :

$$H\Psi \equiv (T_N + H_0)\Psi = E\Psi \quad (5.21)$$

fait apparaître différents termes.

Dans le membre de gauche de (5.21), le premier terme (énergie cinétique cinétique des noyaux) contient des dérivées secondes ($\vec{P} = -i\hbar\vec{\nabla}_Q$) ; en utilisant $(fg)'' = f''g + 2f'g' + fg''$, il devient :

$$\begin{aligned} T_N \chi_n(Q_j) \Phi_n(q_i; Q_j) &= \sum_k \frac{1}{2M_k} \left[(\vec{P}_k^2 \chi_n) \Phi_n + 2(\vec{P}_k \chi_n) (\vec{P}_k \Phi_n) + \chi_n \vec{P}_k^2 \Phi_n \right] \\ &= (T_N \chi_n) \Phi_n + \sum_k \frac{1}{M_k} (\vec{P}_k \chi_n) (\vec{P}_k \Phi_n) + \chi_n (T_N \Phi_n) . \end{aligned} \quad (5.22)$$

Le second terme au premier membre de (5.21) contient les termes du genre :

$$H_0 \chi_n(Q_j) \Phi_n(q_i; Q_j) = \chi_n(Q_j) H_0 \Phi_n(q_i; Q_j) = E_n(Q_j) \chi_n(Q_j) \Phi_n(q_i; Q_j) , \quad (5.23)$$

⁴Sinon on aurait deux Hamiltoniens distincts (pour deux configurations fixes des noyaux), dont les fonctions propres n'ont aucune raison d'être orthogonales.

où a été utilisé le fait que les Q_j dans H_0 n'interviennent que dans des opérateurs multiplicatifs (pas d'opérateurs cinétiques, voir (5.16)) et que, de ce fait $H_0\chi_n = \chi_n H_0$. Compte tenu de tout ceci et après quelques réarrangements, (5.21) devient :

$$\sum_n [(T_N\chi_n)\Phi_n + E_n(Q_j)\chi_n\Phi_n] = \sum_n \left[E\chi_n\Phi_n - \sum_k \frac{1}{M_k} (\vec{P}_k\Phi_n)(\vec{P}_k\chi_n) - (T_N\Phi_n)\chi_n \right] . \quad (5.24)$$

En multipliant maintenant par Φ_m^* , en intégrant sur tous les q_i et en tenant compte de (5.19), il vient :

$$T_N\chi_m(Q_j) + E_m(Q_j)\chi_m(Q_j) = E\chi_m(Q_j) - \sum_n \Lambda_{mn}\chi_n(Q_j) , \quad (5.25)$$

où Λ_{mn} est un opérateur contenant les coordonnées et les moments conjugués des noyaux :

$$\Lambda_{mn} = \int \prod_i dq_i \sum_n \Phi_m^* \left[\sum_k \frac{1}{M_k} (\vec{P}_k\Phi_n)\vec{P}_k + (T_N\Phi_n) \right] . \quad (5.26)$$

L'équation (5.25) peut s'écrire en notation de Dirac, en convenant que les *brackets* sous-entendent l'intégration sur tous les q_i et en considérant les coefficients $\chi_m(Q_j)$ comme des fonctions d'onde $\langle Q_j|\chi_m\rangle$:

$$[T_N(Q_j) + E_m(Q_j)]|\chi_m\rangle = E|\chi_m\rangle - \sum_n \left(\langle \Phi_m|T_N|\Phi_n\rangle + \sum_k \langle \Phi_m|\frac{\vec{P}_k}{M_k}|\Phi_n\rangle \cdot \vec{P}_k \right) |\chi_n\rangle . \quad (5.27)$$

Ceci peut encore s'écrire plus formellement :

$$[T_N(Q_j) + E_m(Q_j)]|\chi_m\rangle = E|\chi_m\rangle - \sum_n \Lambda_{mn}|\chi_n\rangle . \quad (5.28)$$

Clairement, Λ_{mn} constitue la matrice d'un certain opérateur représentant les couplages cinétiques entre les états électroniques obtenus dans la 1^{ère} étape de l'approximation de Born et Oppenheimer, et induisant *de facto* une dépendance entre les coefficients $\chi_m(Q_j)$ du développement (5.20).

5.2.3 Approximation adiabatique

Les équations (5.25) - (5.28) sont en principe exactes, dans l'hypothèse où les $\{\Phi_n\}$ forment une base complète. L'approximation dite adiabatique⁵ consiste à laisser tomber purement et simplement tous les couplages Λ_{mn} et à ne retenir que :

$$[T_N + E_m(Q_j)]\chi_m(Q_j) \simeq E\chi_m(Q_j) . \quad (5.29)$$

Alors, ou bien $\chi_m(Q_j) = 0$, ou bien $\chi_m(Q_j)$ est l'une des fonctions propres de l'équation (5.29), le nombre quantique m étant fixé. En d'autres termes, le développement (5.20) se réduit alors à un seul terme, celui associé au m choisi et la fonction d'onde totale est un simple produit :

$$\Psi(q_i; Q_j) \simeq \chi_m(Q_j)\Phi_m(q_i; Q_j) \equiv \Psi_m^{\text{adiab}}(q_i; Q_j) . \quad (5.30)$$

La justification détaillée de cette approximation n'est pas aisée. Elle signifie, physiquement, que le mouvement des électrons se fait à m donné, c'est-à-dire que le mouvement (lent, quasi-statique) des noyaux intervient de fait dans les états électroniques comme une simple variation *paramétrique* de chacun d'entre eux : on néglige de ce fait les transitions induites d'un état Φ_n à un autre état $\Phi_{n'}$; dit autrement, les électrons s'adaptent instantanément à la configuration lentement variable des noyaux, d'où le nom d'approximation adiabatique. D'un

⁵ Il existe aussi un *théorème adiabatique*, dû à Gell-Mann et Low[35], stipulant que quand on "branche" infiniment lentement une perturbation V , un système initialement dans un état propre de H_0 se retrouve dans un état propre de $H = H_0 + V$ à l'issue du branchement (état qui n'est d'ailleurs pas forcément le fondamental). Une démonstration peut être trouvée dans l'ouvrage de Fetter et Walecka[36].

autre côté, la fonction nucléaire obéit à l'équation (5.29) où la dynamique des électrons n'apparaît plus explicitement (les coordonnées électroniques q_i n'y figurent pas, elles ont été "moyennées" sous la forme $\langle \Phi_m | H_0 | \Phi_m \rangle$) : les noyaux voient en fait le mouvement *moyen* des électrons, en conformité avec l'idée que ceux-ci vont très vite.

La lenteur du mouvement nucléaire assure en général qu'une fréquence de vibration ω_{vibr} (et *a fortiori* une fréquence de rotation ω_{rot}) est toujours très petite devant ω_{el} , de sorte que les différents états électroniques, même munis de leur structure fine vibrationnelle et rotationnelle, restent le plus souvent bien séparés en énergie. *A contrario*, cette remarque permet de prévoir que l'approximation adiabatique est sûrement en défaut lorsque deux énergies E_n et E_m en viennent à se croiser quand les coordonnées nucléaires varient. Clairement, au voisinage du point de croisement, on ne peut faire l'économie d'un traitement plus raffiné du genre théorie de perturbation pour deux niveaux quasi-dégénérés, impliquant au moins un mélange des deux états Φ_n et Φ_m – à moins bien sûr que les éléments de matrice de Λ_{mn} soient nuls, pour des raisons de symétrie par exemple. Lorsque tel n'est pas le cas, on assiste comme toujours en pareil cas à un croisement de niveaux *évité* où la dégénérescence est levée⁶. Il en résulte des changements de propriétés physiques de la molécule (distorsion, modification des spectres, altération des mécanismes de réaction).

Hormis cette dernière situation, donnant lieu à des effets observables⁷, la validation de l'approximation adiabatique est finalement fournie par l'explication remarquable des spectres moléculaires, tant dans le domaine UV, qu'IR ou micro-ondes.

5.3 Discussion

En résumé, dans l'approximation adiabatique, la fonction d'onde totale de la molécule est de la forme :

$$\Psi(q_i; Q_j) = \chi_m(Q_j) \Phi_m(q_i; Q_j) , \quad (5.31)$$

avec :

$$H_0 \Phi_m = E_m(Q_j) \Phi_m , \quad (5.32)$$

et :

$$[T_N + E_m(Q_j)] \chi_m(Q_j) = E \chi_m(Q_j) . \quad (5.33)$$

Φ_m est la fonction d'onde des électrons, les noyaux étant fixés, E_m est la valeur propre correspondante, qui dépend précisément des Q_j . χ_m décrit le mouvement des noyaux et satisfait une équation aux valeurs propres où la valeur propre électronique $E_m(Q_j)$ joue finalement le rôle d'une énergie potentielle effective, représentant les interactions entre noyaux "habillées" par le mouvement moyen des électrons. L'énergie propre E figurant dans (5.33), en tant que valeur propre (scalaire), est l'énergie totale de la molécule, une fois prise la moyenne (au sens quantique) sur les variables électroniques. L'équation pour χ_m contient l'opérateur cinétique des noyaux, E représente donc bien l'énergie totale de la molécule (électrons et noyaux), dans l'approximation ainsi définie.

5.3.1 À propos du mouvement des électrons

Les états électroniques (Φ_m, E_m) dépendent des positions nucléaires, un fait qui va être illustré dans le cas d'une molécule diatomique AB. \vec{R} désignant le vecteur \vec{AB} , ces états ne dépendent évidemment que de la distance $R = AB$ (isotropie de l'espace). E_m est donc une fonction de R , qui a forcément une limite quand $R \rightarrow +\infty$ (sans que l'on sache précisément suivant quelle loi cette limite est atteinte) ; par ailleurs, elle tend vers l'infini quand $R \rightarrow 0$ (dans le cas contraire, la molécule serait instable aux courtes distances), traduisant la résistance de chaque noyau (ou chaque cœur) l'invasion par l'autre de son domaine personnel. Ceci étant, deux cas sont essentiellement possibles :

⁶Une situation analogue se présente pour un atome en champ magnétique, \mathcal{B} . En champ très faible, on perturbe les niveaux $|nJM_JLS\rangle$, ce qui produit des corrections d'énergie linéaires par rapport à \mathcal{B} . Pour deux multiplets voisins $|nJM_JLS\rangle$ et $|nJ'M_J'LS\rangle$, l'extrapolation en champ intermédiaire ou fort conduit inévitablement à des croisements de niveaux, pour lesquels il convient de reprendre le traitement de perturbation (exemple : les deux premiers multiplets $^2P_{1/2}$ et $^2P_{3/2}$ de l'atome d'hydrogène).

⁷Les plus connus portent le nom d'effet Renner, effet Jahn - Teller, effet Landau - Zener.

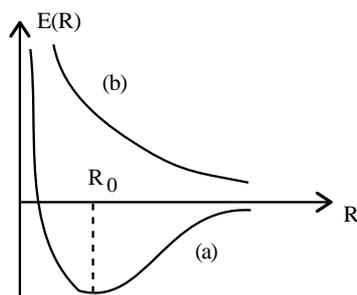


Figure 5.2: Variation de l'énergie électronique pour un état stable (a) ou instable (b) d'une molécule diatomique. Les valeurs des énergies en $R = +\infty$ ne sont pas forcément les mêmes.

- $E_m(R)$ présente un minimum pour une valeur R_0 (voir fig. 5.2, (a)), ce qui signifie que la molécule est stable ; l'état Φ_m est alors dit *liant*. Si, en outre, E_m est la plus petite valeur propre électronique de (5.32), R_0 s'interprète comme la longueur de la liaison chimique AB. L'état ainsi décrit est stable, et il faut une énergie E_{diss} pour la dissocier en deux constituants séparés : c'est l'énergie de la liaison chimique. Si E_m n'est pas la plus petite énergie propre de (5.32), R_0 est la longueur de liaison (usuellement étirée) lorsque la molécule est dans un état excité stable vis-à-vis de la configuration nucléaire mais évidemment instable radiativement et/ou par conversion interne.
- E_m n'a pas de *minimum* (voir fig. 5.2, (b)) ; l'état Φ_m est alors dit *antiliant*. Ceci signifie que pour cet état électronique la molécule n'a pas d'état stable vis-à-vis de la configuration nucléaire et se dissocie spontanément. Ce type d'état est l'intermédiaire obligé de la photodissociation. Si E_m est la plus petite valeur propre de (5.32), l'absence de minimum signifie que cet édifice moléculaire est une espèce éphémère.

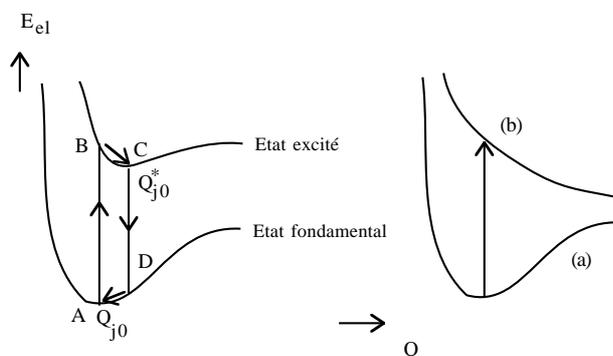


Figure 5.3: À gauche : illustration schématique de l'absorption/émission d'une molécule ; l'absorption (AB) est suivie de l'émission (CD). Les petites flèches représentent la relaxation *via* les niveaux de rotation et vibration (non-représentés). À droite : photodissociation par excitation radiative d'un état stable (a) à un état instable (b).

Ce sont les transitions entre états électroniques qui donnent les grands traits du spectre de la molécule dans l'optique ou l'UV (fig. 5.3 à gauche). Comme les mouvements nucléaires sont lents, la transition est quasi-verticale (les Q_j ne changent pas pendant l'absorption, flèche AB). Lorsque l'état excité a une durée de vie radiative assez longue, une relaxation vibrationnelle peut se produire⁸, la molécule à l'état (électronique) excité se retrouvant dans la configuration d'équilibre caractérisée par Q_{j0}^* . L'émission radiative qui survient tôt

⁸À condition bien sûr que la molécule ne soit pas isolée. En phase vapeur, ce sont les collisions qui permettent cette relaxation vibrationnelle, avec la complicité éventuelle des degrés de liberté de rotation. Le décalage entre absorption et émission est l'un des

ou tard⁹ se fait à nouveau verticalement (flèche CD) ; après relaxation dans l'état fondamental, la molécule retrouve la configuration Q_{j0} . Il en résulte qu'absorption et émission ne se produisent pas au même endroit, cette dernière donnant un pic décalé vers le rouge par rapport à l'absorption. En outre, les intensités sont en général assez différentes, en raison de l'intervention des produits scalaires des états de vibration de départ et d'arrivée (facteurs dits *de Franck - Condon*).

De la même façon, les états électroniques interviennent dans les processus de photodissociation : sous l'effet d'une irradiation optique ou UV, la molécule est portée de l'état fondamental (stable) à un état excité instable (fig. 5.3 à droite).

5.3.2 À propos du mouvement des noyaux

En ce qui concerne les mouvements nucléaires, une molécule à N noyaux possède $3N$ degrés de liberté. 3 d'entre eux sont associés à la translation en bloc, 3 autres représentent la rotation en bloc. Ces mouvements "solides" n'affectent pas le mouvement interne, de sorte que la valeur propre électronique E_m ne dépend en fait que de $3N - 6$ coordonnées nucléaires¹⁰. Si R_j ($1 \leq j \leq 3N - 6$ ou $3N - 5$) désigne un degré de liberté de mouvement interne (associé à une *déformation* de la molécule), on peut écrire, au voisinage d'un minimum absolu de E_m survenant pour les valeurs R_{j0} :

$$\left[T_N + E_m(R_{j0}) + \frac{1}{2} \sum_j k_{jj'} (R_j - R_{j0}) (R_{j'} - R_{j'0}) + \dots \right] \chi_m(R_j) = E \chi_m(R_j) , \quad (5.34)$$

où les \dots représentent des termes anharmoniques. En s'en tenant aux seuls termes écrits (approximation harmonique), on fait apparaître les petites vibrations des noyaux autour des positions d'équilibre. La forme quadratique se diagonalise en introduisant les bonnes combinaisons linéaires X_λ des coordonnées R_j , appelées modes normaux de vibration représentant des déformations *collectives* du squelette moléculaire. La détermination des vibrations normales est grandement facilitée par l'usage de la Théorie des groupes, plus particulièrement par la Théorie de la représentation linéaire des groupes : l'usage systématique de la symétrie permet, comme toujours, de prédiagonaliser la forme quadratique. Par ailleurs, de nouvelles constantes de raideur K_λ apparaissent, qui s'expriment naturellement à l'aide des $k_{jj'}$.

Le terme T_N contient tous les moments conjugués (il y en a $3N$), 6 (ou 5) d'entre eux étant associés aux mouvements en bloc. Les 3 moments conjugués qui représentent la translation en bloc ne présentent pas d'intérêt ; les 3 (ou 2) autres, associés à la rotation¹¹, contiennent les moments d'inertie, calculables *a priori* pour n'importe quelle configuration, pas seulement la configuration d'équilibre R_{j0} . Les $3N - 6(5)$ impliquant les modes normaux associés à la forme quadratique $\frac{1}{2} \sum_\lambda K_\lambda X_\lambda^2$ (avec K_λ strictement positif) donnent $3N - 6(5)$ oscillateurs harmoniques indépendants.

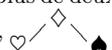
La variation des moments d'inertie avec l'état de vibration des liaisons induit un couplage entre la vibration et la rotation. Cet effet est souvent assez petit, et peut être omis dans un premier temps ; lorsqu'il est mis de côté, la rotation se découple de la vibration, les moments d'inertie étant alors naturellement définis relativement à la configuration d'équilibre. Dans ces conditions, l'énergie totale de la molécule¹² se décompose en une *somme* de trois termes, la simple somme résultant du découplage des degrés de liberté :

$$E = E_m(R_{j0}) + E_{\text{vibr}} + E_{\text{rot}} . \quad (5.35)$$

facteurs rendant difficile la résonance optique pour une molécule (pour un atome, l'écart est essentiellement dû au recul du noyau, le plus souvent petit par rapport à la largeur naturelle – pas d'effet Mossbauer pour une transition optique !)

⁹en l'absence d'autres processus de conversion interne.

¹⁰Pour une molécule linéaire, il n'y a en fait que $3N - 5$ degrés de liberté internes. En effet, l'un des moments d'inertie est nul, et il ne reste que deux moments d'inertie donnant une énergie de rotation. L'un des degrés de liberté est non pertinent, c'est l'angle représentant la rotation autour de l'axe de la molécule. En pareil cas, il y a toujours 3 degrés de translation, mais seulement 2 de rotation.

Notons que ce décompte repose sur le présupposé que la molécule reste linéaire quoi qu'il arrive. Pour une molécule à plus de deux noyaux, $\heartsuit - \diamond - \spadesuit$, ceci veut dire que toute déformation perpendiculaire à l'axe défini par la configuration d'équilibre, , possède une constante de raideur très élevée.

¹¹décrite à l'aide des angles d'Euler (voir [11], p. 447)

¹²dans le repère de son centre de masse.

$E_m(R_{j0})$ est la valeur de l'énergie électronique calculée dans la configuration d'équilibre ; E_{vibr} est l'énergie de l'ensemble des oscillateurs au-dessus du minimum d'énergie électronique ; enfin, E_{rot} est l'énergie de rotation, calculée avec les noyaux occupant leurs positions d'équilibre et donc *ipso facto* découplée de la vibration. Les variations de ces trois termes se produisent sur des échelles d'énergie très différentes, respectivement UV ou optique, IR et micro-ondes. La discussion détaillée des deux types de mouvement fait l'objet des deux chapitres suivants.

Chapitre 6

Structure électronique des molécules. Nature physique de la liaison chimique

Dans le chapitre précédent, on a vu comment il est possible de séparer les différents degrés de liberté d'un édifice complexe comme une molécule. Tout naturellement, l'étude des propriétés moléculaires commence par l'analyse des états électroniques, qui sont les plus énergétiques et qui déterminent les propriétés chimiques. Sur un plan conceptuel, c'est bien la Mécanique Quantique qui fournit une description rationnelle de l'origine de la liaison chimique et, par là, permet de comprendre la stabilité des molécules. Il s'agit ici d'exposer les idées principales – en prenant pour exemples les édifices moléculaires les plus simples – sans rentrer dans une zoologie quasi-inépuisable.

6.1 L'ion moléculaire H_2^+

La molécule la plus simple – au sens d'assemblage de plusieurs noyaux – est formée de deux protons avec juste ce qu'il faut d'électrons (en fait un seul) pour fabriquer un ciment susceptible de s'opposer à la répulsion électrostatique des deux protons nus. *A priori*, rien ne permet d'affirmer qu'un tel système existe à l'état stable : l'électron unique est tirailé entre les deux protons, ce qui introduit une interaction effective *attractive* entre ces derniers. La question de savoir si cette attraction est suffisante pour lutter contre la répulsion coulombienne entre protons ne peut être tranchée que par une étude quantitative, donnée ci-dessous dans le cadre de l'approximation adiabatique ; ici, la seule coordonnée nucléaire pertinente est la distance R entre les protons.

En tout état de cause, on voit de suite que c'est le *partage* d'un électron entre les protons qui est susceptible de produire une liaison stable, prototype de la liaison dite covalente. On rencontre ici un exemple du schéma universel représentant toute interaction physique : deux particules en interaction sont deux particules échangeant entre elles d'autres particules (ce sont ces dernières qui véhiculent l'échange d'"information" se traduisant par une interaction). Ici, les allers-retours de l'électron induisent une interaction *effective* attractive entre les protons, quelle que soit leur distance R ; complétée par la répulsion $\frac{e^2}{R}$, elle constitue l'énergie¹ $E(R)$ figurant dans l'équation nucléaire au titre d'une énergie potentielle. On note que si un état stable existe, cette interaction ne ressemble pas du tout à l'interaction nue des deux protons – qui se repoussent suivant la loi de Coulomb ("renormalisation" des interactions) : l'échange d'électron peut finalement créer une paire liée proton - proton².

¹Notée $E(Q_j)$ dans le chapitre précédent, l'énergie électronique ne dépend ici que d'une seule variable, R .

²*Mutatis mutandis*, le même scénario peut se produire dans un métal à basse température pour donner lieu à la supraconductivité : alors, ce sont les phonons du réseau (vibrations quantifiées) qui servent d'intermédiaires et permettent à deux électrons de former une paire liée électron - électron ; cette paire liée constitue un boson susceptible (avec ses congénères) de former un "superfluide" au sein d'une mer d'électrons "normaux".

6.1.1 Solution exacte

S'agissant de regarder les états électroniques tels que les définit l'approximation de Born - Oppenheimer, il faut d'abord résoudre l'équation aux valeurs propres où les noyaux sont fixes, séparés par la distance R . Dans ces conditions, il faut trouver les solutions de :

$$H_0\Phi = E\Phi, \quad (6.1)$$

H_0 étant le Hamiltonien électrostatique à noyaux fixes :

$$H_0 = \frac{\vec{p}^2}{2m} - \frac{e'^2}{r_1} - \frac{e'^2}{r_2} + \frac{e'^2}{R}. \quad (6.2)$$

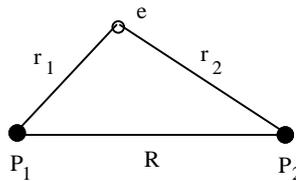


Figure 6.1: L'ion moléculaire H_2^+ .

Φ est une fonction de la position \vec{r} de l'électron (exprimée d'une façon ou d'une autre), paramétrée par la distance³ internucléaire R , soit $\Phi(\vec{r}; R)$. r_1 et r_2 désignent les distances de l'unique électron aux deux protons (voir fig. 6.1) ; noter que les r_i contiennent la distance R . Dans l'expression (6.2), $\frac{e'^2}{R}$ est une simple constante additive, qu'il ne faut cependant pas omettre : elle est une contribution importante à l'énergie totale qui joue un rôle décisif pour l'apparition d'un minimum d'énergie quand la distance internucléaire varie. L'étape importante est donc de fait la résolution de⁴ :

$$\left[\frac{\vec{p}^2}{2m} - \frac{e'^2}{r_1} - \frac{e'^2}{r_2} \right] \Phi(\vec{r}; R) = E_{\text{el}}(R)\Phi(\vec{r}; R), \quad (6.3)$$

étant entendu que l'énergie E définie en (6.1) est :

$$E(R) = E_{\text{el}}(R) + \frac{e'^2}{R}. \quad (6.4)$$

Il se trouve que pour ce système très simple, il est possible de résoudre exactement l'équation aux valeurs propres⁵. Le calcul complet étant très technique, on s'en tiendra à quelques éléments purement descriptifs.

Le problème a clairement une symétrie cylindrique ; l'angle ϕ entre le plan contenant l'électron et les noyaux, et un plan de référence arbitraire est une variable remarquable par son absence de H_0 (variable cyclique). De fait, si \vec{l} désigne le moment cinétique de l'électron par rapport à un point quelconque de l'axe internucléaire (choisi comme axe Oz), alors l_z est une constante du mouvement :

$$[H_0, l_z] = 0, \quad (6.5)$$

relation traduisant la symétrie cylindrique du système, et la variable ϕ se sépare. Les fonctions propres de $l_z = -i\hbar\frac{\partial}{\partial\phi}$ sont de la forme $e^{im\phi}$, avec m entier comme l'exige la théorie générale du moment cinétique orbital dès que les fonctions propres cherchées sont normalisables sur la sphère unité⁶. L'usage a consacré la notation

³Évidemment, les états propres de (6.1) ne dépendent pas de l'orientation de l'axe portant P₁ et P₂.

⁴Ici, il n'y a qu'un électron : la fonction Φ est donc *de facto* une orbitale ; comme elle est décrit un électron dans un assemblage de plusieurs noyaux, on l'appelle *orbitale moléculaire*. Dans (6.3), \vec{r} est (par exemple) le rayon-vecteur de l'électron en prenant l'origine au milieu de P₁P₂.

⁵Burrau était un spécialiste du problème des trois corps en Mécanique. L'article [33] consacré à H₂⁺ est donné dans les références.

⁶Le caractère entier de m (et de l pour \vec{L}^2) est une conséquence directe de l'algèbre des composantes du moment cinétique (orbital), $[l_u, l_v] = i\hbar\epsilon_{uvw}l_w$, et du fait que l'on manipule des éléments de matrice impliquant la norme des états propres ; ceux-ci n'ont de sens que si cette norme est finie. Le même scénario – et les mêmes prérequis – est à l'œuvre dans la quantification des états propres de l'oscillateur harmonique à l'aide des opérateurs a et a^\dagger .

$\sigma, \pi, \delta, \dots$, suivant que $|m| = 0, 1, 2, \dots$, en réminiscence de la notation s, p, d, ... pour les états de l'atome d'hydrogène. Par ailleurs, H_0 étant invariant par renversement du temps⁷ – opération qui inverse le moment cinétique – on peut affirmer d'emblée que l'énergie propre associée à Φ – qui dépend évidemment de m puisqu'il n'y a pas de symétrie sphérique comme dans un atome – ne dépend en fait que de $|m|$, pas du signe de m .

Compte tenu de la géométrie du système, les coordonnées elliptiques s'imposent :

$$\xi = \frac{r_1 + r_2}{R}, \quad 1 \leq \xi < +\infty, \quad \mu = \frac{r_1 - r_2}{R}, \quad -1 \leq \mu \leq +1; \quad (6.6)$$

ce choix étant fait, tous les états propres Φ peuvent être cherchés sous la forme⁸ :

$$\Phi = F(\xi, \mu) e^{im\phi} \quad (m \in \mathbb{Z}), \quad (6.7)$$

où F est une fonction à déterminer. Le Laplacien en coordonnées elliptiques est :

$$\Delta = \frac{4}{R^2(\xi^2 - \mu^2)} \left[(\xi^2 - 1) \frac{\partial^2}{\partial \xi^2} + 2\xi \frac{\partial}{\partial \xi} - (\mu^2 - 1) \frac{\partial^2}{\partial \mu^2} - 2\mu \frac{\partial}{\partial \mu} + \left(\frac{1}{\xi^2 - 1} - \frac{1}{\mu^2 - 1} \right) \frac{\partial^2}{\partial \phi^2} \right]; \quad (6.8)$$

la séparation (6.7) remplace simplement $\frac{\partial^2}{\partial \phi^2}$ par $-m^2$. L'équation aux fonctions propres pour $F(\xi, \mu)$ s'écrit ainsi :

$$-\frac{\hbar^2}{mR^2} \left[(\xi^2 - 1) \frac{\partial^2}{\partial \xi^2} + 2\xi \frac{\partial}{\partial \xi} - (\mu^2 - 1) \frac{\partial^2}{\partial \mu^2} - 2\mu \frac{\partial}{\partial \mu} - m^2 \left(\frac{1}{\xi^2 - 1} - \frac{1}{\mu^2 - 1} \right) \right] F(\xi, \mu) - \frac{e'^2}{R} \left(2\xi - \frac{1}{2}(\xi^2 - \mu^2) \right) F(\xi, \mu) = \frac{E}{2} (\xi^2 - \mu^2) F(\xi, \mu); \quad (6.9)$$

Visiblement, les deux variables ξ et μ se séparent aussi dans (6.9) ; sans perte de généralité, on peut donc poser :

$$F(\xi, \mu) = f(\xi) g(\mu); \quad (6.10)$$

le report dans (6.9) montre que chacune des fonctions f et g satisfait une équation différentielle associée de Legendre⁹ [32], paramétrée par m^2 (et R), dont les solutions sont bien connues et répertoriées. Comme toujours en Mécanique Quantique, il convient de faire le tri parmi toutes les solutions possibles, pour ne retenir que celles qui sont physiquement acceptables. Il s'agit ici d'obtenir les états liés, donc normalisables. Cette condition agit comme un crible et ne retient qu'un nombre infini dénombrable de solutions associés à des énergies discrètes (quantifiées) $\{E_{n,\text{el}}(R)\}_{n \in \mathbb{N}^*}$, toujours dépendantes¹⁰ de R .

Cette solution exacte du problème ne donne pas d'expressions simples, ni pour les fonctions f et g (ce sont des séries de polynômes de Legendre¹¹ – l'existence d'une série en l traduisant le fait que le carré du moment cinétique orbital n'est pas une constante du mouvement (brisure de la symétrie sphérique)), ni pour les énergies $E_{n,\text{el}}(R)$. On peut néanmoins tracer numériquement la variation de ces dernières en fonction de la distance internucléaire R . Les énergies $E_n(R)$ définies par (6.4) pour les premiers états sont données dans l'ouvrage de Slater [34]. Il apparaît deux types d'états, stables et instables (ce sont les états *liants* et *antiliants* déjà introduits), pouvant d'ailleurs se croiser¹². Les deux états de plus basse énergie forment une paire liante/antiliante, ce que l'on peut comprendre par un argument peu rigoureux de théorie de perturbation :

⁷ H_0 ne contient que le *carré* de la vitesse.

⁸ La fonction F est paramétrée par m , nombre qui apparaît explicitement dans les équations pour les fonctions f et g introduites ci-dessous ; cette dépendance n'est toutefois pas explicitée pour ne pas alourdir les notations.

⁹ L'équation associée de Legendre est la suivante ([1], p. 419) :

$$\left[(1 - u^2) \frac{d^2}{du^2} - 2u \frac{d}{du} + l(l + 1) - \frac{m^2}{1 - u^2} \right] L(u) = 0 \quad (m \in \mathbb{Z}). \quad (6.11)$$

Cette équation a notamment des solutions polynômiales (quand l est entier), usuellement notées P_l^m , de degré l , apparaissant déjà dans les solutions propres de \bar{l}^2 (associées à $e^{im\phi}$, ce sont essentiellement les harmoniques sphériques Y_{lm}).

¹⁰ R est contenu dans les coordonnées elliptiques.

¹¹ Par exemple, $g(\mu) = \sum_{l \geq m} \alpha_l P_l^{l-m}(\mu)$, où α_l est un certain coefficient.

¹² au voisinage des points de croisement, l'approximation adiabatique doit être reconsidérée – et peut se révéler d'ailleurs encore valide si, pour des raisons de symétrie par exemple, les deux états restent découplés.

si $R = +\infty$, chaque niveau est dégénéré deux fois, l'électron pouvant se trouver soit sur un proton, soit sur l'autre. Quel que soit R fini¹³, les deux éléments diagonaux de H sont encore égaux par symétrie, mais l'élément non-diagonal n'est plus nul : c'est lui qui va lever la dégénérescence, à peu près symétriquement tant que R n'est pas trop petit ; on obtient deux niveaux distincts de nature opposée, formés avec les deux combinaisons linéaires \pm des deux seuls états atomiques introduits.

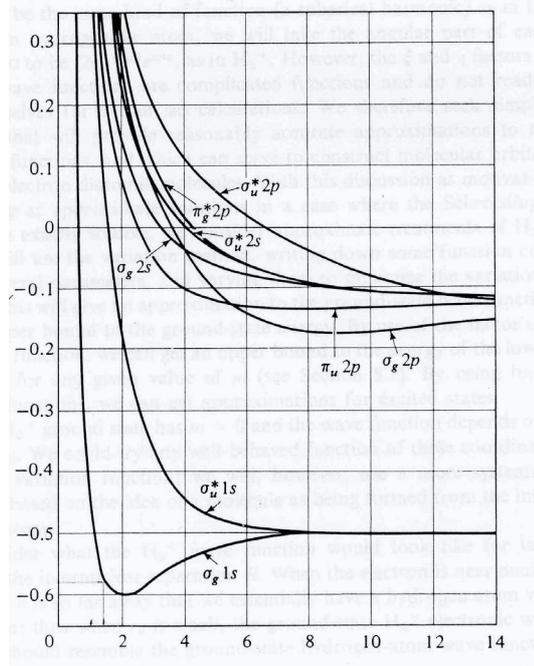


Figure 6.2: Variation de l'énergie totale $E_n(R)$ pour les premiers états de l'ion H_2^+ (figure tirée de [34]). $\frac{R}{a_0}$ est porté en abscisse, $\frac{E}{2E_1}$ est en ordonnée ($2E_1 \simeq 27$ eV est l'unité atomique d'énergie, aussi appelée Rydberg). Noter l'existence de croisements, pour lesquels il convient de réexaminer au cas par cas l'approximation adiabatique.

La divergence de $E_n(R)$ à l'origine provient exclusivement de $\frac{e'^2}{R}$. En ce qui concerne les $E_{n,\text{el}}(R)$, leurs limites sont faciles à trouver :

- $R \rightarrow 0$: on a un ion He^+ ($Z = 2$), dont les énergies sont $-\frac{2^2 m e'^4}{2n^2 \hbar^2} = 4E_{n,\text{hydrogène}} = -4\frac{E_1}{n^2}$ où :

$$E_1 = -\frac{\hbar^2}{2ma_0^2} = -\frac{e'^2}{2a_0} \simeq -13 \text{ eV} , \quad (6.12)$$

d'où :

$$\lim_{R \rightarrow 0} E_{n,\text{el}}(R) = -4\frac{E_1}{n^2} . \quad (6.13)$$

- $R \rightarrow \infty$: on se retrouve avec un proton nu et un atome d'hydrogène, système d'énergie $-\frac{E_1}{n^2}$.

Pour l'énergie totale (6.4) (les noyaux étant toujours fixes), on a donc :

$$E_n(R) \simeq -4\frac{E_1}{n^2} + \frac{e'^2}{R} \simeq \frac{e'^2}{R} \quad (R \rightarrow 0) , \quad E_n(R) \simeq -\frac{E_1}{n^2} + \frac{e'^2}{R} \simeq -\frac{E_1}{n^2} \quad (R \rightarrow +\infty) . \quad (6.14)$$

Pour l'état fondamental de H_2^+ , les différentes contributions à l'énergie se combinent pour donner effectivement un état stable, et comme $\frac{e'^2}{R}$ diverge en $R = 0$, l'existence d'un tel état liant est finalement conditionné

¹³mais pas trop petit afin que l'on puisse s'en tenir aux deux états 1s des atomes séparés. Aux très petites distances, c'est l'ensemble des états atomiques qu'il faut considérer.

par la décroissance assez rapide de $E_{n,\text{el}}(R)$ quand R décroît (voir plus loin la discussion physique). Quantitativement, le minimum est ici assez plat et l'énergie de liaison E_{diss} est corrélativement faible ; les valeurs numériques sont :

$$R_0 = 2a_0 \simeq 1.06 \text{ \AA} \quad E_{\text{diss}} \simeq 2.79 \text{ eV} \simeq 64 \text{ kcal/mole} \simeq 0.17 E_1 . \quad (6.15)$$

L'énergie de dissociation ne représente donc que 17% de l'énergie totale au minimum. Au total, la molécule est assez "molle" (fragile), ce qui se comprend bien puisqu'il n'y a qu'un seul électron pour assurer la liaison.

La notation traditionnelle des différents états fait référence à leurs propriétés de symétrie. En particulier, H_0 est invariant dans l'inversion d'espace Π , aussi appelée *parité* ; cette opération change ϕ en $\phi + \pi$ sans modifier ni r_1 , ni r_2 et n'affecte donc visiblement pas H_0 ($[\Pi, H_0] = 0$). En l'absence de dégénérescence, les états propres sont donc soit pairs (notés g, pour *gerade*), soit impairs (notés u, pour *ungerade*) dans l'inversion :

$$\Pi\Phi_{\text{g}} = \Phi_{\text{g}} , \quad \Pi\Phi_{\text{u}} = -\Phi_{\text{u}} . \quad (6.16)$$

Un autre élément de symétrie physiquement utile à préciser est le comportement des fonctions propres dans une réflexion par rapport au plan médiateur des deux protons P_1 et P_2 , puisqu'une fonction impaire dans cette opération est nulle partout dans ce plan (on s'attend donc à ce qu'elle soit antiliante). Cette propriété de symétrie n'est pas indépendante des précédentes puisqu'elle est le produit d'une rotation d'un angle π autour de l'axe internucléaire (qui donne $(-1)^m$ par le facteur $e^{im\phi}$) et d'une inversion d'espace. Le comportement des fonctions propres sous l'effet de la symétrie-miroir est donc :

- pas de changement de signe pour les états g ayant m pair et les états u avec m impair
- changement de signe pour les états g ayant m impair et les états u avec m pair

Malgré la redondance, l'usage est d'ajouter une astérisque¹⁴ pour les états impairs dans la symétrie-miroir ($1\sigma_{\text{u}}^*$, $2\pi_{\text{g}}^*$, ...) : l'astérisque permet ainsi d'identifier à l'œil les orbitales antiliantes.

6.1.2 Méthodologie pour une description approchée

La méthode exposée ci-dessus est d'une part déjà complexe – alors qu'il s'agit de la molécule la plus simple –, d'autre part n'est pas généralisable au cas d'édifices plus complexes¹⁵. Il faut donc trouver d'autres méthodes de description (approchée, cette fois), qui constituent un véritable arsenal, mais dont les idées fondamentales peuvent déjà s'énoncer à propos de H_2^+ .

L'une de ces méthodes (dite LCAO, pour *Linear Combination of Atomic Orbitals*) consiste à définir une base de fonctions d'origine atomique pour exprimer, sous la forme d'un développement linéaire, les orbitales moléculaires d'un électron¹⁶. L'idée physique sous-jacente est élémentaire. Chaque proton (ou noyau) offre à l'électron une infinité d'états liés et non-liés (états de l'atome d'hydrogène), mais il n'y a pas lieu de privilégier l'un de ces ensembles d'états – dont la réunion forme d'ailleurs un système surcomplet si l'on incorpore les états de diffusion. S'agissant de décrire le fondamental de l'ion et/ou ses premiers états excités, on choisit naturellement d'introduire les états atomiques de plus basse énergie. N'en retenant que deux – pour l'exemple, et pouvoir faire le calcul à la main –, on considère ainsi les deux états $1s$ normalisés, centrés sur les deux protons, soit la fonction $\psi_{nlm} = \psi_{100} \equiv \phi_1$ centrée sur P_1 , et sa jumelle centrée sur P_2 :

$$\phi_2(\vec{r}) = \psi_{100}(\vec{r} - \vec{R}) , \quad (6.17)$$

¹⁴Il n'existe pas d'états du type σ_{g}^* , par exemple.

¹⁵On ne voit pas bien ce que viendraient faire les coordonnées elliptiques dans une molécule à $M > 2$ noyaux...

¹⁶Pour une molécule à plusieurs électrons, on utilisera les orbitales moléculaires pour contruire un (des) déterminant(s) de Slater, tout comme on utilise des orbitales atomiques pour contruire les états antisymétriques d'un atome polyélectronique.

image de ϕ_1 par la translation $\mathcal{T}(\vec{R})$ (on prend l'origine du repère sur P_1). Ces deux fonctions (réelles) ne sont pas orthogonales ; soit S leur produit scalaire :

$$\int_{\mathbb{R}^3} d^3r \phi_1(\vec{r})\phi_2(\vec{r}) = S , \quad (6.18)$$

qui est bien sûr une fonction décroissante de R , en gros de façon exponentielle. D'ailleurs, en écrivant la positivité du produit scalaire $\langle \phi_1 + \lambda\phi_2 | \phi_1 + \lambda\phi_2 \rangle \forall \lambda$, on voit facilement que $S^2 \leq 1$. Ici, S est visiblement positif, de sorte que :

$$0 \leq S \leq 1 . \quad (6.19)$$

Le calcul le plus simple consiste donc à s'en tenir à ces deux fonctions, arguant du fait que les autres états sont nettement plus hauts en énergie ; la solution ainsi obtenue sera évidemment une solution approchée. On écrit ainsi, pour la fonction électronique (mais ici, c'est une simple orbitale moléculaire puisqu'il n'y a qu'un électron) :

$$\Phi = c_1\phi_1 + c_2\phi_2 , \quad (6.20)$$

et la question est maintenant de trouver les deux coefficients c_1 et c_2 . En reportant dans l'équation aux valeurs propres (6.1), en multipliant à gauche par $\langle \phi_1 |$ et par $\langle \phi_2 |$, on obtient le système¹⁷ :

$$\begin{cases} \langle \phi_1 | H_0 | \phi_1 \rangle c_1 + \langle \phi_1 | H_0 | \phi_2 \rangle c_2 = E(c_1 + S c_2) \\ \langle \phi_2 | H_0 | \phi_1 \rangle c_1 + \langle \phi_2 | H_0 | \phi_2 \rangle c_2 = E(S c_1 + c_2) \end{cases} . \quad (6.21)$$

En posant :

$$\langle \phi_1 | H_0 | \phi_1 \rangle = \langle \phi_2 | H_0 | \phi_2 \rangle = \varepsilon , \quad \langle \phi_1 | H_0 | \phi_2 \rangle = \langle \phi_2 | H_0 | \phi_1 \rangle = v , \quad (6.22)$$

on obtient :

$$\begin{cases} (\varepsilon - E) c_1 + (v - ES) c_2 = 0 \\ (v - ES) c_1 + (\varepsilon - E) c_2 = 0 \end{cases} . \quad (6.23)$$

Ce système n'a de solution non-nulle que si son déterminant est nul :

$$\begin{vmatrix} \varepsilon - E & v - ES \\ v - ES & \varepsilon - E \end{vmatrix} = 0 \iff \varepsilon - E = \pm(v - ES) \iff E = \frac{\varepsilon \pm v}{1 \pm S} \equiv E_{\mp} . \quad (6.24)$$

Pour E_+ , on a $\varepsilon - E_+ = (v - E_+S)$, donc $c_1 + c_2 = 0$, soit $c_1 = -c_2$, l'état propre est de la forme $\phi_1 - \phi_2$ (état antiliant, densité nulle dans le plan médiateur) ; au contraire, Pour E_- , on a $\varepsilon - E_- = -(v - E_-S)$, donc $c_1 - c_2 = 0$, soit $c_1 = c_2 = 0$: l'état propre est de la forme $\phi_1 + \phi_2$ (état liant, accumulation de densité entre les protons). Les deux fonctions propres (approchées) ainsi obtenues, une fois normalisées, sont donc :

$$\Phi_{E_{\pm}} = \frac{1}{\sqrt{2(1 \mp S)}} (\phi_1 \mp \phi_2) , \quad E_-(R) < E_+(R) . \quad (6.25)$$

Les deux énergies $E_{\pm}(R)$ varient comme indiqué sur la figure 6.3. Ce schéma très simple rend donc compte qualitativement de l'existence d'un fondamental stable. Bien sûr, on a l'inégalité :

$$E_-(R) > E_{n=1}(R) , \quad (6.26)$$

où $E_{n=1}(R)$ est l'énergie fondamentale exacte obtenue dans la section 6.1.1, conformément au Principe Variationnel.

La fig. 6.4 donne par comparaison les deux énergies électroniques correspondantes, dont les limites en $R = 0$ ne sont pas (et n'ont aucune raison d'être) les mêmes que pour la fonction exacte (comparer avec (6.13) – avec quel n d'ailleurs ?). Pour l'état liant, le minimum est en $R = 0$, pour l'état antiliant, le minimum (très

¹⁷Comme les deux orbitales atomiques ϕ_1 et ϕ_2 ne sont pas orthogonales, une quantité du type $\langle \phi_j | H_0 | \phi_{j'} \rangle$ n'est pas l'élément de matrice $(H_0)_{jj'}$ de H_0 entre les deux états $|\phi_j\rangle$ et $|\phi_{j'}\rangle$. En termes de ces quantités, l'équation donnant les valeurs propres E est $\text{Det}(\langle \phi_j | H_0 | \phi_{j'} \rangle - E \delta_{jj'}) = 0$, équivalente à $\text{Det}((H_0)_{jj'} - E \delta_{jj'}) = 0$.

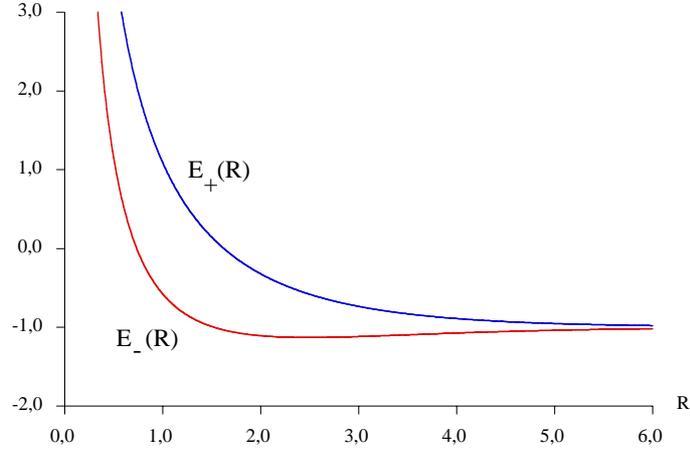


Figure 6.3: Variation de l'énergie totale pour les deux orbitales moléculaires liante et antiliante $E_{\mp}(R)$ (voir (6.24)). R est compté en unités a_0 et les énergies en unités E_I ; au minimum de $E_-(R)$, la différence d'énergie est voisine de 10 eV.

plat !) est pour un R'_0 fini. Dans les deux cas, l'interaction entre les deux protons médiée par l'électron est de fait attractive.

Cela vaut la peine de dire un mot sur les quantités S , ε et v (pour plus de détails, voir [8], complément G_{XI}). D'abord, on trouve sans difficulté¹⁸ :

$$S = e^{-\rho} \left(1 + \rho + \frac{\rho^2}{3}\right) \quad \left(\rho = \frac{R}{a_0}\right) . \quad (6.27)$$

Par ailleurs, en raison de la définition¹⁹ (6.22) :

$$\varepsilon = \langle \phi_1 | \frac{\vec{p}^2}{2m} - \frac{e'^2}{r_1} - \frac{e'^2}{r_2} | \phi_1 \rangle + \frac{e'^2}{R} = -E_I - \langle \phi_1 | \frac{e'^2}{r_2} | \phi_1 \rangle + \frac{e'^2}{R} ; \quad (6.28)$$

on trouve :

$$\langle \phi_1 | \frac{e'^2}{r_2} | \phi_1 \rangle = \frac{2}{\rho} [1 - e^{-2\rho} (1 + \rho)] E_I , \quad (6.29)$$

d'où :

$$\varepsilon = -E_I - \frac{2}{\rho} [1 - e^{-2\rho} (1 + \rho)] E_I + \frac{e'^2}{R} = -E_I + \frac{2}{\rho} e^{-2\rho} (1 + \rho) E_I . \quad (6.30)$$

Enfin, l'intégrale v est :

$$v = \langle \phi_1 | H_0 | \phi_2 \rangle = -S E_I - \langle \phi_1 | \frac{e'^2}{r_2} | \phi_2 \rangle + S \frac{e'^2}{R} ; \quad (6.31)$$

un calcul simple donne l'expression finale de v :

$$v = -S E_I - 2 e^{-\rho} (1 + \rho) E_I + S \frac{e'^2}{R} . \quad (6.32)$$

En définitive, c'est essentiellement²⁰ l'intégrale v (parfois appelée intégrale de résonance) qui offre la possibilité à l'électron de sauter par effet-tunnel d'un proton à l'autre, à une fréquence²¹ $\nu = h^{-1}(E_+ - E_-)$:

¹⁸Le calcul se fait facilement en coordonnées elliptiques ([8], complément G_{XI}).

¹⁹Se souvenir que $\phi_1 \equiv \psi_{100}$ est propre de la somme des deux premiers termes apparaissant dans (6.2).

²⁰Si v était mis à zéro, on aurait encore $E_- < E_+$ à cause du dénominateur $1 \pm S$, mais la stabilisation serait bien moindre, et ne résulterait que du recouvrement des orbitales atomiques.

²¹Avec $E_+ - E_- \sim 10$ eV, $\nu \sim 10^{15}$ Hz.

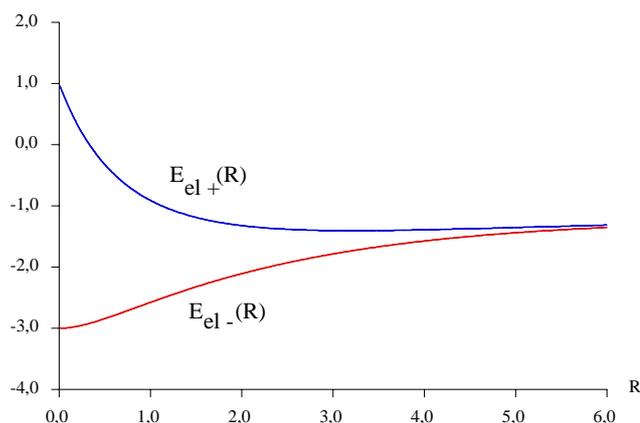


Figure 6.4: Variation de l'énergie électronique $E_{\text{el}\pm} = E_{\pm} - \frac{e^2}{R}$ pour les deux orbitales moléculaires liante et antiliante $E_{\mp}(R)$ (voir (6.24)). R est compté en unités a_0 et les énergies en unités E_I .

si on imagine l'électron "préparé" dans l'état purement atomique $\Psi(t=0) = \phi_1$ centré sur P_1 , l'état à l'instant t est²² $\Psi(t) = \cos 2\pi\nu t \phi_1 + i \sin 2\pi\nu t \phi_2$. Les états ϕ_i sont localisés autour de chaque proton ; au contraire, les deux orbitales moléculaires sont délocalisées au sens où elles donnent une densité notable sur tout l'espace moléculaire. Cette délocalisation induite par la proximité de ces derniers produit un état d'énergie plus basse que $-E_I$; on parle alors de stabilisation par résonance²³. On retrouvera ci-dessous le fait que c'est bien le mouvement de l'électron sur tout l'espace offert qui est à l'origine d'une liaison stable (l'énergie cinétique de celui-ci *augmente* quand la liaison se forme, voir section 6.2).

Quantitativement, les résultats ne sont pas très fameux, mais la méthode peut, cette fois, être améliorée de façon évidente :

- en incluant dans les orbitales atomiques des paramètres variationnels (par exemple, une charge nucléaire ajustable) et en minimisant l'énergie liante, conformément au Principe variationnel. La charge effective obtenue n'est pas interprétable, au contraire d'un système à plusieurs électrons, par un effet d'écran ; c'est en fait un simple facteur d'échelle, comme on l'a vu à propos de l'atome d'Hélium
- en élargissant la base par l'introduction d'autres orbitales atomiques, ce qui revient à travailler dans un espace vectoriel plus vaste. D'une façon générale, si on introduit N orbitales atomiques, on obtiendra autant d'orbitales moléculaires qui seront des approximations pour les fonctions d'onde des N premiers états (y compris le fondamental) électroniques Φ_n . Les énergies correspondantes $E_n^{(N)}(R)$ ($1 \leq n \leq N$) sont les énergies approchées, dont on espère qu'elles convergent vite vers leurs limites exactes quand N devient assez grand. Elles obéissent à certaines relations d'ordre quand on compare les résultats d'un premier calcul avec N orbitales atomiques, et ceux d'un second effectué en ajoutant deux orbitales atomiques à la base introduite dans le premier calcul. En ordonnant suivant $E_n^{(N)} < E_{n+1}^{(N)}$ les énergies obtenues avec N orbitales atomiques, un théorème dû à Fröbenius²⁴ permet d'affirmer que $E_n^{(N+1)} < E_n^{(N)}$; en particulier, les énergies $E_{n=1}^{(N)}$ de l'état fondamental forment une suite décroissante, toujours en accord avec le Principe variationnel.

On conçoit aisément que ces généralisations, qui améliorent sensiblement la description pour l'ion moléculaire considéré, s'étendent sans peine à des molécules plus complexes – sans préjuger par ailleurs du mode de traitement de l'interaction entre les électrons au sein de la molécule²⁵.

²²Le i devant $\sin 2\pi\nu t$ est essentiel pour assurer la conservation de la norme de $\Psi(t)$ alors que les ϕ_i ne sont pas orthogonales.

²³Un autre cas de stabilisation par résonance est celui de la molécule de benzène.

²⁴aussi connu sous le nom de théorème de Rayleigh - Ritz.

²⁵Pour une molécule contenant plusieurs électrons, on peut encore former des orbitales moléculaires comme ci-dessus, que l'on remplit les unes après les autres en épuisant le *stock* d'électrons et avec lesquelles on forme un déterminant de Slater. Il est clair que la fonction d'onde ainsi construite est caractéristique un modèle à électrons indépendants. On sait faire beaucoup mieux : de

6.2 Nature physique de la liaison chimique

L'origine de la stabilité de la liaison résulte clairement d'une compétition entre deux interactions antagonistes : les deux protons se repoussent, mais chacun d'entre eux attire le même électron ; celui-ci joue le rôle d'une "colle" entre deux particules qui, nues, ont tendance à se repousser fortement. Cette image est universelle en Physique : toute interaction entre deux systèmes, fondamentale ou comme ici effective, résulte de l'échange de particules entre ces deux systèmes. Les deux protons coopèrent en attirant chacun l'électron et ce sont ces efforts conjugués qui assurent la stabilité de l'édifice.

En réalité, il s'agit d'ingrédients *nécessaires* ; l'existence ou non d'une liaison stable résulte évidemment du fait que l'interaction effective induite par les allers-retours de l'électron est assez forte pour donner à $E(R)$ un minimum ; plus précisément, c'est la nature même de la distribution électronique liée à une orbitale moléculaire ou une autre qui conditionne l'existence d'une liaison, laquelle résulte de la compétition subtile entre les différentes contributions à l'énergie.

Le Théorème du Viriel permet d'analyser finement cette compétition, en fournissant des relations simples entre les valeurs moyennes des énergies cinétique et potentielle. La démonstration de ce théorème repose sur la considération de la quantité $\vec{r} \cdot \vec{p}$ (pour un seul électron), tout comme en Mécanique Classique ; le commutateur de cette quantité avec $H_0 = T + V$ est :

$$[H_0, \vec{r} \cdot \vec{p}] = \sum_{u=x, y, z} [H_0, up_u] = \sum_{u=x, y, z} ([H_0, u] p_u + u[H_0, p_u]) \quad . \quad (6.33)$$

Compte tenu de :

$$[H_0, u] = \frac{1}{2m} [\vec{p}^2, u] = \frac{1}{2m} [p_u^2, u] = -i\frac{\hbar}{m} p_u \quad , \quad [H_0, p_u] = [V, p_u] = -i\hbar [V, \frac{\partial}{\partial u}] = +i\hbar \frac{\partial V}{\partial u} \quad , \quad (6.34)$$

il vient :

$$[H_0, \vec{r} \cdot \vec{p}] = \sum_{u=x, y, z} \left(-i\frac{\hbar}{m} p_u^2 + i\hbar u \frac{\partial V}{\partial u} \right) = i\hbar (-2T + \vec{r} \cdot \vec{\nabla} V) \quad . \quad (6.35)$$

Par ailleurs, V est une fonction homogène de degré -1 :

$$V(\vec{r}, \vec{R}) = f(x, y, z, R) \quad \text{avec : } f(\lambda x, \lambda y, \lambda z, \lambda R) = \frac{1}{\lambda} f(x, y, z, R) \quad ; \quad (6.36)$$

la dérivation en λ de la dernière égalité donne :

$$x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} + z \frac{\partial f}{\partial z} + R \frac{\partial f}{\partial R} = -\frac{1}{\lambda^2} f(x, y, z, R) \quad (6.37)$$

et si on fait $\lambda = 1$ partout, il vient :

$$\vec{r} \cdot \vec{\nabla} f + R \frac{\partial f}{\partial R} = -f(x, y, z, R) \quad \iff \quad \vec{r} \cdot \vec{\nabla} V + R \frac{\partial V}{\partial R} = -V \quad . \quad (6.38)$$

Cette relation permet d'écrire le commutateur (6.35) sous la forme :

$$[H_0, \vec{r} \cdot \vec{p}] = i\hbar \left(-2T - V - R \frac{\partial V}{\partial R} \right) \quad . \quad (6.39)$$

Maintenant :

1. la valeur moyenne d'un commutateur $[H_0, \Omega]$ est nulle dans tout état propre normalisable $|E\rangle$ de H_0 :

$$\langle E|[H_0, \Omega]|E\rangle = (E - E) \langle E|\Omega|E\rangle = 0 \quad , \quad (6.40)$$

toute évidence, la base d'orbitales atomiques choisie judicieusement peut aussi être utilisée pour *représenter* (approximativement) les orbitales moléculaires de Hartree - Fock, afin de prendre (partiellement) en compte la répulsion entre les électrons (le schéma de Hartree-Fock n'étant qu'une approximation parmi d'autres).

la dernière égalité supposant que $\langle E|\Omega|E\rangle$ est une quantité finie, ce qui est vrai pour tout état lié²⁶. Il résulte alors de (6.39) que la moyenne²⁷ prise sur tout état électronique normalisable satisfait :

$$\langle 2T + V + R \frac{\partial V}{\partial R} \rangle = 0 \quad (6.41)$$

2. par ailleurs, soit un Hamiltonien dépendant d'un paramètre ζ , $H(\zeta)$; ses valeurs propres E et ses fonctions propres ψ_E , dépendent aussi de ce paramètre, soit $E(\zeta)$ et $\psi_E(\zeta)$. Comme, pour une fonction normalisable, $\frac{d}{d\zeta} \langle \psi_E(\zeta) | \psi_E(\zeta) \rangle = 0$, il vient :

$$\frac{dE}{d\zeta} \equiv \frac{d}{d\zeta} \langle \psi_E(\zeta) | H(\zeta) | \psi_E(\zeta) \rangle = \langle \psi_E(\zeta) | \frac{\partial H}{\partial \zeta} | \psi_E(\zeta) \rangle \equiv \langle \frac{\partial H}{\partial \zeta} \rangle \iff \frac{dE}{d\zeta} = \langle \frac{\partial H}{\partial \zeta} \rangle, \quad (6.42)$$

la dernière égalité est l'expression du théorème de Hellmann - Feynman.

avec $H \equiv H_0$ et $\lambda \equiv R$, (6.42) donne :

$$\frac{dE}{dR} = \langle \frac{\partial H_0}{\partial R} \rangle = \langle \frac{\partial V}{\partial R} \rangle, \quad (6.43)$$

ce qui permet de récrire (6.41) sous la forme :

$$\langle 2T \rangle + \langle V \rangle + R \frac{dE}{dR} = 0, \quad (6.44)$$

où V désigne toujours l'énergie potentielle totale. Par ailleurs, on a évidemment :

$$E = \langle T \rangle + \langle V \rangle, \quad (6.45)$$

d'où, par élimination soit de $\langle V \rangle$ soit de $\langle T \rangle$ entre (6.44) et (6.45), les deux relations :

$$\langle T \rangle = -E - R \frac{dE}{dR}, \quad \langle V \rangle = 2E + R \frac{dE}{dR}. \quad (6.46)$$

Ces dernières expressions permettent de montrer que l'existence d'une liaison chimique stable résulte de la *diminution* de l'énergie potentielle totale et d'une *augmentation* de l'énergie cinétique relativement à leurs valeurs pour la molécule dissociée ($R = \infty$). En effet, s'il existe une position d'équilibre R_0 , alors, en ce point (où la dérivée s'annule) :

$$\langle T \rangle_{R_0} = -E(R_0) \quad \langle V \rangle_{R_0} = 2E(R_0). \quad (6.47)$$

Par ailleurs, pour $R = +\infty$, on a aussi²⁸ :

$$\langle T \rangle_{\infty} = -E(\infty) \quad \langle V \rangle_{\infty} = 2E(\infty). \quad (6.48)$$

D'où :

$$\langle T \rangle_{R_0} - \langle T \rangle_{\infty} = E(\infty) - E(R_0) > 0, \quad \langle V \rangle_{R_0} - \langle V \rangle_{\infty} = 2[E(R_0) - E(\infty)] < 0. \quad (6.49)$$

Physiquement, l'augmentation d'énergie cinétique lors de la formation de la liaison se comprend bien : initialement localisé près d'un proton pour les très grands R , l'électron peut en fait aller sur l'autre quand R devient assez petit. Ces excursions se traduisent naturellement par une augmentation de T . La liaison est donc d'autant plus forte que l'électron "va vite" d'un proton à l'autre.

Enfin, en terme de densité électronique, on note que tout naturellement la liaison stable (état liant) contient une accumulation d'électron entre les protons, alors qu'une liaison instable présente un déficit d'électrons entre ceux-ci.

²⁶et pour toute observable suffisamment régulière. On peut aussi dire : quand H_0 et Ω sont hermitiques, leur commutateur est anti-hermitique ; sa valeur moyenne est donc de la forme $i \times$ nombre réel, un nombre qui ne peut être que zéro.

²⁷Toutes ces moyennes résultent d'une intégration exclusive sur \vec{r} et restent donc des fonctions de R .

²⁸On suppose que $\lim_{R \rightarrow +\infty} R \frac{dE}{dR} = 0$, ce qui suppose $\alpha > 0$ si $E(R) \sim E_{\infty} + CR^{-\alpha}$ (on exclut par exemple une dépendance logarithmique $\sim [\ln \frac{R}{\sigma_0}]^{-1}$).

Chapitre 7

Mouvement des noyaux. Spectres de vibration et de rotation

L'analyse de la structure électronique, présentée sommairement au chapitre précédent, permet de comprendre l'origine de la liaison chimique et, plus généralement, fournit un schéma explicatif de la stabilité des molécules. De surcroît, la mise en évidence des états électroniques – où les noyaux sont fixes par définition – permet de comprendre l'existence des spectres UV et visible des molécules – à condition toutefois de ne pas rentrer trop dans le détail¹.

La description précise du mouvement des noyaux (une fois moyenné celui des électrons) est le fondement théorique d'une autre spectroscopie (IR et micro-ondes²) apportant de précieux renseignements sur la conformation moléculaire et les champs de forces effectifs entre noyaux. Principalement, la spectroscopie IR (spectres de vibration) donne des renseignements sur les constantes de force d'équilibre et donc sur la "force" des liaisons chimiques, tandis que la spectroscopie micro-ondes permet de "voir" les états rotationnels, la molécule étant considérée (en première approximation) comme un rotateur rigide, défini par des axes principaux et des moments d'inertie ; la détermination de ces derniers fournit notamment les distances d'équilibre des liaisons. Au total, les mesures spectroscopiques mettant en jeu les degrés de libertés vibrationnels et rotationnels permettent de remonter à la configuration géométrique de la molécule et aux forces qui la régissent.

Comme on l'a vu au chapitre 5, la fonction d'onde adiabatique d'une molécule est de la forme :

$$\Psi(q_i; Q_j) = \chi_m(Q_j) \Phi_m(q_i; Q_j) , \quad (7.1)$$

où la fonction d'onde électronique Φ_m satisfait :

$$H_0 \Phi_m \equiv (T_e + V) \Phi_m = E_m(Q_j) \Phi_m , \quad (7.2)$$

T_e étant l'opérateur cinétique des électrons, cependant que la fonction d'onde des noyaux $\chi_m(Q_j)$ obéit à :

$$[T_N + E_m(Q_j)] \chi_m(Q_j) = E \chi_m(Q_j) . \quad (7.3)$$

Au chapitre précédent, on a montré sommairement que l'existence d'états propres électroniques liés – et stables vis-à-vis de (petites) déformations de la configuration nucléaire – est due à de subtiles compétitions entre l'énergie cinétique des électrons, la répulsion entre noyaux et entre électrons, et l'interaction attractive entre les noyaux et les électrons.

¹Par exemple, toute raie d'émission s'accompagne de satellites provenant de transitions électroniques avec modification de l'état de vibration intramoléculaire.

²La spectroscopie Raman, très utile pour les molécules dénuées de moment dipolaire électrique – voir ci-dessous – met en jeu des phénomènes généralisant la diffusion Rayleigh classique, et utilise donc une source d'excitation de fréquence optique, typiquement.

Pour une molécule à N noyaux, il y a $3N$ degrés de liberté nucléaires. Comme on l'a vu, seuls $3N - 6$ d'entre eux – ou $3N - 5$ pour une molécule linéaire³ – correspondent à des mouvements internes (c'est-à-dire produisant des déformations du squelette moléculaire). Parmi les autres degrés, ceux de translation ne donnent jamais de quantification (translation en bloc dans \mathbb{R}^3 sans conditions aux limites) ; en revanche, la rotation en bloc est quantifiée, en conséquence du fait que les variables dynamiques sont des *angles*, définis sur un intervalle *fini*, $[0, 2\pi]$ ou $[0, \pi]$ suivant les cas. En Mécanique Quantique, toute variable confinée produit la quantification *via* le jeu des conditions aux limites imposées par le sens physique accordé aux fonctions d'onde. En conséquence de la théorie du moment cinétique, la fonction est inchangée par une rotation de 2π des variables *spatiales* puisqu'il n'est ici question que d'un moment cinétique *orbital*⁴.

L'étude détaillée des mouvements nucléaires est assez laborieuse pour une molécule un tant soi peu complexe⁵ ; on se bornera principalement ici à exposer les idées essentielles, données ci-dessous dans le cas le plus simple, celui des molécules diatomiques.

7.1 Molécules diatomiques

Dans cette section, on considère exclusivement le cas d'une molécule AB, hétéronucléaire ($A \neq B$) ou homonucléaire ($A = B$). Dans cette seconde étape de l'approximation adiabatique, on dispose d'une énergie électronique E_m , jouant le rôle d'une énergie potentielle dans une équation aux valeurs propres pour (ici) deux noyaux ponctuels de masses M_A et M_B . Évidemment, E_m ne dépend ici que d'une seule coordonnée, la distance R entre les deux noyaux (homogénéité et isotropie galiléennes de l'espace). S'agissant d'étudier les mouvements de la molécule constituée, la fonction $E_m(R)$ est supposée posséder un minimum (unique) en $R = R_0$.

7.1.1 Fonctions propres de vibration - rotation

L'équation nucléaire a la forme explicite :

$$\left[-\frac{\hbar^2}{2M_A} \Delta_A - \frac{\hbar^2}{2M_B} \Delta_B + E_m(R) \right] \chi_m(\vec{R}_A, \vec{R}_B) = E \chi_m(\vec{R}_A, \vec{R}_B) . \quad (7.4)$$

Ceci représente l'équation aux valeurs propres standard pour deux particules en interaction centrale par $E_m(R)$. La séparation du centre de masse fait apparaître χ_m sous la forme :

$$\chi_m(\vec{R}_A, \vec{R}_B) = e^{i\vec{K} \cdot \vec{X}} F_m(\vec{R}) \quad (\vec{R} = \vec{R}_B - \vec{R}_A) , \quad (7.5)$$

où \vec{X} donne le centre de masse :

$$\vec{X} = \frac{M_A \vec{R}_A + M_B \vec{R}_B}{M_A + M_B} , \quad (7.6)$$

et le vecteur d'onde \vec{K} est relié à l'énergie de translation en bloc par :

$$E_{\text{translation}} = \frac{\hbar^2 \vec{K}^2}{2(M_A + M_B)} . \quad (7.7)$$

Par ailleurs, la fonction F_m satisfait ($\mu^{-1} = M_A^{-1} + M_B^{-1}$) :

$$\left[-\frac{\hbar^2}{2\mu} \Delta_{\vec{R}} + E_m(R) \right] F_m(\vec{R}) = \left(E - \frac{\hbar^2 \vec{K}^2}{2M} \right) F_m(\vec{R}) . \quad (7.8)$$

³On entend par là une molécule dont la configuration d'équilibre est linéaire et qui ne peut se déformer "transversalement", cas-limite où les constantes de torsion transverses sont très grandes.

⁴Apparition de nombres exclusivement entiers (et non entiers ou demi-entiers comme pour un moment de spin).

⁵Quelques indications sont données dans la section 7.3.

μ étant une masse (réduite) *nucléaire*, le terme cinétique est petit par rapport à la profondeur du puits de potentiel que constitue $E_m(R)$. Dans le repère du centre de masse – le seul pertinent ici –, on a :

$$\left[-\frac{\hbar^2}{2\mu} \Delta_{\vec{R}} + E_m(R) \right] F_m(\vec{R}) = E F_m(\vec{R}) . \quad (7.9)$$

Ceci fait renouer avec un cadre familier : (7.9) est une équation aux valeurs propres pour une particule de masse μ dans le potentiel radial $E_m(R)$. La séparation des variables angulaires procède donc exactement comme pour l'atome d'hydrogène et fait apparaître les harmoniques sphériques. En désignant par (R, θ, ϕ) les variables sphériques conventionnelles (voir fig. 7.1), on pose⁶ :

$$F_m(\vec{R}) = \frac{1}{R} f(R) Y_{LM}(\theta, \phi) \quad (L = 0, 1, 2, \dots, \quad M = -L, -L+1, -L+2, \dots, L-1, L) , \quad (7.10)$$

et la fonction radiale f satisfait :

$$-\frac{\hbar^2}{2\mu} \frac{d^2 f}{dR^2} + V_{\text{eff}}(R) f(R) = E f(R) , \quad (7.11)$$

avec :

$$V_{\text{eff}}(R) = E_m(R) + \frac{\hbar^2 L(L+1)}{2\mu R^2} . \quad (7.12)$$

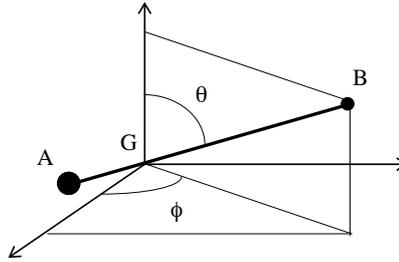


Figure 7.1: Repère utilisé pour fixer la position d'une molécule diatomique AB.

Comme d'habitude, le potentiel effectif est la somme de l'énergie potentielle purement radiale et du terme centrifuge $\frac{\hbar^2 L^2}{2\mu R^2}$, provenant de l'opérateur cinétique T_N . De toute évidence, la variable R décrit la vibration de la molécule, alors que les variables angulaires sont associées à sa rotation. Il est utile à ce stade de mettre des nombres dans (7.12), ce qui permet de réaliser que les deux contributions à V_{eff} ne sont pas du même ordre de grandeur. En effet (m est ici la masse de l'électron) :

$$\frac{\hbar^2 L(L+1)}{2\mu R^2} \sim \frac{\hbar^2}{2\mu(2a_0)^2} = \frac{\hbar^2}{8\mu(\frac{\hbar^2}{me'^2})^2} = \frac{m}{\mu} \frac{me'^4}{8\hbar^2} = \frac{m}{4\mu} E_I , \quad (7.13)$$

d'une part. D'autre part, $E_m \sim E_I$, de sorte que :

$$\frac{\hbar^2 L(L+1)}{2\mu R^2} \sim \frac{m}{4\mu} E_m \ll E_m ; \quad (7.14)$$

sans surprise, l'énergie de rotation est bien une petite correction devant l'énergie E_m ; en première approximation :

$$V_{\text{eff}} \simeq E_m . \quad (7.15)$$

Par ailleurs, ce sont les vibrations de la molécule constituée qui présentent un intérêt, en tant que source d'information sur l'édifice moléculaire tel qu'il existe. Implicitement, ceci suppose que les vibrations sont de

⁶L'indice m de F_m continue à désigner un ensemble de nombres quantiques. Par ailleurs, la fonction radiale f dépend explicitement de L , mais ceci est omis par simplicité.

petite amplitude (autrement, la liaison chimique se casse). Il est donc licite, pour la question posée, de développer V_{eff} près de son minimum et d'écrire⁷ :

$$V_{\text{eff}}(R) \simeq E_m(R_0) + \frac{1}{2} E_m''(R_0) (R - R_0)^2 + \frac{\hbar^2 L(L+1)}{2\mu R_0^2} - \frac{\hbar^2 L(L+1)}{\mu R_0^3} (R - R_0) , \quad (7.16)$$

Le troisième terme dans (7.16) est l'énergie du rotateur rigide, la distance R étant fixée à sa valeur d'équilibre R_0 . Le dernier terme est visiblement un couplage entre vibration et rotation : sous l'effet de celle-ci, la liaison s'étire un peu et le point d'équilibre change (un peu), sans pour autant changer la fréquence⁸ de fond de puits (qui ne dépend que de la dérivée seconde). Dans la suite, en oubliant ce type de correction et se bornant au terme harmonique venant de E_m , l'équation à résoudre, (7.11), prend la forme simplifiée :

$$-\frac{\hbar^2}{2\mu} \frac{d^2 f}{dR^2} + \left[E_m(R_0) + \frac{1}{2} \mu \omega_m^2 (R - R_0)^2 + B_m \hbar L(L+1) \right] f(R) = E f(R) , \quad (7.17)$$

où ont été introduites la pulsation de vibration ω_m et la notation traditionnelle pour le terme centrifuge⁹ :

$$\mu \omega_m^2 = E_m''(R_0) , \quad B_m = \frac{\hbar}{4\pi \mu R_0^2} \equiv \frac{\hbar}{4\pi I_0} , \quad (7.18)$$

I_0 désignant le moment d'inertie à la distance d'équilibre R_0 .

Clairement, on peut attendre des effets isotopiques importants, exclusivement dus au terme cinétique T_N : par exemple, entre l'hydrogène H_2 et son isotope D_2 , on gagne un facteur 2 pour la masse réduite μ . D'un autre côté, la dérivée seconde $E_m''(R_0)$ ne dépend pas des masses nucléaires¹⁰ : la variation de masse réduite se transporte donc fidèlement sur les pulsations. Compte tenu de la première relation (7.18), la pulsation propre doit donc être divisée par $\sqrt{2}$ et c'est bien ce que l'on observe : la vibration strictement harmonique se produit à 4401 cm^{-1} pour H_2 , et à 3112 cm^{-1} pour D_2 . D'une façon générale, la pulsation ω_m est d'autant plus grande que la masse réduite est faible et la liaison stable (le minimum de $E_m(R)$ est aigu). Pour une molécule de gaz rare (en interaction par van der Waals), ω_m est beaucoup plus petit (quelques dizaines de cm^{-1} au plus).

L'équation (7.17) est une approximation de (7.11), et représente un oscillateur harmonique linéaire¹¹ pour l'écart $R - R_0$, dont le spectre est décalé de la constante $B_m \hbar L(L+1)$. D'après ce que l'on sait de l'oscillateur harmonique linéaire, on peut dès lors affirmer que la valeur propre E de (7.17) est de la forme :

$$E = E_m(R_0) + \hbar \omega_m \left(v + \frac{1}{2} \right) + B_m \hbar L(L+1) \quad (v \in \mathbb{N}, L \in \mathbb{N}) . \quad (7.19)$$

Cette expression met bien en évidence le découplage des différents degrés de liberté, l'énergie apparaissant sous la forme d'une *somme* de trois termes, écrits dans l'ordre des énergies décroissantes :

$$E = E_{\text{el}} + E_{\text{vibr}} + E_{\text{rot}} . \quad (7.20)$$

Compte tenu des estimations des énergies de vibration et de rotation, on a $B_m \ll \frac{\omega_m}{2\pi}$, de sorte que les niveaux de vibration et de rotation sont disposés comme illustré sur la figure 7.2.

En ce qui concerne la fonction propre f , c'est – formellement – une fonction propre d'oscillateur harmonique centrée en R_0 , à ceci près que la variable R est ici cantonnée aux valeurs positives – donc la variable $R - R_0$ donnant l'écart à la position d'équilibre varie seulement de $-R_0$ à $+\infty$, et non entre $\pm\infty$. Ce *distinguo* n'est pas gênant en pratique, puisque toutes les fonctions d'onde décroissent très vite de part et d'autre de R_0 (l'enveloppe est gaussienne). Les différences deviennent sensibles seulement pour les grandes valeurs du nombre quantique de vibration v , mais, de toute façon, l'hypothèse des petites oscillations exclut la considération d'états de vibration très excités. Plus précisément, les fonctions propres harmoniques sont localisées autour de

⁷ R_0 dépend de l'état électronique considéré ; on devrait donc plutôt noter R_{m0} la valeur d'équilibre, ce que l'on ne fait pas pour alléger l'écriture.

⁸ Le terme suivant du développement de l'énergie centrifuge est en $(R - R_0)^2$ et "renormalise" donc la fréquence de vibration.

⁹ La constante B_m dépend de l'état électronique considéré *via* la distance d'équilibre R_0 de la liaison ; B_m est homogène à une fréquence.

¹⁰ L'énergie $E_m(R)$ se trouve dans l'étape où les noyaux sont fixes, et est donc forcément la même pour tous les isotopes.

¹¹ À une petite réserve près liée au fait qu'ici R est une variable positive – voir plus loin.

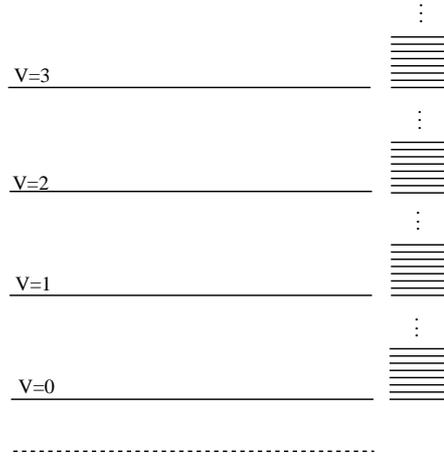


Figure 7.2: Représentation schématique des niveaux de vibration et de rotation d'une molécule diatomique. Le trait en pointillés schématise le zéro d'énergie. Les niveaux de vibration ont l'énergie $\hbar\omega_m(v + \frac{1}{2})$, $v \in \mathbb{N}$; les peignes très fins schématisent les niveaux de rotation, décalés de $\hbar B_m L(L + 1)$ par rapport aux niveaux de vibration pure.

R_0 sur un intervalle ΔR ; pour que la queue gaussienne du côté $R < 0$ soit invisible et donc inoffensive, il faut $\Delta R \ll R_0$. Comme $\Delta R \sim (\frac{\hbar}{\mu\omega_m})^{1/2}$, on doit donc avoir :

$$\sqrt{\frac{\hbar}{\mu\omega_m}} \ll R_0 \iff \sqrt{\frac{\hbar}{\mu\sqrt{E''_m(R_0)}/\mu}} \ll R_0 \iff \frac{\hbar^2}{\mu E''_m(R_0)R_0^2} \ll R_0^2 . \quad (7.21)$$

Par ailleurs, $E''_m(R_0)R_0^2 \sim \frac{\hbar^2}{ma_0^2}$; en reportant cette estimation dans (7.21), on trouve que cette inégalité est équivalente à :

$$\frac{m}{\mu} \ll 1 , \quad (7.22)$$

inégalité qui est clairement bien vérifiée.

Au total, les fonctions propres de vibration - rotation sont les fonctions :

$$F_{vLM} = \frac{1}{R} f_v(R) Y_{LM}(\theta, \phi) \equiv u_v(R) Y_{LM}(\theta, \phi) . \quad (7.23)$$

$f_v(R)$ est une gaussienne multipliée par un polynôme de Hermite de degré v , H_v :

$$f_v(R) = C_v e^{-\frac{\mu\omega_m}{2\hbar}(R-R_0)^2} H_v \left(\sqrt{\frac{\mu\omega_m}{2\hbar}} (R - R_0) \right) , \quad (7.24)$$

où C_v est la constante de normalisation. Y_{LM} est une harmonique sphérique qui décrit la rotation du "bâton" rigide¹² que constitue la molécule diatomique à noyaux fixés à la distance R_0 , et dont l'orientation est définie par les deux angles θ et ϕ .

¹²ce que l'on appelle un rotateur rigide.

7.1.2 Spectres de vibration et de rotation

À l'ordre le plus bas, l'interaction entre un système microscopique et le champ électromagnétique classique est du genre $-\vec{d}\cdot\vec{\mathcal{E}}$, où $\vec{\mathcal{E}}$ est le champ électrique de l'onde¹³ et où \vec{d} est le moment dipolaire électrique total¹⁴ de la molécule (neutre) moyenné sur le mouvement électronique associé à l'état électronique considéré : s'agissant ici d'examiner les transitions entre états dynamiques des noyaux, on raisonne à état électronique donné, en ne considérant que les degrés de liberté nucléaires. \vec{d} est donc :

$$\vec{d} = \langle \Phi_m | \sum_i e\vec{r}_i + \sum_j Z_j |e| \vec{R}_j | \Phi_m \rangle = \langle \Phi_m | \sum_i e\vec{r}_i | \Phi_m \rangle + \sum_j Z_j |e| \vec{R}_j ; \quad (7.26)$$

ici, les $\langle \dots | \dots | \dots \rangle$ représentent des intégrations sur les seules coordonnées électroniques \vec{r}_i .

Pour une molécule diatomique, le moment \vec{d} est forcément porté par l'axe internucléaire \overrightarrow{AB} puisque, en tant que propriété de la molécule, il est invariant dans toute opération de symétrie :

$$\vec{d} = d(R) \frac{\vec{R}}{R} . \quad (7.27)$$

Par ailleurs, si la molécule est homonucléaire ($A = B$, A_2 en notation chimique), le système est symétrique dans l'inversion par rapport au milieu de la liaison. Le seul vecteur invariant dans cette opération est le vecteur nul : dépourvue de moment dipolaire, une molécule de ce type est donc inactive dans l'infra-rouge et les micro-ondes¹⁵. Seules les molécules hétéronucléaires sont actives en IR et micro-ondes ; pour une telle molécule, l'existence d'un moment dipolaire permanent \vec{d} résulte physiquement du fait que, compte tenu de la dissymétrie des charges nucléaires, les électrons sont plus attirés par un noyau que par l'autre (celui qui est le plus électronégatif).

Comme ce sont la vibration et la rotation de la molécule telle qu'elle existe qui sont pertinentes, la forme utile du moment dipolaire \vec{d} résulte d'un développement de Taylor autour de R_0 , étant entendu que \vec{d} est le long de $\vec{R} \equiv \overrightarrow{AB}$:

$$d(R) = d(R_0) + q_{\text{eff}} \cdot (R - R_0) + \dots , \quad (7.28)$$

où la constante q_{eff} est une charge électrique effective. Si le champ est polarisé suivant Oz , le produit scalaire $-\vec{d}\cdot\vec{\mathcal{E}}$ donne $d(R) \cos\theta$. Les intensités des raies de vibration - rotation sont donc gouvernées¹⁶ par le module carré de l'élément de matrice :

$$\langle F_{vLM} | d(R) \cos\theta | F_{v'L'M'} \rangle . \quad (7.29)$$

Les variables sont séparées, de sorte que (7.29) est le produit de trois intégrales. L'une, de 0 à 2π , porte sur l'angle azimutal ϕ , son intégrand est $e^{i(M'-M)\phi}$ et n'est non-nulle que si $M' = M$, d'où la règle de sélection¹⁷ $\Delta M = 0$. Ceci étant fait, il reste le produit des deux intégrales :

$$\langle u_v | d(R) | u_{v'} \rangle \langle Y_{LM} | \cos\theta | Y_{L'M'} \rangle . \quad (7.30)$$

Comme on l'a déjà vu à propos des transitions dipolaires électriques pour un atome (ch. 1), le facteur angulaire n'est différent de zéro que si $|L - L'| = 1$, d'où une deuxième règle de sélection de type E1 :

$$L' = L \pm 1 . \quad (7.31)$$

¹³Il s'agit ici de la diffusion élastique résonnante de photons : tout comme pour un atome à deux niveaux, la susceptibilité est exaltée lorsque la fréquence du champ coïncide avec l'une des fréquences de Bohr du système. L'image classique est une vibration forcée du dipôle microscopique qui, en tant que telle, s'effectue à la fréquence de l'onde excitatrice.

Schématiquement, il s'agit de processus élémentaires du premier ordre vis-à-vis de la théorie des perturbations dépendant du temps. On sait que pour une onde quasi-monochromatique de pulsation ω et de durée T , la probabilité de transition fait apparaître une pseudo-fonction de Dirac $\delta_T(E_f - E_i - \hbar\omega)$ traduisant la conservation de l'énergie au sens de Bohr :

$$E_f = E_i + \hbar\omega \quad (\text{absorption}) . \quad (7.25)$$

¹⁴La molécule étant neutre, \vec{d} est invariant dans tout changement de l'origine du repère.

¹⁵Voir plus loin la description qualitative de l'effet Raman.

¹⁶d'un point de vue purement mécanique. En temps utile, il faudra prendre en compte l'aspect *statistique* décrit par les probabilités de Boltzmann de l'état de départ.

¹⁷Pour un champ électrique polarisé circulairement autour de Oz , la règle correspondante est $\Delta M = \pm 1$.

Le facteur radial de (7.29) est :

$$\left\langle \frac{1}{R} f_v | d(R_0) + q_{\text{eff}}(R - R_0) | \frac{1}{R} f_{v'} \right\rangle . \quad (7.32)$$

Les *bra - kets* désignent une intégration du genre $\int_0^{+\infty} R^2(\#)dR$: compte tenu des deux facteurs $\frac{1}{R}$, il apparaît strictement un élément de matrice d'oscillateur à une dimension :

$$\langle f_v | d(R_0) + q_{\text{eff}}(R - R_0) | f_{v'} \rangle = d(R_0) \delta_{vv'} + q_{\text{eff}} \langle f_v | (R - R_0) | f_{v'} \rangle , \quad (7.33)$$

avec maintenant $\langle f | (\#) | g \rangle = \int_0^{+\infty} dR f^*(R)(\#)g(R)$. On sait de l'étude de l'oscillateur harmonique que la coordonnée est proportionnelle à la somme $a + a^\dagger$: la matrice de la coordonnée sur les états propres du Hamiltonien n'a donc d'éléments de matrice non-nuls qu'entre deux états différant d'un quantum. D'où la troisième règle de sélection¹⁸ :

$$v' = v \quad \text{ou} \quad v' = v \pm 1 . \quad (7.34)$$

Spectre de rotation pure

Les transitions $v' = v$ se font donc sans changement de l'état de vibration : elles forment ce que l'on appelle le spectre de rotation pure. Compte tenu de (7.19), deux niveaux consécutifs de rotation pure sont séparés en énergie par (voir figure 7.3) :

$$B_m h[(L+1)(L+2) - L(L+1)] = 2B_m h(L+1) \quad (L \geq 0) . \quad (7.35)$$

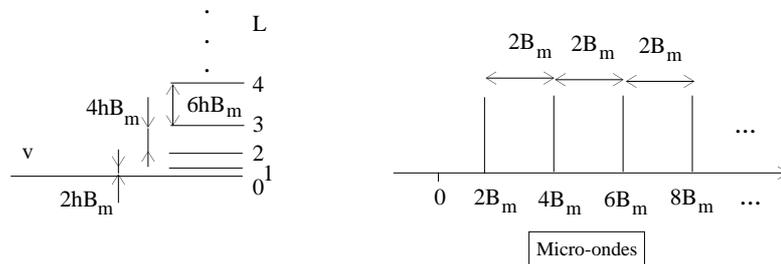


Figure 7.3: Schéma des transitions donnant le spectre de rotation pure ($v \rightarrow v, L \longleftrightarrow L+1$).

En vertu de la règle $\Delta L = \pm 1$, le spectre de rotation pure est donc une série de raies impliquant les différences d'énergie $hB_m[(L+1)(L+2) - L(L+1)] = hB_m(2L+2)$ ($L \geq 0$). Elles sont donc équidistantes, centrées en fréquence aux points $2B_m, 4B_m, 6B_m, \text{etc.}$ Outre la dépendance vis-à-vis de L et v des éléments de matrice, les intensités sont pondérées par les poids de Boltzmann¹⁹ donnant les populations thermiques des états de départ ; comme $g_L = 2L+1$ est la dégénérescence de chaque niveau L de rotation, ces probabilités sont de la forme $(2L+1)e^{-A L(L+1)} \sim (2L+1)e^{-A L^2}$. La distribution de probabilité présente donc un maximum quand L varie, survenant pour $L^2 hB \sim k_B T$.

Spectre de vibration-rotation

Les transitions $v' = v \pm 1$ forment le spectre dit de vibration - rotation. Compte tenu des ordres de grandeurs, les degrés de liberté de rotation donnent en fait une structure fine aux raies de vibration, lesquelles se situent dans l'infra-rouge. On peut distinguer deux groupes de transitions :

¹⁸ À la petite réserve près énoncée dans la note 11.

¹⁹ À l'ambiante, on a $k_B T \simeq 25$ meV. Cette énergie est nettement plus grande que l'échelle d'énergie donnant les différences d'énergie rotationnelle ($\nu_{\text{rot}} \sim 5 \times 10^{10}$ Hz \longleftrightarrow 0.2 meV). Il en résulte qu'à l'ambiante, les états rotationnels sont effectivement distribués.

1. $L \rightarrow L + 1$, correspondant aux différences d'énergie :

$$\hbar\omega_m + B_m h[(L + 1)(L + 2) - L(L + 1)] = \hbar\omega_m + 2B_m h(L + 1) . \quad (L \geq 0) \quad (7.36)$$

2. $L \rightarrow L - 1$, correspondant aux différences d'énergie :

$$\hbar\omega_m + B_m h[(L - 1)L - L(L + 1)] = \hbar\omega_m - 2B_m h L . \quad (L \geq 1) \quad (7.37)$$

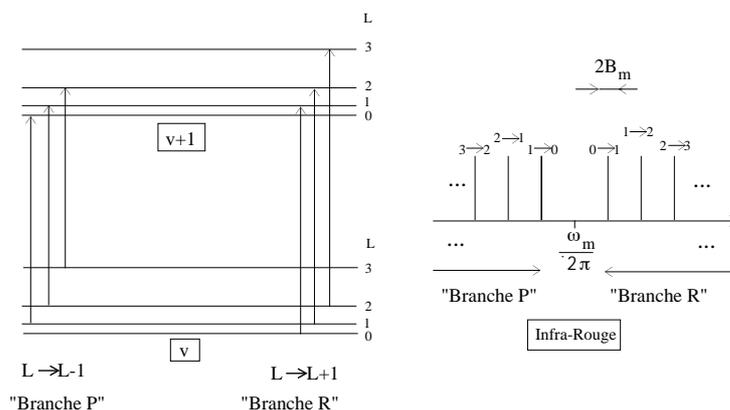


Figure 7.4: Schéma des transitions donnant le spectre de vibration-rotation ($v \rightarrow v + 1$, $L \rightarrow L \pm 1$).

Au total, on obtient deux groupes de raies équidistantes, séparées de $2B_m$ en fréquence, symétriques de part et d'autre de $\nu_m = \frac{\omega_m}{2\pi}$. Finalement, tout se passe comme si la raie ($\Delta v = \pm 1$, $\Delta L = 0$) (qui n'existe pas), était décomposée par la structure fine due à la rotation (voir fig. 7.4). Tout comme pour la vibration pure, les intensités sont pondérées par des poids de Boltzmann ($\nu_{\text{vibr}} \sim 5 \times 10^{12}$ Hz \longleftrightarrow 20 meV).

Ce qui précède rend compte, pour l'essentiel, des spectres de vibration - rotation. Il existe bien sûr diverses corrections à effectuer le cas échéant pour expliquer certaines finesses. Le plus souvent, la plus importante d'entre elles s'obtient en développant le terme centrifuge :

$$\frac{\hbar^2 L(L + 1)}{2\mu R^2} = \frac{\hbar^2 L(L + 1)}{2\mu} \left[\frac{1}{R_0^2} - 2(R - R_0) \frac{1}{R_0^3} + 3(R - R_0)^2 \frac{1}{R_0^4} + \dots \right] . \quad (7.38)$$

Le terme linéaire décale légèrement la position d'équilibre de R_0 en \tilde{R}_0 , cependant que la correction quadratique "renormalise" la pulsation ω_m en $\tilde{\omega}_m$. Enfin, il arrive aussi que l'on doive considérer les termes anharmoniques, c'est-à-dire introduire le terme en $(R - R_0)^3$ dans le développement de E_m autour de R_0 : alors les niveaux de vibration ne sont plus strictement équidistants (les écarts entre deux niveaux consécutifs diminuent quand on monte en énergie²⁰). Le calcul convenable des diverses corrections exige du savoir-faire : il faut d'une part effectuer des calculs de perturbation aux ordres successifs, d'autre part comparer entre elles les importances relatives des différentes perturbations.

7.2 Effet Raman

On a vu qu'une molécule homonucléaire est inactive dans l'IR ou les micro-ondes : dans cette gamme spectrale, tous les éléments de matrice (dipolaires électriques) sont nuls. Pour obtenir le même type d'informations que celles obtenues sur les molécules AB par les moyens décrits ci-dessus, il convient donc de s'y prendre autrement, en exploitant un effet, dit effet Raman, dont le principe est donné ci-après en raisonnant semi-classiquement.

²⁰Retenir l'image : plus le potentiel se raidit, plus l'écart entre deux niveaux d'énergie consécutifs augmente, et inversement. Le spectre de l'oscillateur harmonique est uniforme (écarts constants, E_n varie comme n), les énergies du puits carré infini augmentent comme n^2 (écarts $E_{n+1} - E_n \propto n$). Le puits infini est visiblement plus raide que le puits harmonique.

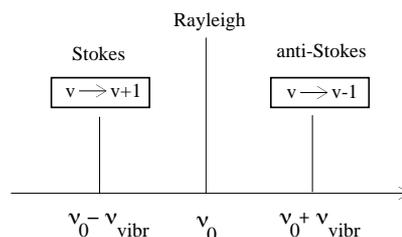


Figure 7.5: Schéma du spectre Raman de vibration ; quand la molécule gagne un *quantum* de vibration ($v \rightarrow v + 1$), c'est au détriment du photon diffusé, qui a une fréquence plus *petite* que le photon incident, et inversement.

On connaît la diffusion Rayleigh, qui est la mise en vibration forcée des électrons d'un atome ou d'une molécule, à des fréquences petites devant la plus petite fréquence de résonance. D'une façon générale, lorsqu'un champ électromagnétique est envoyé sur un tel système, et si sa fréquence²¹ ν_0 ne coïncide pas avec une résonance, on obtient une diffusion élastique des photons dont la section efficace varie comme ν_0^4 à basse fréquence (diffusion Rayleigh) et tend vers une constante²² σ_T à haute fréquence (diffusion élastique des rayons X, par exemple). Sous l'effet du champ excitateur, le petit dipôle induit est mis en oscillation forcée et rayonne à la fréquence ν_0 : la quantité importante est donc la *polarisabilité* de l'atome ou de la molécule. Cette polarisabilité dépend évidemment de la géométrie de la molécule ; si les noyaux sont fixes, une seule fréquence existe, c'est ν_0 et l'on obtient une simple diffusion élastique.

À l'inverse, si les noyaux vibrent à la fréquence ν_{vibr} , le dipôle induit sera (lentement) modulé en amplitude à la fréquence ν_{vibr} , très petite devant ν_0 ([8], complément A_V). La combinaison des ces oscillations fait alors apparaître naturellement deux composantes supplémentaires $\nu_0 \pm \nu_{\text{vibr}}$: outre la raie élastique Rayleigh (toujours présente, la molécule ne change pas d'état vibrationnel durant la diffusion du photon), deux raies symétriquement déplacées apparaissent (raies inélastiques), l'une en $\nu_0 + \nu_{\text{vibr}}$ (raie dite anti-Stokes), l'autre en $\nu_0 - \nu_{\text{vibr}}$ (raie dite Stokes). Ces deux raies correspondent au cas où la diffusion modifie l'état vibrationnel de la molécule ($v \rightarrow v \pm 1$). La raie Stokes ($\nu < \nu_0$) est celle où la molécule gagne un *quantum* de vibration au détriment du photon, la raie anti-Stokes ($\nu > \nu_0$) celle où la molécule en perd un et où on récupère un photon diffusé de plus grande énergie (voir fig. 7.5).

Le même effet vaut pour la rotation ([8], complément C_{VI}) : une molécule qui tourne sur elle-même présente un moment dipolaire oscillant (à une fréquence *double* : quand la molécule homonucléaire a fait un demi-tour, le dipôle induit est revenu à l'identique). On attend donc aussi un effet Raman, et donc des raies Stokes et anti-Stokes, sur les spectres de rotation. Une différence existe cependant par rapport à la vibration : alors que l'oscillateur a une fréquence unique (le spectre d'énergie est une échelle à barreaux équidistants), il n'en va pas de même pour un rotateur. Il existe donc *a priori* une infinité de fréquences $\nu_0 \pm 2\nu_{\text{rot}}$ où $\nu_{\text{rot}} = 2B, 4B, \dots$, de sorte qu'il existe – *modulo* la règle de sélection²³ $\Delta L = 0, \pm 2$ – un très grand nombre de raies visibles de chaque sorte (Stokes ou anti-Stokes) (voir [8], complément C_{VI}).

7.3 Molécules polyatomiques

7.3.1 Coordonnées normales

Pour une molécule ayant $N > 2$ noyaux, il existe $N_V = 3N - 6$ ou $3N - 5$ degrés de liberté de vibration, notés R_j . Le jeu des R_j conduit à une *déformation* de la molécule relativement à une configuration d'équilibre stable,

²¹En pratique, pour la spectroscopie Raman, ν_0 est une fréquence optique.

²² $\sigma_T = 8\pi \frac{r_e^2}{3}$, où $r_e = \frac{e^2}{mc^2}$ est le rayon classique de l'électron.

²³avec toujours $\Delta M = 0$.

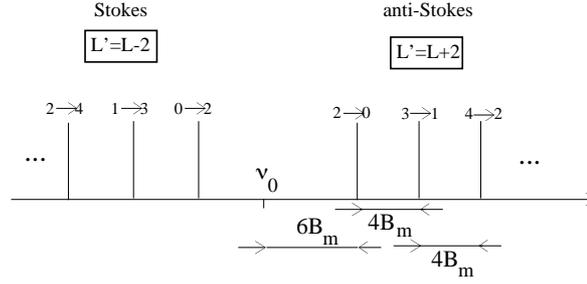


Figure 7.6: Schéma du spectre Raman de rotation ; quand la molécule passe de L à $L + 2$, la fréquence du photon diffusé est plus *petite* que ν_0 , et inversement.

prise comme référence :

$$E_m(R_j) \simeq E_m(R_{j0}) + \sum_{j,j'=1}^{N_V} \frac{1}{2} (R_j - R_{j0})(R_{j'} - R_{j'0}) \left(\frac{\partial^2 E_m}{\partial R_j \partial R_{j'}} \right)_0, \quad (7.39)$$

configuration autour de laquelle les noyaux effectuent de petites vibrations approximées par le développement harmonique ci-dessus. Comme la configuration est supposée stable, la fonction jouant le rôle d'une énergie potentielle harmonique est une forme quadratique définie positive que l'on peut diagonaliser en formant les bonnes combinaisons linéaires des R_j , appelées coordonnées normales et notées X_l . Au total, l'énergie électronique, dans l'approximation harmonique, peut se mettre sous forme diagonale :

$$E_m(R_j) = E_m(R_{j0}) + \sum_{l=1}^{N_V} \frac{1}{2} k_l X_l^2. \quad (7.40)$$

Toutes les constantes de raideur k_l sont strictement positives. En effectuant le même changement de variables dans les termes cinétiques constituant T_N , l'équation pour la fonction nucléaire prend la forme :

$$\left(\sum_{l=1}^{N_V} H_{\text{harm},l} + H_{\text{rot}} \right) \chi = E \chi. \quad (7.41)$$

Le premier terme est une somme de N_V oscillateurs harmoniques à une dimension. Le second est le Hamiltonien du rotateur rigide constitué par la molécule dont les noyaux sont fixés aux positions d'équilibre²⁴ et incorpore les derniers degrés de liberté, la translation en bloc étant toujours mise de côté. Notons qu'une vibration normale implique en général tous les noyaux : c'est donc bien un mouvement *collectif*²⁵ de ces derniers.

Remarque

L'introduction des coordonnées normales peut aussi s'effectuer de façon plus formelle. On part de :

$$H_N = \sum_{i=1}^N \frac{1}{2} M_i \vec{V}_i^2 + E_m(Q_j). \quad (7.42)$$

Posant :

$$\vec{s}_j = \sqrt{M_j} \vec{V}_j, \quad \vec{s}_j = \sqrt{M_j} (s_j - s_{j0}), \quad (7.43)$$

le terme cinétique devient :

$$2T_N = \sum_{i=1}^N \sum_{u=x,y,z} s_{ju}^2 \equiv \sum_{\lambda=1}^{3N} \dot{s}_\lambda^2, \quad (7.44)$$

²⁴On néglige ici tout effet de couplage entre vibration et rotation.

²⁵analogue à celui que l'on rencontrera pas la suite à propos des vibrations de réseau, dont la quantification produit les *phonons*.

d'où :

$$H_N = \frac{1}{2} \sum_{\lambda=1}^{3N} \dot{s}_\lambda^2 + V(s_\lambda) \quad V(s_\lambda) = E_m(Q_j(s_\lambda)) . \quad (7.45)$$

En développant V autour de la position d'équilibre :

$$V(s_\lambda) = \sum_{\lambda,\mu=1}^{3N} \frac{1}{2} f_{\lambda\mu} s_\lambda s_\mu , \quad (7.46)$$

on trouve finalement :

$$H_N = \frac{1}{2} \sum_{\lambda=1}^{3N} \dot{s}_\lambda^2 + \sum_{\lambda,\mu=1}^{3N} \frac{1}{2} f_{\lambda\mu} s_\lambda s_\mu . \quad (7.47)$$

Maintenant, V est diagonalisé par une transformation orthogonale :

$$X_l = \sum_{\lambda=1}^{3N} O_{l\lambda} s_\lambda . \quad (7.48)$$

Cette transformation laisse invariant le carré de la norme ; il en résulte que le terme cinétique s'exprime comme suit :

$$T_N = \frac{1}{2} \sum_{\lambda=1}^{3N} \dot{X}_l^2 , \quad (7.49)$$

cependant que, par définition, l'énergie potentielle devient :

$$V = \frac{1}{2} \sum_{\lambda,\mu=1}^{3N} \frac{1}{2} k_l X_l^2 . \quad (7.50)$$

On peut démontrer – mais c'est une évidence physique – qu'il existe 6 (5) constantes k_l qui sont nulles, associées à la rotation et à la translation en bloc. Autrement dit, l'équation caractéristique :

$$\begin{vmatrix} f_{11} - \Lambda & f_{12} & \dots & f_{1\ 3N} \\ f_{21} & f_{22} - \Lambda & \dots & f_{2\ 3N} \\ \vdots & \vdots & \ddots & \vdots \\ f_{3N1} & f_{3N2} & \dots & f_{3N\ 3N} - \Lambda \end{vmatrix} \quad (7.51)$$

possède 6 (5) fois la valeur propre $\Lambda = 0$.

La détermination des coordonnées normales est grandement facilitée par l'utilisation explicite de la symétrie moléculaire (rotations, réflexions, ...). En effet, soit S une opération de symétrie de la configuration d'équilibre ; par nature, effectuer cette opération ne change rien et, en particulier, ne modifie pas l'énergie potentielle. On a donc d'une part ($\Delta Q_j = Q_j - Q_{j0}$) :

$$E_m(SQ_j) - E_m(SQ_{j0}) = \sum_{j,j'} \frac{1}{2} (S\Delta Q_j) (S\Delta Q_{j'}) H_{jj'} . \quad (7.52)$$

et, d'autre part :

$$E_m(SQ_j) - E_m(SQ_{j0}) = E_m(Q_j) - E_m(Q_{j0}) . \quad (7.53)$$

Par comparaison, on en déduit :

$$\sum_{j,j'} (S\Delta Q_j) (S\Delta Q_{j'}) H_{jj'} = \sum_{j,j'} \Delta Q_j \Delta Q_{j'} H_{jj'} . \quad (7.54)$$

L'opération de symétrie est une transformation linéaire, représentée par une matrice orthogonale²⁶ :

$$S\Delta Q_j = \sum_k S_{jk} \Delta Q_k \quad {}^t S = S^{-1} . \quad (7.55)$$

²⁶Puisque toute opération de symétrie conserve les angles et les distances.

D'où, successivement :

$$\sum_{j,j'} \sum_{k,k'} S_{jk} \Delta Q_k S_{j'k'} \Delta Q_{k'} H_{jj'} = \sum_{j,j'} \Delta Q_j \Delta Q_{j'} H_{jj'} , \quad (7.56)$$

$$\sum_{k,k'} \left(\sum_{j,j'} S_{jk} H_{jj'} S_{j'k'} \right) \Delta Q_k \Delta Q_{k'} = \sum_{j,j'} \Delta Q_j \Delta Q_{j'} H_{jj'} , \quad (7.57)$$

$$\sum_{k,k'} ({}^t S H S)_{kk'} \Delta Q_k \Delta Q_{k'} = \sum_{j,j'} \Delta Q_j \Delta Q_{j'} H_{jj'} . \quad (7.58)$$

Les indices étant muets, on en déduit :

$$({}^t S H S)_{jj'} = H_{jj'} \iff S^{-1} H S = H \iff H S = S H ; \quad (7.59)$$

la dernière égalité n'exprime rien d'autre que, S étant une opération de symétrie, H et S commutent. Maintenant, soit X_l une certaine combinaison linéaire propre de S :

$$S X_l = \sigma_l X_l . \quad (7.60)$$

Comme $[H, S] = 0$, on a :

$$H S X_l = H(\sigma_l X_l) = \sigma_l H X_l = S H X_l \iff S(H X_l) = \sigma_l(H X_l) . \quad (7.61)$$

Ceci montre que si X_l est propre de S , alors $H X_l$ est encore propre de S , avec la même valeur propre σ_l . Si σ_l n'est pas dégénérée, alors $S X_l$ est proportionnel à X_l , c'est-à-dire que X_l est aussi propre de H :

$$H X_l = h_l X_l . \quad (7.62)$$

Autrement dit, X_l est l'une des directions principales de la forme quadratique donnant E_m dans l'approximation harmonique.

Si σ_l est dégénérée, deux fois par exemple, il existe deux combinaisons linéaires telles que :

$$S X_{l1} = \sigma_l X_{l1} , \quad S X_{l2} = \sigma_l X_{l2} . \quad (7.63)$$

On a toujours :

$$S(H X_{l1}) = \sigma_l(H X_{l1}) , \quad S(H X_{l2}) = \sigma_l(H X_{l2}) . \quad (7.64)$$

$H X_{li}$ ($i = 1, 2$) est propre de S avec la valeur propre σ_l ; donc chaque $H X_{li}$ est une combinaison linéaire de X_{l1} et X_{l2} : on peut donc diagonaliser H dans ce sous-espace sans pour autant dédiagonaliser la restriction de S à ce sous-espace. Donc, même en cas de dégénérescence, on rend simultanément diagonale les deux matrices de H et S .

Ainsi, au lieu de procéder directement à la diagonalisation de la matrice de l'énergie potentielle, mieux vaut considérer d'emblée les opérations de symétrie et rechercher – au moyen des outils de la Théorie de la Représentation Linéaire des Groupes – les combinaisons propres, ce qui assure une diagonalisation (ou au moins une prédiagonalisation) de l'énergie potentielle (7.39).

7.3.2 Rotation des molécules polyatomiques

Pour terminer, disons un mot de la rotation en bloc (de la configuration d'équilibre) – qui donne lieu à une quantification en raison du caractère périodique des variables angulaires et de la nécessité pour la fonction d'onde d'être monovaluée. En tant que corps solide, la molécule dans sa configuration d'équilibre possède un tenseur d'inertie I_{uv} ($u, v = x, y, z$), dont les éléments sont définis comme suit :

$$I_{xx} = \sum_{\text{Noyaux}} M_i (y_i^2 + z_i^2) , \quad I_{xy} = \sum_{\text{Noyaux}} M_i x_i y_i , \quad \dots , \quad (7.65)$$

où les coordonnées sont définies par rapport à un repère lié à la molécule et dont l'origine est le centre de masse G . Ce tenseur d'inertie est réel symétrique et peut donc être diagonalisé, ce qui conduit à trois axes principaux X , Y et Z . Sur ces directions, le tenseur d'inertie est représenté par la matrice diagonale :

$$\begin{bmatrix} I_{XX} & 0 & 0 \\ 0 & I_{YY} & 0 \\ 0 & 0 & I_{ZZ} \end{bmatrix} . \quad (7.66)$$

Si les trois moments d'inertie I_{UU} sont égaux entre eux, le rotateur est dit sphérique (ou isotrope)²⁷. Si seulement deux d'entre eux sont égaux entre eux, on parle de toupie symétrique, pour évoquer un objet ayant une symétrie de révolution cylindrique. Enfin, si les trois moments principaux sont tous différents, on a une toupie asymétrique.

L'énergie classique de rotation est purement cinétique (en l'absence de tout champ extérieur) et est de la forme :

$$\frac{1}{2}I_{XX}\omega_{XX}^2 + \frac{1}{2}I_{YY}\omega_{YY}^2 + \frac{1}{2}I_{ZZ}\omega_{ZZ}^2 \quad (7.67)$$

où les ω_{UU} sont les composantes du vecteur rotation $\vec{\Omega}$, donné par les vitesses angulaires de rotation autour des trois axes principaux. C'est bien parce que le tenseur d'inertie est diagonal vis-à-vis de ces derniers que l'énergie cinétique est une simple somme de trois termes indépendants les uns des autres.

²⁷On emploie aussi l'expression "toupie symétrique".

Partie III

Physique de la matière condensée

Chapitre 8

Matière condensée ordonnée

8.1 Classification des solides ordonnés

Il s'agit ici par définition des solides possédant un ordre à longue portée, en principe infinie. Un tel solide résulte du rapprochement d'entités bien définies avant la formation du cristal : atomes, molécules, ions. Bien sûr, la constitution du solide en tant qu'édifice stable s'accompagne d'altérations plus ou moins grandes de ces briques élémentaires constitutives ; on admettra "qu'il en reste toujours quelque chose", ce qui est presque une évidence sur le plan physique. Il y a donc dans le solide la trace des constituants à l'état libre : on parle alors de "cœur" pour désigner ce fantôme du constituant isolé. Les électrons complémentaires sont traditionnellement appelés électrons de valence en ce qui concerne ceux qui sont principalement responsables de la cohésion du solide, ou, s'il en reste, électrons de conduction puisque, presque libres, ils sont aptes à conduire le courant électrique.

Ceci étant, une première grande classification se crée spontanément sur la considération des propriétés électriques permettant de séparer nettement conducteurs et isolants, sur la base de leur résistivité électrique. La véritable explication de cette distinction ressort de la Mécanique Quantique et s'exprime par des considérations dans l'espace des impulsions. A défaut d'un critère précis défini sur la considération de grandeurs dans l'espace physique, il est cependant possible d'effectuer une distinction qualitative dans l'espace réel : en général, dans un métal, la distribution électronique est nettement plus diffuse (diluée, délocalisée) que dans les isolants, où les électrons de valence restent davantage situés près des cœurs, ou entre deux cœurs, pour fabriquer la "colle" assurant la stabilité de la liaison chimique.

Il s'avère très utile d'établir une sous-classification pour les isolants, fondée, elle, précisément sur la densité électronique. On distingue ainsi principalement :

- Les solides covalents, où l'on retrouve les caractéristiques *directionnelles* des liaisons chimiques. Le prototype de cette espèce de solide est le diamant, constitué d'atomes de carbone répartis dans l'espace suivant deux réseaux cubiques à faces centrées entremêlés, résultat de la valence 4 du carbone (structure tétraédrique¹). D'autres exemples : le silicium et le germanium, ou même le carbone sous sa forme graphite² – qui ont pourtant des propriétés électriques très différentes du diamant ! Alors que le diamant est un isolant remarquable, on sait bien que Ge et Si – même purs – sont des semi-conducteurs.

¹C'est alors l'hybridation sp^3 qui est pertinente.

²expliquée par l'hybridation sp^2 produisant les angles de 120° .

La différence tient bien sûr à la structure de bande et aux valeurs *numériques* des gaps d'énergie (à l'ambiante³ : 0.67 eV et 1.12 eV pour Ge et Si respectivement, 5.5 eV pour le diamant⁴).

- Les solides moléculaires, qui sont fondamentalement une juxtaposition d'atomes ou molécules très peu perturbés par rapport à leur état libre en phase diluée. Il y a peu de densité électronique entre les atomes. Les prototypes sont les solides de gaz rares (Néon, Argon, Krypton, Xénon⁵) et les solides de certaines molécules aromatiques (anthracène, par exemple).
- Les solides ioniques, par exemple le sel ordinaire (NaCl). Ils sont formés à partir de couples d'atomes ; l'un des deux est de type métallique (il perd un ou des électrons lors de la formation du solide, c'est ce que fait Na), alors que l'autre est très avide d'électrons (très électronégatif, "antimétallique", c'est le cas de Cl). Les électrons échangés restent très bien localisés près des centres, comme dans un cristal moléculaire, ce qui produit une structure formées de paires d'ions ; mais la réorganisation est importante par rapport à la structure des deux partenaires non-appariés, puisque l'un a perdu des électrons alors que l'autre en a gagné. Au total, les électrons restent bien localisés, l'interaction entre ions n'est pas écrantée et c'est donc l'interaction de Coulomb nue qui est à l'œuvre – bien évidemment complétée par une répulsion de cœur dur.

En tout état de cause, ce sont bien sûr les électrons plus ou moins partagés entre les "briques" du solide qui en constituent la colle, le ciment. Selon le cas, on parlera de liaison métallique, liaison covalente ou liaison ionique. En ce qui concerne les solides moléculaires, il s'agit plutôt d'une pseudo-liaison chimique (fragile) résultant d'une interaction du type van der Waals, interprétable en termes d'échanges de photons (virtuels) entre deux objets (atomes, molécules) globalement neutres. De façon plus élémentaire, il s'agit d'une interaction dipôle induit - dipôle induit : la présence du premier partenaire (et des charges électriques qu'il renferme) induit un moment dipolaire électrique d_2 au sein du deuxième partenaire, *via* la polarisabilité α :

$$d_2 \sim \alpha \mathcal{E} = \alpha \frac{d_1}{R^3} . \quad (8.1)$$

Comme l'interaction entre deux dipôles est du genre $\frac{d_1 d_2}{R^3}$, on en déduit une énergie d'interaction attractive variant comme $\frac{d_1}{R^3} \alpha \frac{d_1}{R^3} \propto R^{-6}$. Par $\vec{F} = -\vec{\nabla} V$, il en résulte que les forces de van der Waals décroissent comme R^{-7} – aux distances grandes par rapport à l'échelle atomique, mais pas trop grandes : quand les photons (virtuels) – qui sont les vecteurs de l'interaction électromagnétique – mettent trop longtemps pour aller d'un atome à l'autre, il apparaît des corrections donnant finalement une décroissance plus rapide avec la distance (par exemple : $V \propto R^{-7}$, force en R^{-8}).

8.2 Énergie de cohésion des solides ordonnés

8.2.1 Généralités

L'énergie de cohésion d'un solide est l'énergie qu'il faut fournir pour le séparer en ses constituants à l'état libre. Une telle définition est clairement ambiguë, car tout dépend des fragments que l'on définit. D'une façon générale, on choisit les éléments libres au vu de considérations physiques mais tout choix reste forcément un peu conventionnel ; les distinctions qui en résultent sont de peu d'intérêt. Par exemple, s'agissant de l'azote solide, on pourrait décider de décomposer le solide en molécules N_2 plutôt qu'en atomes N , mais peu importe au fond, puisque l'on connaît l'énergie de dissociation de la molécule N_2 .

D'une façon générale, un solide se forme par *condensation* d'objets élémentaires (typiquement des atomes ou des petites molécules) initialement en phase vapeur. Lors de la formation du solide, ces entités peuvent

³Au voisinage de l'ambiante, le gap varie linéairement en température. Ceci provient de la dilatation du réseau, ce qui induit une expansion du potentiel périodique.

⁴Ces valeurs ne sont pas d'un ordre de grandeur différent, mais rentrent dans des exponentielles, d'où les ordres de grandeurs différents pour la conductivité – sans compter avec les valeurs prises par la densité électronique.

⁵L'hélium fait bande à part : en vertu de sa petite masse, les effets quantiques ne peuvent jamais être ignorés et toute tentative de description classique est vouée à l'échec.

perdre une partie de leur intégrité, ce qui donne lieu précisément aux différents types de solides (moléculaires, ioniques, métalliques, *etc.*), comme brièvement discuté ci-dessus. On continuera ici à désigner par *cœurs* ce qui reste des objets libres après ce réarrangement ; en général, ceux-ci se définissent d'eux-mêmes sans aucune ambiguïté.

Afin de rester au niveau le plus simple, la discussion ci-dessous de l'énergie de cohésion est faite dans l'hypothèse où les cœurs sont fixes, et sont situés aux nœuds d'un réseau géométrique : on n'envisagera donc que des solides ordonnés. Il s'agit bien sûr d'une approximation, à la fois qualitative et quantitative : d'après la Mécanique Quantique, une particule ne peut être localisée strictement et est soumise aux inégalités de Heisenberg du genre $\Delta x \Delta p_x \sim \hbar$. Il en résulte que pour une particule de masse M , à une incertitude en position Δx correspond une énergie $\Delta E \sim \frac{\Delta p^2}{2M} \sim \frac{\hbar^2}{M\Delta x^2}$ – appelée énergie de point zéro quand il s'agit de vibrations – et qui, en toute rigueur, devrait être prise en compte dans le calcul de l'énergie de cohésion. Pour fixer les idées, avec $\Delta x \sim 0.5 \text{ \AA}$, et pour un atome de carbone, on trouve $\Delta E \sim 0.05 \text{ eV}$, ce qui est loin d'être négligeable devant l'énergie par atome que l'on trouvera dans l'hypothèse où le réseau est rigide. Clairement, traiter les cœurs comme des objets ponctuels classiques sera, de ce point de vue, d'autant meilleure que leur masse est élevée.

Il est possible de construire une théorie élémentaire de la cohésion pour les solides moléculaires et ioniques, n'utilisant que les idées classiques. En revanche, rien de tel n'existe pour les solides covalents – sous-jacente, il s'agit en fait de la théorie de la liaison chimique. Quant aux métaux, la situation est difficile à décrire convenablement : on ne peut pas ignorer les électrons presque libres – qu'il faut traiter quantiquement – et leurs interactions, même si elles sont en général assez fortement écrantées.

8.2.2 Cohésion des solides moléculaires

Le prototype du solide moléculaire est le cristal de gaz rare. Excluant l'hélium – si léger que les effets quantiques de point zéro ne peuvent être ignorés – on se place dans l'approximation où les atomes constitutifs sont traités comme des objets ponctuels classiques. Par définition d'un cristal moléculaire, ses atomes sont très peu modifiés par rapport à leur état libre lors de la formation du cristal ; c'est le cas quand il s'agit d'atomes très robustes comme les gaz rares avec leurs couches électroniques complètes.

L'interaction entre deux objets neutres, robustes et bien localisés se compose de deux contributions dont l'origine physique est évidente :

1. une répulsion forte à courte distance, chaque atome ayant par nature un domaine propre réservé résistant à toute invasion extérieure, en particulier à l'investissement d'un autre atome
2. une interaction attractive à longue portée, connue sous le nom d'interaction de van der Waals, donnant une énergie potentielle décroissant comme R^{-6} :

$$V(R) \simeq -\frac{C^{\text{ste}}}{R^6} \quad (R \gg a_0) . \quad (8.2)$$

Au total, l'interaction entre deux atomes neutres a l'allure représentée sur la figure (8.1).

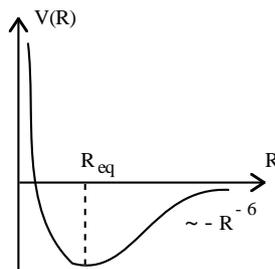


Figure 8.1: Allure de l'énergie potentielle d'interaction entre deux objets neutres.

Autant la forme à longue distance est parfaitement déterminée, et en un sens universelle, autant la répulsion de cœur varie d'un élément à l'autre. La modélisation précise de cette dernière n'a pas beaucoup d'importance (ce qui importe, c'est d'inclure de fait une répulsion), puisqu'elle intervient sur une petite échelle de distance, très peu "explorée" pour le problème ici analysé⁶. Le choix d'une dépendance en R ou d'une autre est largement dicté par le souci de la simplicité analytique. Un choix très courant consiste à poser :

$$V(R) \simeq + \frac{A}{R^{12}} \quad (R \lesssim a_0) , \quad (8.3)$$

et on obtient alors le "potentiel 6-12" de Lennard - Jones :

$$V(R) = 4v_0 \left[\left(\frac{d}{R} \right)^{12} - \left(\frac{d}{R} \right)^6 \right] \quad (v_0 > 0) . \quad (8.4)$$

Le 12 n'est finalement choisi que par commodité ($2 \times 6 = 12$!), aucune raison théorique ne le justifie, la seule contrainte est qu'il soit fortement répulsif⁷ – et il l'est. La distance d'équilibre pour une paire unique est $R_{\text{eq}} = 2^{1/6} d \simeq 1.12d$, ce qui donne le sens physique du paramètre d dans (8.4) ; quant à l'énergie de dissociation d'une paire, elle vaut $V(R_{\text{eq}})$; avec le calibrage (8.4), on a tout simplement $V(R_{\text{eq}}) = -v_0$.

L'existence d'une interaction entre les atomes entraîne que le gaz qu'ils constituent ne saurait être parfait. On peut donc espérer trouver les deux paramètres phénoménologiques⁸ d et v_0 en mesurant l'écart à la loi $PV = Nk_B T$. De fait, il existe un développement de l'équation d'état des fluides en puissances de la densité ρ de particules ($\rho = \frac{N}{V}$), appelé développement du Viriel. Celui-ci permet d'exprimer toutes les grandeurs thermodynamiques sous la forme d'un développement du même type ; en particulier, la pression P apparaît maintenant sous la forme :

$$P = N \frac{k_B T}{V} [1 - B(T)\rho - C(T)\rho^2 - \dots] . \quad (8.5)$$

Dans cette expression, le premier terme redonne la loi des gaz parfaits. Les fonctions $B(T)$, $C(T)$, \dots , sont par définition les coefficients du Viriel⁹, et s'expriment à l'aide d'intégrales impliquant justement l'interaction entre deux atomes ; par exemple :

$$B = 2\pi \int_0^{+\infty} r^2 dr [e^{-\beta V(r)} - 1] . \quad (8.6)$$

La détermination de la quantité¹⁰ B (via la mesure des isothermes par exemple) permet de caractériser $V(R)$ et donne donc accès à v_0 et d – étant à nouveau entendu que des calculs de chimie quantique permettent aussi d'obtenir des valeurs numériques plus ou moins précises. B dépend de la température, mais pas de la densité ρ : ceci permet de comprendre que les valeurs obtenues par l'étude du gaz conservent toute leur utilité pour l'étude de la phase condensée. Quelques valeurs numériques :

	Ne	Ar	Kr	Xe	
v_0 (eV)	0.0031	0.0104	0.0140	0.0200	.
R_{eq} (Å)	2.74	3.40	3.65	3.98	

Comme le choix de la partie répulsive n'est pas critique, il arrive aussi souvent que l'on prenne un potentiel de sphère dure pour la représenter. Alors :

$$V(R) = \begin{cases} +\infty & \text{si } R < d \\ -\frac{C^{\text{ste}}}{R^6} & \text{si } R > d \end{cases} , \quad (8.8)$$

où d est une longueur atomique, d'ordre a_0 (voir figure 8.2).

⁶Tôt ou tard, des exponentielles de Boltzmann du genre $e^{-\beta V(R)}$ interviennent dans des intégrales ; à très courte distance et aux températures usuelles, ces exponentielles tendent vers zéro à toute vitesse et viennent donc "tuer" toute contribution à l'intégrale ; de ce fait, la forme précise de $V(R)$ à courte distance est en pratique sans grande importance.

⁷En tout cas, il doit diverger plus vite à l'origine que R^{-6} . D'une façon générale, les calculs quantiques avec les fonctions d'onde donnent plutôt une "divergence" exponentielle du genre $e^{2\alpha(R_{\text{eq}}-R)}$ ($R < R_{\text{eq}}$), que l'on trouve dans le potentiel de Morse $V(x) = -V_0[e^{-2\alpha(R-R_{\text{eq}})} - 2e^{-\alpha(R-R_{\text{eq}})}]$.

⁸Rappelons qu'en toute rigueur, seul le coefficient du terme en R^{-12} est d'essence phénoménologique. On peut en principe calculer le terme de van der Waals si l'on connaît les fonctions d'onde des atomes (ou des molécules).

⁹Il est clair qu'un tel développement n'a d'intérêt et d'utilité que pour des densités pas trop élevées. Par ailleurs, son existence est plausible physiquement, à condition d'être loin d'un point critique.

¹⁰Noter que B est une fonction de la température.

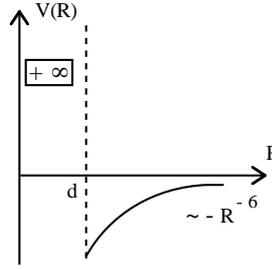


Figure 8.2: Allure de l'énergie potentielle d'interaction entre deux objets neutres avec un cœur dur d'extension d .

Dans tous les cas, les énergies de liaison par paire sont faibles – comme en attestent les valeurs de v_0 , comparées aux valeurs usuelles de la physique atomique ou de la chimie, ce qui n'est pas surprenant compte tenu de la nature physique d'un cristal moléculaire. En tout cas, c'est bien parce que les composants sont très peu altérés lors de la condensation que l'on peut utiliser les paramètres déduits de l'analyse en phase vapeur.

L'interaction d'une paire étant précisée, il convient maintenant de considérer effectivement un *cristal*, construit en disposant aux nœuds d'un réseau géométrique les atomes en interaction par $V(\vec{R})$. Si \vec{R}_i repère un point de ce réseau, l'interaction totale est :

$$E = \frac{1}{2} \sum_{\vec{R}_i, \vec{R}_j, i \neq j} V(\vec{R}_i - \vec{R}_j) , \quad (8.9)$$

où le facteur $\frac{1}{2}$ évite de compter deux fois l'interaction au sein d'une même paire. Dans la limite d'un cristal macroscopique contenant N ($\sim 10^{23}$) atomes, il y a une quasi-invariance par translation (discrète), l'énergie E est une grandeur extensive¹¹ et il vient simplement, aux termes de surface près :

$$E \rightarrow \frac{N}{2} \sum_{\vec{R} \neq \vec{0}} V(\vec{R}) \quad (N \gg 1) . \quad (8.10)$$

Dans cette limite, on peut donc définir une énergie de cohésion par particule, ε :

$$\varepsilon \equiv \frac{E}{N} = \frac{1}{2} \sum_{\vec{R}} V(\vec{R}) . \quad (8.11)$$

On peut aussi dire que, fixant une particule (à l'origine par exemple), son interaction avec toutes les autres est $\sum_{\vec{R} \neq \vec{0}} V(\vec{R})$, ce qui fournit $N \sum_{\vec{R} \neq \vec{0}} V(\vec{R})$ pour les N particules ; mais alors, on a compté deux fois l'interaction d'une même paire (i, j) , d'où le facteur $\frac{1}{2}$.

Si l'on choisit le potentiel (8.4), il faut donc calculer :

$$\varepsilon = 4v_0 \frac{1}{2} \sum_{\vec{R}} \left[\left(\frac{d}{R} \right)^{12} - \left(\frac{d}{R} \right)^6 \right] . \quad (8.12)$$

où la somme court sur tous les nœuds du réseau, chacun d'entre eux étant caractérisé par un vecteur \vec{R} . Par exemple, pour un réseau cubique de côté a , on doit calculer des sommes¹² du genre :

$$\left(\frac{d}{a} \right)^n \sum_{m_1, m_2, m_3} \frac{1}{\| m_1 \vec{i} + m_2 \vec{j} + m_3 \vec{k} \|^n} \equiv \left(\frac{d}{a} \right)^n A_n \quad (8.13)$$

¹¹L'argument ne vaut que dans la mesure où les forces ne sont pas à portée infinie : de ce fait, les effets de surface/volume s'évanouissent à la limite thermodynamique. En effet, si ξ désigne la portée des interactions, le nombre d'atomes de surface en interaction mutuelle est $\sim (N/L^3) \xi L^2$, pour un échantillon cubique de côté L , soit $N (\xi/L)$. Dans la limite thermodynamique, pour une grandeur extensive comme l'énergie, les corrections de surface s'annulent comme L^{-1} , soit comme $N^{-1/3}$ lorsque la densité est maintenue constante.

¹²dites *sommes de réseau*.

avec $n = 6, 12$ et où les m_i sont des entiers naturels ; le calcul de telles sommes (triples !) n'est pas si simple et des méthodes ingénieuses ont été mises au point pour accélérer la convergence.

Dans le cas d'un réseau quelconque, il est commode de poser $\vec{R} = l\vec{\rho}(\vec{R})$, où l est la distance entre plus proches voisins (par exemple, pour un cubique faces centrées (CFC) de maille conventionnelle de côté a , $l = a\frac{\sqrt{2}}{2}$, pour un cubique centré (CC), $l = a\frac{\sqrt{3}}{2}$). Avec ces notations, il faut calculer des sommes du genre S_n :

$$\left(\frac{d}{l}\right)^n \sum_{\vec{R} \in \mathcal{B}} \frac{1}{\|\vec{\rho}(\vec{R})\|^n} \equiv \left(\frac{d}{l}\right)^n A_n \quad (8.14)$$

Pour un réseau CFC (le cas des gaz rares), les sommes de réseau donnent :

$$A_6 \simeq 14.45, \quad A_{12} \simeq 12.13, \quad (8.15)$$

Il en résulte que l'énergie de cohésion par particule est :

$$\varepsilon = 2v_0 \left[A_{12} \left(\frac{d}{l}\right)^{12} - A_6 \left(\frac{d}{l}\right)^6 \right]. \quad (8.16)$$

Par minimisation, on trouve la distance d'équilibre l_{eq} , reliée en principe le côté de la maille cubique conventionnelle du cristal CFC constitué ($a_{\text{eq}} = \sqrt{2}l_{\text{eq}}$) :

$$l_{\text{eq}} = \left(\frac{2A_{12}}{A_6}\right)^{1/6} d \simeq 1.09 d, \quad a_{\text{eq}} \simeq 1.54 d. \quad (8.17)$$

La distance l_{eq} inférieure est à la valeur d'équilibre d'une seule paire ($\simeq 1.12 d$), traduisant l'augmentation de densité d'un "dimère" quand on passe en phase solide – indépendamment de la structure cristalline (compacte puisqu'il s'agit d'un CFC) : une grande quantité d'atomes qui tous s'attirent ont évidemment tendance à former un assemblage relativement compact – seule la répulsion à courte distance prévenant le *collapse*. Les valeurs numériques sont en assez bon accord avec l'expérience :

	Ne	Ar	Kr	Xe	
l_{eq} (expérimental, Å)	3.13	3.75	3.99	4.33	.
$l_{\text{eq}} = 1.09 d$ (Å)	2.99	3.71	3.98	4.34	

On constate d'ailleurs que l'accord est d'autant meilleur que l'atome est plus lourd : l'oubli des termes de point-zéro est bien de moins en moins préjudiciable quand la masse des atomes augmente. D'une façon générale, les vibrations irréductibles d'après Heisenberg jouent le rôle d'une répulsion effective : quand les atomes vibrent, ils occupent plus de place et leur distance d'équilibre se trouve, de fait, augmentée.

Une fois obtenue la distance d'équilibre, on trouve l'énergie de cohésion par particule en reportant l'expression (8.17) de l_{eq} dans (8.16) :

$$\varepsilon_{\text{eq}} = 2v_0 \left[A_{12} \left(\frac{A_6}{2A_{12}}\right)^2 - A_6 \frac{A_6}{2A_{12}} \right] = -v_0 \frac{A_6^2}{2A_{12}}. \quad (8.19)$$

Pour le CFC, on a ainsi :

$$\varepsilon_{\text{eq}} \simeq -8.60 v_0. \quad (8.20)$$

L'accord est convenable, comme le montre la table ci-dessous ; à nouveau, il est d'autant meilleur que l'atome est plus massif. Les énergies expérimentales sont plus élevées (en valeurs algébriques, l'énergie de liaison est plus faible, la liaison est donc plus fragile) puisque l'on a négligé une contribution *positive* à l'énergie, à savoir l'énergie de vibration¹³. On note la faiblesse de la liaison comparée à une liaison chimique ordinaire, plutôt de l'ordre de l'eV.

	Ne	Ar	Kr	Xe	
ε (expérimental, eV)	-0.02	-0.08	-0.11	-0.17	.
$\varepsilon_{\text{eq}} = -8.60 v_0$ (eV)	-0.027	-0.089	-0.120	-0.172	

On note qu'il y a presque un facteur 10 entre ε_{eq} et v_0 : rapportée à un constituant, l'énergie de cohésion du solide CFC X est beaucoup plus grande que l'énergie d'une molécule X_2 .

¹³L'énergie de vibration contient deux termes (cinétique et potentiel), tous deux positifs.

8.2.3 Cohésion des réseaux ioniques

On considère à nouveau le cas le plus simple : un ensemble d'ions fixés aux nœuds d'un réseau et en prenant comme fragments les ions isolés. Il n'y a toujours pas d'énergie cinétique (les ions sont fixés) et l'interaction dominante à une distance $\gtrsim a_0$ est maintenant celle de Coulomb, qui décroît très lentement comme R^{-1} . Il existe toujours une répulsion à courte distance, de sorte qu'il faut écrire (dans la limite macroscopique) :

$$E = \frac{N}{2} \sum_{\vec{R}} [V_{\text{cœurs}}(\vec{R}) + V_{\text{Coulomb}}(\vec{R})] \equiv E_{\text{cœurs}} + E_{\text{Coulomb}} . \quad (8.22)$$

En raison de la décroissance lente de l'interaction de Coulomb, le calcul de la somme impliquant V_{Coulomb} n'est pas aussi simple qu'il pourrait paraître à première vue. La difficulté est visible immédiatement pour un réseau à une dimension, où alternent des ions positifs et des ions négatifs, portant chacun une charge élémentaire et situés à la distance a les uns des autres. L'énergie potentielle de l'ion situé à l'origine avec tous les autres est :

$$2 \frac{e'^2}{a} \left[-1 + \frac{1}{2} - \frac{1}{3} + \dots \right] . \quad (8.23)$$

Le facteur 2 global prend en compte les deux moitiés du réseau, à gauche et droite de l'origine. La série harmonique alternée est convergente :

$$\sum_{n=1}^{+\infty} \frac{(-1)^n}{n} = -\ln 2 , \quad (8.24)$$

mais converge avec une lenteur désespérante. D'ailleurs, on pourrait décider de compter autrement, en prenant d'abord l'interaction entre l'ion à l'origine avec tous ses partenaires de même signe ; ainsi viendrait la première série :

$$2 \frac{e'^2}{a} \sum_{n=1}^{+\infty} \frac{1}{2n} . \quad (8.25)$$

L'interaction entre le même ion central et ceux de charges opposées est :

$$-2 \frac{e'^2}{a} \sum_{n=1}^{+\infty} \frac{1}{2n+1} . \quad (8.26)$$

Les deux séries (8.25) et (8.26) sont ... divergentes !

Le même type de difficulté se produit dans le cas tridimensionnel ; la série est mal conditionnée et il faut en fait construire le cristal de façon astucieuse, afin d'obtenir rapidement la bonne réponse. Par exemple, le cristal de sel, NaCl, est un cubique à faces centrées : deux réseaux CFC de même côté (paramètre) a , mais décalés de $\frac{a}{2}$ le long d'un côté du cube. La distance entre l'ion à l'origine (Na^+ , par exemple) et ses homologues situés en \vec{R} est de la forme $\rho(\vec{R})a$, où ρ est la fonction vectorielle donnant, en unités a , la distance entre deux ions Na^+ . L'énergie électrostatique prend alors la forme :

$$E_{\text{Coulomb}} = \frac{1}{2} \frac{Ne'^2}{a} \left[- \sum_{\forall \vec{R}} \frac{1}{\rho(\vec{R} + \vec{d})} + \sum_{\forall \vec{R} \neq \vec{0}} \frac{1}{\rho(\vec{R})} \right] , \quad (8.27)$$

où \vec{d} est la translation d'un demi-côté du cube. L'énergie par paire d'ions¹⁴ est donc :

$$\varepsilon_{\text{Coulomb}} = \frac{e'^2}{a} \left[-\frac{1}{\rho(\vec{d})} - \sum_{\forall \vec{R} \neq \vec{0}} \left(\frac{1}{\rho(\vec{R} + \vec{d})} - \frac{1}{\rho(\vec{R})} \right) \right] . \quad (8.28)$$

Compte tenu de la décroissance lente de l'interaction de Coulomb, la série apparaissant dans (8.28) est *conditionnellement* convergente : suivant les regroupements de termes effectués, on peut lui donner à peu près n'importe quelle valeur !

¹⁴S'agissant des cristaux ioniques, l'usage est de compter l'énergie par paire d'ions, plutôt que par ion.

La difficulté peut aussi se sentir qualitativement en invoquant le théorème de Gauss, à propos d'une distribution de charge à symétrie sphérique, de rayon R . On sait que le champ ressenti par une charge-test située à la distance $r > R$ est le même que si toute la charge était concentrée au centre. Autrement dit, quand on ajoute des charges en surface, l'effet global est le même, indépendant de la distance R de ces charges aux charges situées vers le centre de la distribution. Ainsi, une distribution de charge, aussi étendue qu'elle soit, est sensible à ce qui se passe en surface.

Pour contourner la difficulté, on peut construire mentalement le cristal en le complétant à chaque fois par des fragments de charge totale nulle. Clairement, l'adjonction d'un tel fragment donnera des contributions positives et négatives à l'énergie d'interaction et, la distance aidant, les termes supplémentaires seront d'autant plus petits que la taille du système est grande. Dans ces conditions, la série devient (plus ou moins) rapidement convergente.

Dans le cas d'un cristal à symétrie cubique, il est possible de fournir un argument plus quantitatif permettant de comprendre ce qui se passe, et par là de construire des méthodes introduisant des séries rapidement convergentes. Le développement multipolaire s'obtient en faisant un développement de Taylor, ce qui peut s'écrire en utilisant la relation formelle (employée en Mécanique quantique à plusieurs reprises) :

$$f(\vec{r} + \vec{a}) = e^{\vec{a} \cdot \vec{\nabla}} f(\vec{r}) = [1 + \vec{a} \cdot \vec{\nabla} + \frac{1}{2!} (\vec{a} \cdot \vec{\nabla})^2 + \dots] f(\vec{r}) ; \quad (8.29)$$

ainsi, en développant autour du point \vec{r} ($\vec{a} = -\vec{r}'$), il vient :

$$\frac{1}{\|\vec{r} - \vec{r}'\|} = \frac{1}{r} - (\vec{r}' \cdot \vec{\nabla}) \frac{1}{r} + \frac{1}{2} (\vec{r}' \cdot \vec{\nabla})^2 \frac{1}{r} + \dots = \frac{1}{r} + \frac{\vec{r} \cdot \vec{r}'}{r^3} + \frac{3(\vec{r} \cdot \vec{r}')^2 - \vec{r}^2 \vec{r}'^2}{r^5} + \frac{1}{r} \mathcal{O} \left(\frac{r'}{r} \right)^3 . \quad (8.30)$$

Avec :

$$\vec{r}' \cdot \vec{\nabla} \frac{1}{r} = \vec{r}' \cdot \frac{\vec{r}}{r^3} , \quad (\vec{r}' \cdot \vec{\nabla})^2 = \frac{3(\vec{r} \cdot \vec{r}')^2 - \vec{r}^2 \vec{r}'^2}{r^5} , \quad (8.31)$$

l'interaction entre le cristal déjà constitué de densité $\rho(\vec{r})$, et celle du fragment ajouté de densité $\rho_{\text{fr}}(\vec{r}')$ est :

$$\begin{aligned} \frac{1}{2} \int d^3 r \int_{\mathbb{R}^3} d^3 r' \frac{\rho(\vec{r}) \rho_{\text{fr}}(\vec{r}')}{4\pi\epsilon_0 \|\vec{r} - \vec{r}'\|} d^3 r' = \\ \frac{1}{2} \int d^3 r \frac{\rho(\vec{r})}{4\pi\epsilon_0} \left[\int d^3 r' \frac{1}{r} \rho_{\text{fr}}(\vec{r}') + \int d^3 r' \frac{\vec{r} \cdot \vec{r}'}{r^3} \rho_{\text{fr}}(\vec{r}') + \int d^3 r' \frac{3(\vec{r} \cdot \vec{r}')^2 - \vec{r}^2 \vec{r}'^2}{r^5} \rho_{\text{fr}}(\vec{r}') + \dots \right] . \end{aligned} \quad (8.32)$$

Quand le fragment ajouté est de charge totale nulle, le premier terme du crochet est nul. Le second l'est tout autant si $\rho_{\text{fr}}(\vec{r}')$ est pair dans l'inversion d'espace, c'est-à-dire si la densité du morceau ajouté est invariante dans le changement $\vec{r}' \rightarrow -\vec{r}'$. Enfin, si le fragment supplémentaire a la symétrie cubique du cristal, le terme en r^{-5} est aussi nul (rapporté au vecteur unitaire \vec{r}/r , il est en r^{-3}), puisque la symétrie cubique assure que $3 \cos^2 \theta - 1$ est nul. Le terme suivant (non-écrit) – en r^{-4} suivant la même convention – est encore nul par parité. Finalement le premier terme non nul est en r^{-5} : clairement, la série devient rapidement convergente¹⁵.

D'autres méthodes de sommation existent et donnent d'une façon ou d'une autre l'énergie de Coulomb ; pour un réseau cubique de côté a , elle est nécessairement de la forme :

$$E_{\text{Coulomb}} \propto -N \frac{e'^2}{a} . \quad (8.33)$$

En comptant par paire d'ions (l'habitude pour les cristaux ioniques), on définit précisément une constante A , dite *constante de Madelung*, par la relation :

$$E_{\text{Coulomb}} = -A \frac{N}{2} \frac{e'^2}{a} ; \quad (8.34)$$

¹⁵À la réflexion, l'argumentation peut paraître un peu spéculaire, car la Nature se moque de savoir comment le Physicien s'y prend pour faire ses calculs. Et, en définitive, comment peut-on être sûr que le calcul, fait d'une façon ou d'une autre, donne le bon résultat ? En réalité, il faut bien comprendre ceci : quoiqu'il en soit, les cristaux sont toujours de taille *finie*, contenant N ions. Alors, quel que soit N , aussi gigantesque que l'on veut, le résultat est unique, indépendant de la procédure de sommation utilisée. Finalement, s'agissant de calculer la somme d'un grand nombre de termes, c'est le souci d'une bonne efficacité numérique qui dicte le choix d'une méthode ou d'une autre. Physiquement, la réponse est satisfaisante si, une fois la somme calculée par la méthode la plus rapide, l'énergie devient – à l'approximation souhaitée – une grandeur extensive.

A est un simple facteur numérique. En définitive, l'enjeu des différentes méthodes de sommation se réduit au calcul de cette constante. Par paire d'ions, (8.34) donne :

$$\varepsilon_{\text{Coulomb}} = -A \frac{e'^2}{a} . \quad (8.35)$$

La constante de Madelung ne dépend que de la structure cristalline et peut être calculée numériquement ; par exemple :

Structure cristalline	Constante de Madelung, A	
CsCl	1.7627	(8.36)
NaCl	1.7476	
Pechblende	1.6381	

Ceci est une borne supérieure de l'énergie de liaison, puisque la répulsion de cœur (déstabilisante) n'est pas prise en compte. En adoptant une loi-puissance pour cette dernière, l'énergie de cohésion dans le modèle à ions fixes est :

$$E = \frac{N}{2} \varepsilon \quad \varepsilon = -A \frac{e'^2}{a} + \frac{C}{a^m} \quad (N \gg 1) . \quad (8.37)$$

La distance d'équilibre s'obtient par minimisation :

$$A \frac{e'^2}{a^2} - m \frac{C}{a^{m+1}} = 0 \quad \Leftrightarrow \quad a_{\text{eq}} = \left(\frac{mC}{Ae'^2} \right)^{1/(m-1)} . \quad (8.38)$$

On en tire l'énergie par paire d'ions à l'équilibre en utilisant (8.38) au passage :

$$\varepsilon_{\text{eq}} = -A \frac{e'^2}{a_{\text{eq}}} + \frac{C}{a_{\text{eq}}^m} = -A \frac{e'^2}{a_{\text{eq}}} + \frac{A}{m} \frac{e'^2}{a_{\text{eq}}} = -A \frac{m-1}{m} \frac{e'^2}{a_{\text{eq}}} . \quad (8.39)$$

Pour m assez grand, ce dernier résultat est très peu différent de $-A(e'^2/a_{\text{eq}})$; a_{eq} dépend faiblement de m : finalement, pour les cristaux ioniques, le cœur fortement répulsif joue peu de rôle. En raison de son intensité et de sa longue portée, l'interaction de Coulomb est, dans ces systèmes, de loin l'interaction dominante.

8.2.4 Cohésion des solides covalents

Avec les solides covalents (exemple : la variété diamant du carbone), on rentre dans le domaine des solides dont les constituants, une fois rassemblés, sont grandement perturbés par rapport à leur état libre isolé (il en va de même pour les métaux). La liaison est maintenant du type liaison chimique, ce qui constitue une théorie en soi, débordant largement le cadre de ce cours. Il n'existe donc pas de théorie simple bâtie sur des objets libres ; il reste toutefois possible de donner quelques idées qualitatives.

Il est facile de comprendre pourquoi le carbone, par exemple, ne relève pas du tout du même traitement qu'un gaz rare. Dans un gaz rare, toutes les couches atomiques sont saturées et l'addition d'un électron demanderait une énergie importante destinée à peupler un état initialement vide : le principe de Pauli joue à fond son rôle. En revanche, le carbone, avec ses 6 électrons, a la configuration $(1s)^2 (2s)^2 (2p)^2$: la sous-couche (2p) est incomplète (elle peut recevoir 6 électrons), il y a donc beaucoup de place disponible, la case (2p) est de ce fait très malléable. La liaison chimique s'établit grâce au partage des électrons entre plusieurs atomes de carbone, cette mise en commun étant rendue possible par la place disponible chez l'un des électrons de l'autre. Ceci produit une délocalisation des électrons partagés et le calcul montre que cette délocalisation produit un abaissement de l'énergie – très schématiquement, quand Δx augmente, Δp_x diminue – et la formation d'une liaison se trouve donc favorisée¹⁶. Une caractéristique importante de la liaison chimique est son aspect directionnel, qui se traduit par une forte densité électronique le long de l'axe joignant les deux membres de la paire d'atomes engagés dans la liaison. Physiquement, c'est cette directivité qui est en grande partie responsable de la "rigidité" des molécules.

¹⁶C'est ce que l'on appelle en Chimie la stabilisation par *résonance* (exemple : les formules résonnantes du benzène).

8.2.5 Cohésion des métaux

En première approximation, un métal peut se représenter comme un réseau d'ions tous positifs noyés dans une mer d'électrons presque libres, mais néanmoins soumis au potentiel périodique du réseau. Ce sont ces électrons, libérés lors de la formation du cristal (dont les fragments sont neutres au départ), qui constituent la colle assurant la cohésion du cristal – autrement, tous les ions positifs se repoussant, l'édifice ne pourrait être stable et exploserait. Dans la théorie des métaux, il est d'usage d'introduire le rayon de la sphère “occupée” par un électron, noté r_s , défini comme¹⁷ :

$$V = N \frac{4\pi r_s^3}{3} , \quad (8.41)$$

où N est le nombre total d'électrons et V le volume de l'échantillon ; pour les métaux ordinaires, r_s vaut quelques a_0 .

L'énergie de cohésion se définit par référence à l'énergie des constituants séparés. Dans le cas des métaux, des électrons sont libérés lors de la formation du solide et se retrouvent presque libres (dans le cas d'un cristal ionique, chaque électron libéré par l'atome d'une espèce est piégé par un atome de l'autre). L'énergie du cristal constitué doit donc incorporer l'énergie d'interaction entre les ions (comme avant), mais également toutes les énergies relatives aux électrons (énergie potentielle et énergie cinétique).

L'énergie potentielle est celle de l'interaction de Coulomb ; elle donne lieu à plusieurs sortes de contributions :

- les interactions entre ions, répulsives
- les interactions entre électrons, également répulsives
- les interactions entre les ions et les électrons, attractives.

Dans une approche élémentaire, on oublie les vibrations possibles des ions positifs ; ceux-ci se bornent alors à jouer le rôle de charges ponctuelles fixes immergées dans un fond continu négatif compensateur (les électrons), le système étant globalement neutre. Le calcul peut se faire suivant les mêmes techniques que pour les réseaux ioniques, et on trouve, par atome, dans le cas d'un réseau cubique centré :

$$\varepsilon_{\text{Coulomb}} = - \frac{24.35}{r_s/a_0} \text{ eV} . \quad (8.42)$$

Le résultat (8.42) est purement classique et incorpore toutes les interactions coulombiennes *classiques* de la distribution de charge (électrons et ions) ainsi définie (réseau d'ions fixes, de structure déterminée et distributions électronique uniforme et continue).

En fait, le calcul du terme de répulsion entre électrons ne se réduit pas au terme classique inclus dans (8.42) qui, par nature, ignore le Principe de Pauli. Dans l'approximation minimale prenant ce dernier en compte pour des fermions (une fonction d'onde en forme de déterminant de Slater), on sait de l'étude de l'atome d'hélium, qu'apparaissent en plus des termes dits d'échange donnant au total une contribution toujours négative pour deux électrons de même spin (trou de Fermi). Ceci réduit fortement la répulsion de Coulomb (qui est déstabilisante) par rapport à la situation “spinless”¹⁸ ; au total, la prise en compte de ces termes abaisse la moyenne de l'énergie potentielle totale de façon importante. Pour la contribution d'échange, on trouve très précisément¹⁹ :

$$\varepsilon_{\text{échange}} = - \frac{3}{2\pi} \left(\frac{9\pi}{4} \right)^{1/3} \frac{e'^2}{2a_0} \simeq - \frac{12.4}{r_s/a_0} \text{ eV} , \quad (8.43)$$

¹⁷La définition de r_s fluctue d'un auteur à l'autre ; l'autre définition usuelle donne un r_s adimensionné (on compte le rayon de la sphère en unités a_0) :

$$V = N \frac{4\pi r_s^3}{3} a_0^3 . \quad (8.40)$$

¹⁸En effet, pour la moitié des paires d'électrons, la répulsion est fortement diminuée par le trou de Fermi.

¹⁹Ce calcul n'est pas difficile, mais n'est pas pour autant élémentaire. Il passe par le calcul de l'énergie à deux corps dans une approximation de type déterminant de Slater ou, ce qui ici revient au même comme on peut s'en convaincre, un calcul de perturbation au premier ordre.

de sorte que la prise en compte de l'indiscernabilité des électrons donne au total l'énergie de Coulomb :

$$\varepsilon_{\text{Coulomb}} \simeq -\frac{36.8}{r_s/a_0} \text{ eV} . \quad (8.44)$$

Ceci étant, à cause de la faible inertie des électrons et de leur mouvement presque libre au sein du métal, il faut prendre en compte leur énergie cinétique, qui entre bien dans le bilan d'énergie par rapport aux constituants (atomes neutres) séparés. En raison de la rusticité du modèle en cours d'analyse, on va calculer cette énergie cinétique comme si les électrons étaient complètement libres – ce qui suppose que l'énergie potentielle de chaque électron dans le champ des ions positifs est une petite correction²⁰. L'énergie de l'un d'entre eux est alors :

$$\varepsilon(\vec{k}) = \frac{\hbar^2 k^2}{2m} , \quad (8.45)$$

où \vec{k} est le vecteur d'onde. Dans le cas d'un alcalin par exemple, il y a un électron "libre" par maille, c'est-à-dire que s'il y a N ions dans le cristal il y a aussi N électrons, qu'il faut répartir dans les états à une particule conformément au Principe de Pauli. Les électrons étant considérés comme libres, ils sont représentés par des ondes planes $\propto e^{i\vec{k}\cdot\vec{r}}$; en adoptant des conditions cycliques de Born - von Karman dans une boîte cubique de côté L , les valeurs de \vec{k} sont quantifiées, chaque composante étant un multiple entier de $\frac{2\pi}{L}$: l'élément de volume dans l'espace réciproque est donc un petit cube "infinitésimal" de côté $\frac{2\pi}{L}$.

Le principe de Pauli exige de répartir les électrons dans ces états mettant au plus 2 électrons par état, compte tenu du spin. En procédant ainsi, on remplit ceux-ci jusqu'à un certain vecteur d'onde k_F , appelé vecteur d'onde de Fermi. Dans la sphère de rayon k_F , il y a :

$$\frac{\frac{4\pi}{3} k_F^3}{\left(\frac{2\pi}{L}\right)^3} \quad (8.46)$$

petits cubes et chacun peut accepter deux électrons. Pour un alcalin, le nombre d'électrons "libres" est égal au nombre d'atomes, soit N ; au total, on a :

$$2 \frac{\frac{4\pi}{3} k_F^3}{\frac{2\pi^3}{L^3}} = N \iff k_F = \frac{2\pi}{L} \left(\frac{3N}{8\pi}\right)^{1/3} \propto \left(\frac{N}{V}\right)^{1/3} . \quad (8.47)$$

Maintenant, l'énergie cinétique de ces N électrons est :

$$E_{\text{cin}} = \sum_{\text{spin}} \sum_{k \leq k_F} \frac{\hbar^2 k^2}{2m} \rightarrow 2 \left(\frac{L}{2\pi}\right)^3 \int_0^{k_F} 4\pi k^2 dk \frac{\hbar^2 k^2}{2m} . \quad (8.48)$$

Une intégration élémentaire fournit :

$$E_{\text{cin}} = N \frac{3}{5} \frac{\hbar^2 k_F^2}{2m} \equiv N \frac{3}{5} \varepsilon_F , \quad (8.49)$$

la dernière égalité définissant l'énergie de Fermi ε_F pour le gaz d'électrons libres. En introduisant la longueur r_s , on trouve finalement l'énergie cinétique par électron sous la forme :

$$\varepsilon_{\text{cin}} = \frac{3}{5} \frac{e'^2}{2a_0} \left(\frac{9\pi}{4}\right)^{2/3} \left(\frac{a_0}{r_s}\right)^2 \iff \varepsilon_{\text{cin}} \simeq \frac{30.1}{(r_s/a_0)^2} \text{ eV} . \quad (8.50)$$

Notons que comme k_F augmente avec la densité N/V , il en va de même de E_{cin} .

En additionnant maintenant toutes les contributions à l'énergie (8.42), (8.50) et (8.43), on obtient :

$$\varepsilon = \frac{30.1}{(r_s/a_0)^2} - \frac{36.8}{r_s/a_0} \text{ eV} . \quad (8.51)$$

²⁰Cette idée est le fondement de l'approximation dite des *électrons presque libres*, qui sera exposée ultérieurement. Notons toutefois que même si l'effet du réseau est quantitativement petit, il provoque néanmoins une modification qualitative spectaculaire de la loi de dispersion de l'électron : la parabole libre devient une succession d'arcs déformés et disjoints, d'où l'apparition de *gaps* d'énergie.

La minimisation par rapport à r_s donne :

$$\left(\frac{r_s}{a_0}\right)_{\text{eq}} \simeq 1.6 . \quad (8.52)$$

L'énergie de cohésion à l'équilibre s'en déduit²¹ :

$$|\varepsilon(r_{s,\text{eq}})| \simeq 11.2 \text{ eV} . \quad (8.53)$$

Ces résultats sont médiocres. Notamment, les valeurs expérimentales de $\frac{r_s}{a_0}$ varient en fait entre 2 et 6 ; par ailleurs, on trouve évidemment la même valeur d'équilibre pour tous les alcalins²², indépendamment de leur masse atomique, puisque les ions ont été pris ponctuels. Les raisons de ces insuccès sont de plusieurs ordres.

D'abord, comme les ions ne sont pas ponctuels, les électrons de conduction ne peuvent pas en fait s'approcher aussi près des ions qu'on l'a supposé : il existe une sorte de volume exclu. Il en résulte d'une part que l'interaction électrostatique élection - ion n'est pas aussi grande en valeur absolue et que d'autre part, là où elle prend des valeurs notables, la densité électronique est en réalité plus élevée, ce qui produit une énergie cinétique plus grande. En conséquence, le terme $\varepsilon_{\text{Coulomb}}$ est moins négatif et l'énergie ε_{cin} est plus élevée. Ces deux tendances se conjuguent pour donner une valeur d'équilibre plus grande que ne le prévoit la formule simpliste (8.51). En outre, en prenant de fait en compte la variation du rayon ionique avec la masse, un calcul excluant les cœurs peut de fait prédire que les différents alcalins ont bien des distances d'équilibre distinctes.

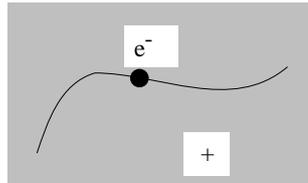


Figure 8.3: Représentation schématique du *jellium*.

Pour les métaux, on dispose aussi d'un autre modèle élémentaire donnant de meilleurs résultats pour r_s – en améliorant le résultat numérique pour les termes de Coulomb – tout en continuant à donner, par construction, une valeur indépendante de l'alcalin considéré. Il s'agit du *jellium*, qui reprend l'essentiel du modèle précédent mais remplace les ions ponctuels par un fond continu positif, d'où son nom. Le calcul de l'énergie cinétique des électrons est évidemment inchangé, et celle-ci est donc toujours donnée par (8.50). En ce qui concerne l'énergie de Coulomb, elle contient toujours des termes “classiques” et des termes d'échange venant du Principe de Pauli. Le calcul donne²³ :

$$\varepsilon_{\text{Coulomb}} = -\frac{3}{2\pi} \left(\frac{9\pi}{4}\right)^{1/3} \frac{e'^2}{2a_0} \simeq -\frac{12.4}{r_s/a_0} \text{ eV} . \quad (8.54)$$

Au total, l'énergie par électron est :

$$\varepsilon(r_s) = \frac{30.1}{(r_s/a_0)^2} - \frac{12.4}{r_s/a_0} \text{ eV} . \quad (8.55)$$

Comme le terme attractif est plus petit que précédemment en valeur absolue, la valeur d'équilibre de r_s s'en trouve augmentée ; elle vaut maintenant :

$$\left(\frac{r_s}{a_0}\right)_{\text{eq}} \simeq 4.8 , \quad (8.56)$$

²¹ $\varepsilon(r_{s,\text{eq}})$ est évidemment négatif.

²²Rappelons que le calcul a été fait pour N atomes (hypothèse nécessaire pour arriver à (8.42)), et que l'on a considéré un nombre égal d'électrons pour obtenir (8.49). Le calcul ne vaut donc bien que pour les éléments de la première colonne, soit Li, Na, K, Rb, Cs.

²³On note, par comparaison avec (8.43), que la contribution totale des interactions coulombiennes est maintenant égale au terme d'échange ; en effet, le calcul détaillé montre que, dans le cas du *jellium*, les termes directs (positifs) compensent très exactement la contribution (qui se trouve être négative) liée à l'existence du fond continu positif (*self*-répulsion de ce fond et attraction entre les électrons et le fond).

d'où l'énergie de cohésion à l'équilibre :

$$|\varepsilon(r_{s,\text{eq}})| \simeq 1.3 \text{ eV} . \quad (8.57)$$

La valeur de $r_{s,\text{eq}}$ est nettement plus satisfaisante, mais ne doit pas faire illusion et faire oublier le caractère un peu artificiel du *jellium*.

Clairement, c'est du côté des hautes densités électroniques que le *jellium* trouve sa pleine validité ; en effet, lorsque r_s est petit, c'est l'énergie cinétique ε_{cin} qui est dominante et la prise en compte élémentaire (*i.e.* par perturbations) de l'interaction coulombienne concerne alors des termes énergétiques relativement petits. À l'autre extrême (faibles densités, grand r_s), le calcul précédent tend à montrer (malgré ses limites) que l'énergie potentielle est dominante, l'énergie cinétique est petite (les électrons sont peu mobiles). En 1938, par de tout autres moyens, Wigner a prévu l'existence d'un cristal bidimensionnel triangulaire d'électrons. Cette prévision théorique a reçu une confirmation expérimentale dans les années '80, par la mesure de la modification des ondes capillaires de l'hélium liquide, à la surface duquel un gaz d'électrons avait été déposé. Les altérations observées sont compatibles avec le réseau triangulaire proposé par Wigner.

D'une façon générale, la théorie de perturbation ordinaire pour traiter l'interaction de Coulomb entre électrons n'est pas satisfaisante : le calcul à des "ordres supérieurs" fait apparaître des termes singuliers du genre $r_s \ln r_s$.

8.3 Structures cristallines

8.3.1 Réseau de Bravais

Soit, dans \mathbb{R}^3 , trois vecteurs linéairement indépendants \vec{a}_1 , \vec{a}_2 , et \vec{a}_3 . On appelle réseau de Bravais l'ensemble des points dont le rayon-vecteur \vec{R} est de la forme :

$$\vec{R} = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3 \quad (n_i \in \mathbb{Z}) . \quad (8.58)$$

À l'ensemble des vecteurs \vec{R} , on peut associer un ensemble de translations $\mathcal{T}(\vec{R})$, appelé translations du réseau ; toute translation du réseau relie deux points du réseau. On peut définir de la même façon un réseau uni- ou bi-dimensionnel ($d=1$ et 2 respectivement) ; généralement, un réseau en dimension d peut être noté \mathbb{Z}^d . Un réseau de Bravais est infini dans toutes ses directions et offre exactement la même vision de quelque point qu'on le regarde (de chaque point, on voit exactement le même paysage que de tous les autres) ; chaque point du réseau est aussi appelé nœud. Les vecteurs \vec{a}_i sont appelés vecteurs primitifs, puisqu'ils permettent d'engendrer le réseau par la relation de définition (8.3.1) ; leur choix n'est pas unique en général, voir l'exemple donné sur la figure 8.4, dans le cas d'un réseau \mathbb{Z}^2 .

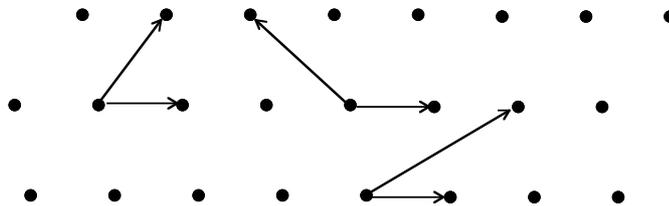


Figure 8.4: Différents choix possibles pour les vecteurs primitifs \vec{a}_i d'un réseau bidimensionnel.

Le mot réseau désigne plus généralement un ensemble infini de points régulièrement répartis dans l'espace suivant une règle précise et non-ambiguë. Tous les réseaux ne sont pas des réseaux de Bravais, l'exemple le plus connu étant le réseau en nid d'abeille (figure 8.5). En effet, si on considère les deux vecteurs \vec{u} et \vec{v} , on peut bien engendrer tout le réseau, mais aussi *en plus* des points (par exemple les centres des hexagones) qui n'appartiennent pas au réseau initialement défini ; il est facile de se convaincre géométriquement qu'il n'existe pas deux vecteurs permettant de construire tous les points du réseau nid d'abeille et eux seulement : le réseau en nid d'abeille n'est pas un réseau de Bravais. Le réseau de Bravais engendré à partir de \vec{u} et \vec{v} (donc avec un nœud au centre de chaque hexagone) s'appelle réseau hexagonal.

L'analyse systématique montre qu'il existe en tout et pour tout 14 réseaux de Bravais distincts (cubique, quadratique, orthorhombique, monoclinique, triclinique), chacun pouvant exister sous différentes formes (par exemple, pour le cubique : simple, centré et à faces centrées).

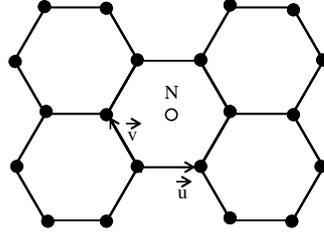


Figure 8.5: Le réseau en nid d'abeille n'est pas un réseau de Bravais : le point N obtenu par la somme $\vec{u} + \vec{v}$ ne lui appartient pas.

Évidemment, outre la symétrie de translation, un réseau peut posséder d'autres éléments de symétrie (rotation, réflexion dans un plan, inversion d'espace, *etc.*), qui se distinguent des translations en ce sens qu'elles possèdent des points fixes (invariants). L'ensemble de ces opérations constitue la symétrie dite ponctuelle²⁴. Un point mérite d'être mentionné : la symétrie d'ordre 5 (une rotation de $\frac{2\pi}{5}$ par exemple) est incompatible avec la symétrie de translation d'un réseau de Bravais. On a toutefois découvert récemment (1984) des matériaux remarquables en ce sens qu'ils exhibent une symétrie pentagonale, tout en donnant des clichés diffraction à l'instar d'un réseau de Bravais. Ces systèmes, appelés quasi-cristaux, semblent donc à première vue invalider le dogme de la cristallographie interdisant la symétrie d'ordre 5. On reviendra sur ce point dans la suite.

Les réseaux cubiques sont des réseaux de Bravais. Le plus simple d'entre eux est le cubique simple (C) – un nœud en chacun des huit sommets –, de côté a , engendré par trois vecteurs \vec{a}_i de même longueur a et orthogonaux entre eux ; si $(\vec{i}, \vec{j}, \vec{k})$ désigne une base orthonormalisée de \mathbb{R}^3 , on peut prendre :

$$\vec{a}_1 = a\vec{i} \quad \vec{a}_2 = a\vec{j} \quad \vec{a}_3 = a\vec{k} . \quad (8.59)$$

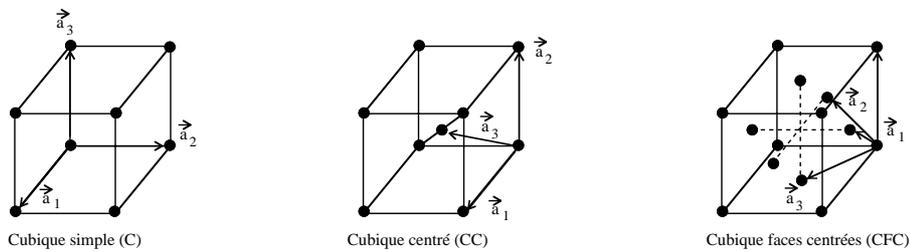


Figure 8.6: Réseaux cubique simple (C), cubique centré (CC) et cubique faces centrées (CFC), avec un choix de vecteurs primitifs.

Un autre réseau cubique est le cubique centré (CC) : il y a un nœud aux huit sommets d'un cube et un autre au centre du cube. Un choix de vecteurs primitifs est :

$$\vec{a}_1 = a\vec{i} \quad \vec{a}_2 = a\vec{j} \quad \vec{a}_3 = \frac{a}{2}(\vec{i} + \vec{j} + \vec{k}) . \quad (8.60)$$

Un autre choix plus symétrique est :

$$\vec{a}_1 = \frac{a}{2}(-\vec{i} + \vec{j} + \vec{k}) \quad \vec{a}_2 = \frac{a}{2}(\vec{i} - \vec{j} + \vec{k}) \quad \vec{a}_3 = \frac{a}{2}(\vec{i} + \vec{j} - \vec{k}) . \quad (8.61)$$

Le dernier réseau cubique important est le cubique faces centrées (CFC), qui, outre les huit nœuds aux sommets d'un cube, a un nœud au centre de chacune des 6 faces de ce dernier ; on peut l'engendrer avec les vecteurs primitifs suivants :

$$\vec{a}_1 = a\vec{i} \quad \vec{a}_2 = \frac{a}{2}(\vec{i} + \vec{j}) \quad \vec{a}_3 = \frac{a}{2}(\vec{j} + \vec{k}) . \quad (8.62)$$

²⁴le qualificatif veut rappeler, précisément, qu'il existe des points fixes.

À nouveau, il existe des choix plus symétriques, par exemple :

$$\vec{a}_1 = \frac{a}{2}(\vec{i} + \vec{j}) \quad \vec{a}_2 = \frac{a}{2}(\vec{j} + \vec{k}) \quad \vec{a}_3 = \frac{a}{2}(\vec{i} + \vec{k}) . \quad (8.63)$$

Ces trois réseaux jouent un rôle éminent, car, dans les conditions ordinaires, de nombreux matériaux cristallisent dans le système cubique. Toutefois, le cubique simple est très rare, un exemple connu étant la phase dite α du Polonium.

Etant donné un point d'un réseau, le nombre de ses plus proches voisins, souvent noté z , est appelé coordination. Pour un cubique simple, $z = 6$ ($z = 2d$) ; de surcroît, on voit facilement que :

$$z_{CC} = 8 \quad z_{CFC} = 12 . \quad (8.64)$$

Dans ces réseaux, pour a donné, la matière est donc nettement plus dense que dans le cubique simple.

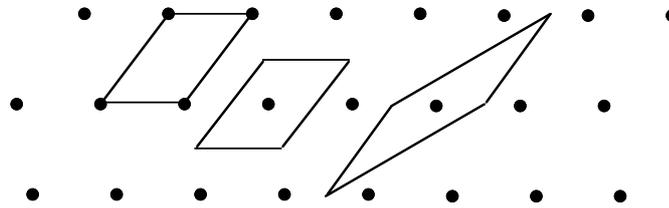


Figure 8.7: Différents choix possibles pour la maille primitive d'un réseau bidimensionnel.

On définit également la maille primitive d'un réseau de Bravais. C'est le volume (ou la surface si $d = 2$) qui, soumis à *toutes* les translations du réseau (construites sur les vecteurs primitifs) remplit complètement l'espace sans aucun recouvrement : c'est la brique élémentaire (incompressible !) pavant l'espace sans laisser le moindre trou. Un réseau étant donné, il y a un choix infini de mailles primitives (voir figure 8.7). Une maille primitive contient exactement un point du réseau et un seul ; tout point situé en surface est partagé entre les mailles adjacentes et compte donc pour une fraction de point – un demi-point s'il est sur la face commune à deux mailles. Soit un ensemble fini de N points construits sur trois vecteurs primitifs, dont le volume est V . La densité de points est donc $n = \frac{N}{V}$; si v désigne le volume d'une maille primitive, on a (aux effets de surface près) :

$$V \simeq N v = (nV) v \quad (N \gg 1) . \quad (8.65)$$

Dans la limite du cristal infini, les effets de surface s'annulent strictement et il vient, après simplification par V :

$$1 = n v \quad \Longleftrightarrow \quad v = \frac{1}{n} . \quad (8.66)$$

Ceci est vrai quelle que soit la maille primitive choisie : toutes les mailles primitives ont donc le même volume v , égal au module du produit mixte construit sur trois vecteurs primitifs :

$$v = |(\vec{a}_1, \vec{a}_2, \vec{a}_3)| . \quad (8.67)$$

On voit en outre que, deux mailles primitives étant choisies, l'une reproduit l'autre exactement quand on la découpe judicieusement en morceaux, auxquels on applique les bonnes translations du réseau.

Une fois définis les vecteurs primitifs, le choix le plus naturel pour la maille primitive est le volume dont tous les points ont pour rayon-vecteur \vec{r} tel que :

$$\vec{r} = x_1 \vec{a}_1 + x_2 \vec{a}_2 + x_3 \vec{a}_3 \quad (0 \leq x_i < 1) . \quad (8.68)$$

Il est d'usage courant, pour la commodité des représentations géométriques, de manipuler des mailles qui ne sont pas primitives au sens où on peut reproduire tout le réseau en leur appliquant seulement un *sous*-ensemble des translations. De telles mailles sont appelées mailles élémentaires conventionnelles. Par exemple, pour le cubique centré, une telle maille est constituée par le cube de côté a , munie de son atome central ; cette

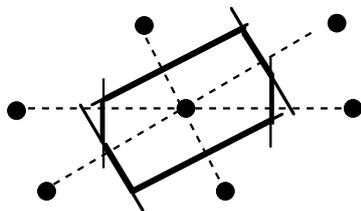


Figure 8.8: La maille primitive de Wigner-Seitz contient tous les points de l'espace plus proches d'un nœud donné que de tous les autres nœuds.

maille conventionnelle est deux fois plus grande en volume que toute maille primitive du CC puisqu'elle contient deux points du réseau²⁵.

Par ailleurs, il est toujours possible de choisir comme maille primitive une maille ayant par elle-même la symétrie (ponctuelle) du réseau. Un exemple est la maille dite²⁶ de Wigner - Seitz, constituée des points de \mathbb{R}^3 qui sont plus proches d'un nœud donné que de tous les autres nœuds. De façon évidente, il s'agit d'un polyèdre (ou d'un polygone pour $d = 2$) dont les faces (ou les côtés) sont dans les plans médiateurs (ou sur les médiatrices) des segments joignant deux voisins du réseau de Bravais (voir figure 8.8). Il s'agit bien d'une maille primitive, car elle restitue convenablement tout l'espace sans recouvrements quand on lui applique les translations du réseau. Par ailleurs, elle présente l'avantage d'être indépendante du choix des vecteurs primitifs (elle est construite directement sur les points du réseau) et c'est d'ailleurs pour cette raison qu'elle possède la même symétrie que ce réseau.

8.3.2 Structure physique d'un réseau de Bravais

Un réseau de Bravais est donc une construction purement géométrique dont les principales caractéristiques viennent d'être énoncées et constituée de points mathématiques. Pour définir un système physique construit de la sorte, il faut préciser la nature des objets situés aux nœuds de ce réseau. On appelle base l'unité physique (atome, molécule, ion, *etc.*) attachée à un nœud et dont la duplication à l'infini par translation reproduit le cristal physique, supposé parfait d'un point de vue structurel. Dans la figure 8.9, on donne l'exemple d'un matériau où une molécule diatomique hétéronucléaire occupe chaque nœud.

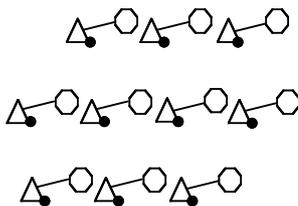


Figure 8.9: Réseau bidimensionnel avec pour base (motif) une molécule diatomique hétéronucléaire.

Un autre exemple intéressant (voir fig 8.10) montre comment un réseau en nid d'abeille – qui n'est pas un réseau de Bravais – peut être construit à partir d'un réseau triangulaire muni d'une base à deux points.

L'ensemble réseau et base définit complètement la structure cristalline.

²⁵Le même argument vaut pour le CFC. Si on prend une maille cubique de côté a , elle contient 6 points (ceux au centre de chaque face) partagés entre deux cubes et 8 points (ceux des sommets) partagés chacun entre 8 cubes. Le nombre de points est donc $6 \times \frac{1}{2} + 8 \times \frac{1}{8} = 4$ et le volume de cette maille conventionnelle est $4 v_{\text{CFC}}$.

²⁶Dans la littérature mathématique des réseaux de points, la maille de Wigner-Seitz s'appelle maille (ou cellule) de Voronoï. Par ailleurs, cette maille a été introduite en premier par Brillouin pour le réseau réciproque.

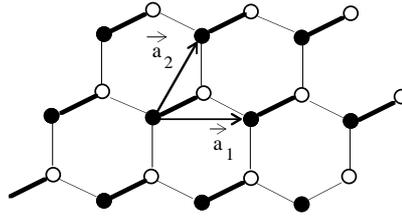


Figure 8.10: Réseau en nid d'abeille construit sur un réseau triangulaire avec deux atomes par nœud.

À l'inverse, pour mettre en avant la symétrie et/ou pour pouvoir raisonner géométriquement plus facilement, certains réseaux de Bravais authentiques peuvent être définis comme des réseaux plus "simples", mais munis d'une base non ponctuelle. Par exemple, le cubique centré peut être défini comme un cubique simple avec sa maille primitive cubique de côté a construite sur une base $(\vec{i}, \vec{j}, \vec{k})$ et muni de la base à deux points de coordonnées :

$$(0, 0, 0) \quad \left(\frac{a}{2}, \frac{a}{2}, \frac{a}{2}\right) . \quad (8.69)$$

De la même façon, le cubique faces centrées peut être construit sur un cubique simple avec une base à quatre points de coordonnées :

$$(0, 0, 0) \quad \left(\frac{a}{2}, \frac{a}{2}, 0\right) \quad \left(\frac{a}{2}, 0, \frac{a}{2}\right) \quad \left(0, \frac{a}{2}, \frac{a}{2}\right) . \quad (8.70)$$

8.4 Réseau réciproque

8.4.1 Définition et propriétés du réseau réciproque

Le réseau réciproque d'un réseau de Bravais apparaît spontanément dès que l'on fait de l'analyse de Fourier. En effet, un réseau de points étant donné, l'objet fondamental de l'analyse de Fourier est l'onde plane $e^{i\vec{k}\cdot\vec{r}}$ où \vec{k} est le vecteur d'onde. Par ailleurs, on connaît l'importance du fait que $e^{i\vec{k}\cdot\vec{r}} = 1$: c'est la condition pour avoir des interférences constructives. On en vient ainsi tout naturellement, un réseau de Bravais $\{\vec{R}\}$ étant donné, à considérer l'ensemble des points caractérisés par un vecteur d'onde, noté \vec{K} pour le distinguer, satisfaisant la relation :

$$e^{i\vec{K}\cdot\vec{R}} = 1 . \quad (8.71)$$

L'ensemble des vecteurs \vec{K} est plongé dans \mathbb{R}^3 , mais ce n'est pas l'espace réel – ses vecteurs sont homogènes à l'inverse d'une longueur (comme un vecteur d'onde) – et s'appelle espace réciproque. Il est clair que l'ensemble des points satisfaisant (8.71) ne forme pas un *continuum* mais a la nature discrète d'un réseau, appelé réseau réciproque. En définitive, le réseau réciproque est l'ensemble des points satisfaisant la relation de définition :

$$e^{i\vec{K}\cdot\vec{R}} = 1 \quad \forall \vec{R} \in \mathcal{B} \quad (8.72)$$

où \mathcal{B} désigne un réseau de Bravais. Le réseau réciproque de \mathcal{B} sera noté $\vec{\mathcal{B}}$. De la relation de définition (8.72), il résulte immédiatement :

$$\forall \vec{r} \in \mathbb{R}^3 : e^{i\vec{K}\cdot(\vec{R}+\vec{r})} = e^{i\vec{K}\cdot\vec{r}} . \quad (8.73)$$

Une fois introduit le réseau réciproque d'un réseau de Bravais, ce dernier est parfois appelé réseau direct, pour la symétrie du langage.

Il est facile de voir que le réseau réciproque est lui-même un réseau de Bravais. En effet, soit \vec{K}_1, \vec{K}_2 et \vec{K}_3 trois vecteurs linéairement indépendants satisfaisant (8.72). Formons maintenant le vecteur \vec{K} :

$$\vec{K} = m_1 \vec{K}_1 + m_2 \vec{K}_2 + m_3 \vec{K}_3 \quad (m_j \in \mathbb{Z}) . \quad (8.74)$$

Alors, $\forall \vec{R} \in \mathcal{B}$ et puisque les \vec{K}_j satisfont (8.72) et que les m_j sont entiers :

$$e^{i \vec{K} \cdot \vec{R}} = e^{i \sum_{j=1}^3 m_j \vec{K}_j \cdot \vec{R}} = \prod_{j=1}^3 e^{i m_j \vec{K}_j \cdot \vec{R}} = \prod_{j=1}^3 1^{m_j} = \prod_{j=1}^3 1 = 1 . \quad (8.75)$$

Le vecteur \vec{K} satisfait (8.72) lui aussi, c'est donc un vecteur de $\tilde{\mathcal{B}}$; sa construction avec des coordonnées entières selon (8.74) montre que $\tilde{\mathcal{B}}$ est un réseau de Bravais. Il s'agit sûrement de la preuve la plus simple et la plus directe de cette propriété fondamentale du réseau réciproque.

Une autre démonstration passe par un choix de vecteurs primitifs pour le réseau direct \mathcal{B} , supposé à trois dimensions. Ceux-ci étant donnés, on considère les trois vecteurs \vec{b}_i définis comme suit :

$$\vec{b}_1 = 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \quad \vec{b}_2 = 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \quad \vec{b}_3 = 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} . \quad (8.76)$$

On va montrer que les \vec{b}_i sont des vecteurs primitifs de $\tilde{\mathcal{B}}$. On note d'abord que $\vec{a}_1 \cdot \vec{b}_1 = 2\pi$ et que $\vec{a}_1 \cdot \vec{b}_2 = 0$, \dots , soit plus généralement :

$$\vec{a}_i \cdot \vec{b}_j = 2\pi \delta_{ij} . \quad (8.77)$$

Chaque \vec{b}_j est orthogonal aux deux \vec{a}_i qui rentrent dans sa définition (8.76), ce qui est évident par la définition du produit vectoriel. Soit maintenant d'une part un vecteur \vec{k} défini comme une combinaison linéaire des \vec{b}_j :

$$\vec{k} = \sum_{j=1}^3 k_j \vec{b}_j \quad (k_j \in \mathbb{R}) , \quad (8.78)$$

d'autre part un vecteur $\vec{R} \in \mathcal{B}$ d'un réseau de Bravais (direct) :

$$\vec{R} = \sum_{i=1}^3 n_i \vec{a}_i \quad (n_i \in \mathbb{Z}) . \quad (8.79)$$

Alors, compte tenu de (8.77) :

$$\vec{k} \cdot \vec{R} = \sum_{i=1}^3 \sum_{j=1}^3 n_i k_j \vec{a}_i \cdot \vec{b}_j = 2\pi \sum_{i=1}^3 n_i k_i \quad \forall \vec{R} \in \mathcal{B} . \quad (8.80)$$

Prenons maintenant $k_j = m_j \in \mathbb{Z}$; alors la somme $\sum_{j=1}^3 n_i k_j$ est elle-même un nombre entier. Il en résulte que tout vecteur \vec{k} de la forme $\sum_{j=1}^3 m_j \vec{b}_j$ est tel que le produit scalaire $\vec{k} \cdot \vec{R}$ est égal à $2\pi \times$ entier, $\forall \vec{R} \in \mathcal{B}$ ²⁷ ; autrement dit, ce vecteur est un vecteur du réseau réciproque $\tilde{\mathcal{B}}$, que l'on peut noter \vec{K} :

$$\vec{K} = \sum_{j=1}^3 m_j \vec{b}_j \quad (m_j \in \mathbb{Z}) \iff \vec{K} \in \tilde{\mathcal{B}} . \quad (8.81)$$

Ceci établit que les \vec{b}_j sont bien des vecteurs primitifs de $\tilde{\mathcal{B}}$.

Le produit mixte $\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)$ est, au signe près, le volume de la maille primitive de \mathcal{B} :

$$|\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)| = v . \quad (8.82)$$

Calculons le volume de la maille primitive de $\tilde{\mathcal{B}}$; on a :

$$\vec{b}_2 \times \vec{b}_3 = \frac{(2\pi)^2}{[\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)]^2} [(\vec{a}_3 \times \vec{a}_1) \times (\vec{a}_1 \times \vec{a}_2)] . \quad (8.83)$$

²⁷Le quantificateur $\forall \vec{R} \in \mathcal{B}$ est essentiel. Pour un choix donné des k_i (par exemple des rationnels), on pourrait trouver des n_i particuliers tels que $\sum_{j=1}^3 n_i k_j$ prend une valeur entière.

En vertu de la relation :

$$(\vec{U} \times \vec{V}) \times \vec{W} = (\vec{W} \cdot \vec{U}) \cdot \vec{V} - (\vec{V} \cdot \vec{W}) \cdot \vec{U} , \quad (8.84)$$

on a :

$$(\vec{a}_3 \times \vec{a}_1) \times (\vec{a}_1 \times \vec{a}_2) = [(\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_3] \cdot \vec{a}_1 - [\vec{a}_1 \cdot (\vec{a}_1 \times \vec{a}_2)] \cdot \vec{a}_3 . \quad (8.85)$$

Le dernier crochet est nul, d'où :

$$\vec{b}_2 \times \vec{b}_3 = \frac{(2\pi)^2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \vec{a}_1 . \quad (8.86)$$

Il en résulte :

$$\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3) = \frac{(2\pi)^2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \vec{b}_1 \cdot \vec{a}_1 = \frac{(2\pi)^3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} , \quad (8.87)$$

la dernière égalité résultant de (8.77). Le volume de $\tilde{\mathcal{B}}$, \tilde{v} , est au signe près le produit mixte $\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)$:

$$\tilde{v} = |\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)| . \quad (8.88)$$

v et \tilde{v} sont donc reliés par :

$$v \tilde{v} = (2\pi)^3 . \quad (8.89)$$

Le produit est bien un nombre pur puisque $[v] = [\text{longueur}^3]$ et que $[\tilde{v}] = [\text{longueur}]^{-3}$.

Le réseau réciproque étant un réseau de Bravais, il possède donc lui aussi un réseau réciproque— qui n'est autre que le réseau direct de départ :

$$\tilde{\tilde{\mathcal{B}}} = \mathcal{B} . \quad (8.90)$$

Ce résultat peut se démontrer de plusieurs façons. Une démonstration directe utilise la définition (8.76) et considère les trois vecteurs \vec{c}_k :

$$\vec{c}_1 = 2\pi \frac{\vec{b}_2 \times \vec{b}_3}{\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)} \quad \vec{c}_2 = 2\pi \frac{\vec{b}_3 \times \vec{b}_1}{\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)} \quad \vec{c}_3 = 2\pi \frac{\vec{b}_1 \times \vec{b}_2}{\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)} . \quad (8.91)$$

D'après (8.86) et (8.87), on a par exemple :

$$\vec{c}_1 = 2\pi \frac{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}{(2\pi)^3} \frac{(2\pi)^2}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \vec{a}_1 = \vec{a}_1 , \quad (8.92)$$

et de même $\vec{c}_2 = \vec{a}_2$, $\vec{c}_3 = \vec{a}_3$. L'identité $\vec{c}_i = \vec{a}_i$ est d'ailleurs évidente géométriquement.

On peut aussi raisonner plus formellement comme suit. Par définition, le réseau réciproque de $\tilde{\mathcal{B}}$ est l'ensemble de tous les vecteurs \vec{T} tels que :

$$e^{i\vec{K} \cdot \vec{T}} = 1 \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (8.93)$$

Par ailleurs, les vecteurs \vec{K} sont tels que :

$$e^{i\vec{K} \cdot \vec{R}} = 1 \quad \forall \vec{R} \in \mathcal{B} . \quad (8.94)$$

Il en résulte que tous les vecteurs \vec{R} appartiennent au réciproque du réciproque :

$$\{\vec{R}\} \subseteq \{\vec{T}\} . \quad (8.95)$$

En outre, soit $\vec{t} = \sum_{i=1}^3 x_i \vec{a}_i$ un vecteur n'appartenant pas à \mathcal{B} (l'un des réels x_i au moins n'est pas entier relatif, par exemple x_1). Prenant $\vec{K} = \vec{b}_1$ et $\vec{T} = \vec{t}$, il vient :

$$e^{i\vec{K} \cdot \vec{T}} = e^{i \sum_i x_i 2\pi \delta_{ij}} = e^{i x_1 2\pi} \neq 1 \quad (x_1 \notin \mathbb{Z}) . \quad (8.96)$$

Ainsi, un vecteur n'appartenant pas à \mathcal{B} n'appartient pas non plus à $\tilde{\tilde{\mathcal{B}}}$. Compte tenu de (8.93) et 8.96, on a donc :

$$\tilde{\tilde{\mathcal{B}}} = \mathcal{B} . \quad (8.97)$$

Toutes les définitions ont été données dans le cas d'un réseau \mathcal{B} à trois dimensions. Elles s'appliquent également en dimension inférieure (en pratique $d = 1$ ou 2). Par exemple, soit à construire le réseau réciproque d'un réseau bidimensionnel engendré par les deux vecteurs primitifs \vec{a}_1 et \vec{a}_2 , faisant entre eux l'angle θ . Rien n'interdit de définir formellement (et provisoirement) une troisième direction, de vecteur primitif \vec{a}_3 ; de surcroît, la prendre perpendiculaire au plan défini par \vec{a}_1 et \vec{a}_2 ne nuit aucunement à la généralité. L'astuce consiste à faire tendre en temps utile, la norme de \vec{a}_3 vers l'infini. Autrement dit, on introduit un réseau monoclinique dont la direction droite est en bout de course dilatée à l'infini.

Compte tenu de la définition (8.76) et du choix de \vec{a}_3 (perpendiculaire aux plans réticulaires), le vecteur \vec{b}_1 est dans le plan (\vec{a}_1, \vec{a}_2) et sa norme est :

$$\|\vec{b}_1\| = 2\pi \frac{a_2 a_3}{a_1 a_2 a_3 \sin \theta} = \frac{2\pi}{a_1 \sin \theta} , \quad (8.98)$$

et, de même :

$$\|\vec{b}_2\| = 2\pi \frac{a_3 a_1}{a_1 a_2 a_3 \sin \theta} = \frac{2\pi}{a_2 \sin \theta} \quad \|\vec{b}_3\| = 2\pi \frac{a_1 a_2 \sin \theta}{a_1 a_2 a_3 \sin \theta} = \frac{2\pi}{a_3} . \quad (8.99)$$

On prend maintenant la limite $a_3 \rightarrow +\infty$; la norme de \vec{b}_3 s'annule et on obtient²⁸ le réseau réciproque, de dimension 2 et situé géométriquement dans le même plan que le réseau direct ; \vec{b}_1 et \vec{b}_2 sont orthogonaux respectivement à \vec{a}_2 et \vec{a}_1 . Enfin, pour un réseau unidimensionnel engendré par \vec{a} , le réseau réciproque engendré par le seul et unique vecteur \vec{b} , colinéaire à \vec{a} (il reste orthogonal à deux dimensions fictives que l'on peut toujours choisir orthogonales à \vec{a}) et de norme $b = 2\pi/a$.

Remarque

Affirmer que le réseau réciproque d'un réseau de Bravais bidimensionnel est de dimension $d = 2$ est conforme à l'usage ; toutefois, si l'on s'en tient à la définition première (8.71), il est plus logique de dire que, en pareil cas, $\vec{\mathcal{B}}$ est constitué de droites infinies perpendiculaires au réseau direct et dont les intersections avec \mathbb{R}^2 constituent le réseau réciproque bidimensionnel défini par le processus de limite ci-dessus. En effet, soit un vecteur \vec{K} ainsi défini dans le plan ($\vec{K} \in \vec{\mathcal{B}}_{d=2} \equiv \mathbb{Z}^2$) et satisfaisant donc :

$$e^{i\vec{K} \cdot \vec{R}} = 1 \quad \forall \vec{R} \in \mathcal{B}_{d=2} \equiv \mathbb{Z}^2 . \quad (8.100)$$

Alors, si on ajoute à \vec{K} (dans le plan) n'importe quel vecteur perpendiculaire \vec{K}_\perp à ce plan :

$$\vec{K}' = \vec{K} + \vec{K}_\perp , \quad (8.101)$$

l'égalité (8.100) sera encore satisfaite. Le même argument permet de dire que le réseau réciproque d'un réseau de Bravais unidimensionnel $\mathcal{B}_{d=1} \equiv \mathbb{Z}$ de pas a est un ensemble de plans parallèles perpendiculaires à ce réseau. Le réseau direct coupe cette famille de plans en des points équidistants d'abscisses $m(2\pi)/a$, $m \in \mathbb{Z}$.

Ces distinctions peuvent jouer un rôle important quand la loi de conservation de l'impulsion est pertinente. Par exemple, si des photons (définis dans \mathbb{R}^3 !) d'impulsion $\hbar\vec{k}$ sont couplés à de la matière bidimensionnelle, la loi de conservation de l'impulsion ne contraint que la *projection* de \vec{k} : tous les photons ne différant que par un vecteur quelconque perpendiculaire au plan de la matière satisfont la conservation de l'impulsion. Autrement dit, une excitation matérielle d'impulsion $\hbar\vec{q}$ se trouve couplée à un *continuum* de photons de vecteur d'onde $\vec{k} = \vec{q} + \vec{q}_\perp$, avec $\vec{q} \cdot \vec{q}_\perp = 0$.

8.4.2 Exemples de réseaux réciproques

Donnons maintenant quelques exemples de réseaux réciproques. Il est évident que le réseau cubique simple de paramètre a a pour réseau réciproque un réseau également cubique simple, de côté $2\pi/a$. Pour le cubique centré,

²⁸En réalité, les choses sont plus subtiles. Il est sûr que si l'on prend la limite $a_3 \rightarrow +\infty$, tout point d'abscisse entière *finie* se retrouve dans le plan. Mais *quid* des points à l'infini ?

choisissons les vecteurs primitifs suivants :

$$\vec{a}_1 = a\vec{i} \quad \vec{a}_2 = a\vec{j} \quad \vec{a}_3 = \frac{a}{2}(\vec{i} + \vec{j} + \vec{k}) . \quad (8.102)$$

On a $(\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_3 = a^2(a/2) = a^3/2$. Maintenant :

$$\vec{b}_1 = \frac{2\pi}{a^3/2} \frac{a^2}{2} \vec{j} \times (\vec{i} + \vec{j} + \vec{k}) = \frac{2\pi}{a} (\vec{i} - \vec{k}) , \quad (8.103)$$

$$\vec{b}_2 = \frac{2\pi}{a^3/2} \frac{a^2}{2} (\vec{i} + \vec{j} + \vec{k}) \times \vec{i} = \frac{2\pi}{a} (\vec{j} - \vec{k}) , \quad (8.104)$$

$$\vec{b}_3 = \frac{2\pi}{a^3/2} a^2 \vec{i} \times \vec{j} = \frac{4\pi}{a} \vec{k} . \quad (8.105)$$

De façon équivalente, on peut aussi introduire :

$$\vec{b}'_1 = \vec{b}_1 + \vec{b}_3 = \frac{2\pi}{a} (\vec{k} + \vec{i}) , \quad \vec{b}'_2 = \vec{b}_2 + \vec{b}_3 = \frac{2\pi}{a} (\vec{j} + \vec{k}) , \quad \vec{b}'_3 = \vec{b}_1 + \vec{b}_2 + \vec{b}_3 = \frac{2\pi}{a} (\vec{i} + \vec{j}) , \quad (8.106)$$

et on reconnaît en $(\vec{b}'_1, \vec{b}'_2, \vec{b}'_3)$ un jeu de vecteurs primitifs du CFC de maille cubique conventionnelle de côté $\frac{4\pi}{a}$. La maille primitive du CC direct a pour volume $v = \frac{a^3}{2}$, celle du réseau réciproque a pour volume $\tilde{v} = \frac{(2\pi)^3}{(a^3/2)} = \frac{16\pi^3}{a^3}$. D'un autre côté, le volume de la maille primitive d'un CFC de côté $4\pi/a$ est $(1/4)(4\pi/a)^3 = 16\pi^3/a^3 = \tilde{v}$, comme il se doit.

Comme le réseau réciproque du réseau réciproque est le réseau direct, le réseau réciproque d'un CFC est un réseau CC. Si le volume v du CFC est $\frac{a^3}{4}$, celui de son réseau réciproque est $\frac{(2\pi)^3}{a^3/4} = \frac{32\pi^3}{a^3} \equiv \frac{1}{2} \left(\frac{4\pi}{a}\right)^3$. Le réciproque d'un CFC de côté a est donc le CC de côté $\frac{4\pi}{a}$.

Comme dernier exemple, prenons un réseau hexagonal, engendré par deux vecteurs \vec{a}_1 et \vec{a}_2 de même norme et faisant entre eux un angle de $\frac{\pi}{3}$, et par un vecteur \vec{c} perpendiculaire au plan défini par \vec{a}_1 et \vec{a}_2 et de longueur arbitraire c . Pour faire le calcul, on définit un repère orthonormé $(\vec{i}, \vec{j}, \vec{k})$ sur lequel :

$$\vec{a}_1 = a\vec{i} \quad \vec{a}_2 = \frac{a}{2}(\vec{i} + \sqrt{3}\vec{j}) \quad \vec{a}_3 = c\vec{k} . \quad (8.107)$$

On trouve alors :

$$\vec{a}_2 \times \vec{a}_3 = \frac{ac}{2}(\sqrt{3}\vec{i} - \vec{j}) \quad \vec{a}_3 \times \vec{a}_1 = ac\vec{j} \quad \vec{a}_1 \times \vec{a}_2 = a^2 \frac{\sqrt{3}}{2} \vec{k} , \quad (8.108)$$

et :

$$(\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_3 = \frac{\sqrt{3}}{2} a^2 c . \quad (8.109)$$

Il en résulte :

$$\vec{b}_1 = \frac{2\pi}{a\sqrt{3}}(\sqrt{3}\vec{i} - \vec{j}) \quad \vec{b}_2 = \frac{4\pi}{a\sqrt{3}}\vec{j} \quad \vec{b}_3 = \frac{2\pi}{c}\vec{k} . \quad (8.110)$$

Ces résultats montrent que le réseau réciproque de l'hexagonal (a, c) est un réseau hexagonal $(4\pi/a\sqrt{3}, 2\pi/c)$, simplement tourné de $\pi/6$ par rapport au réseau direct. Si on prend le réciproque de ce dernier, il faut encore tourner de $\pi/6$ et on retrouve le réseau de départ. Enfin, en faisant tendre c vers l'infini pour construire le réseau réciproque du réseau bidimensionnel, la longueur du vecteur \vec{b}_3 tend vers zéro : les points de \vec{B} le long de \vec{k} deviennent infiniment proches les uns des autres et engendrent une pseudo-droite ayant la puissance de \mathbb{Z} (voir la note 28).

Comme le réseau réciproque est un réseau de Bravais, tout ce qui a été dit concernant le choix des vecteurs primitifs, la maille primitive, la maille conventionnelle, *etc.* est applicable à \vec{B} . En particulier, on peut définir une maille du genre Wigner - Seitz pour \vec{B} . Pour des raisons qui seront plus claires dans la suite, la maille de Wigner - Seitz du réseau réciproque est appelée 1^{ère} zone de Brillouin (en abrégé : BZ1). Ainsi, ce que l'on appelle 1^{ère} zone de Brillouin d'un CFC est géométriquement identique à la maille de Wigner - Seitz d'un CC.

8.4.3 Plans réticulaires

On appelle plan réticulaire un plan contenant au moins trois nœuds d'un réseau de Bravais, c'est-à-dire, partant de là, une infinité. Il existe une relation très étroite entre les plans réticulaires et les vecteurs du réseau réciproque ; cette relation joue un rôle important pour les phénomènes de diffraction et peut être appréhendée sur une stricte base géométrique.

Trois nœuds définissent un plan ; tous les plans parallèles à un même plan forment une famille de plans, équidistants les uns des autres d'une distance d dépendant de la famille considérée. Il y a manifestement une infinité de façons de résoudre un réseau en un ensemble de plans réticulaires.

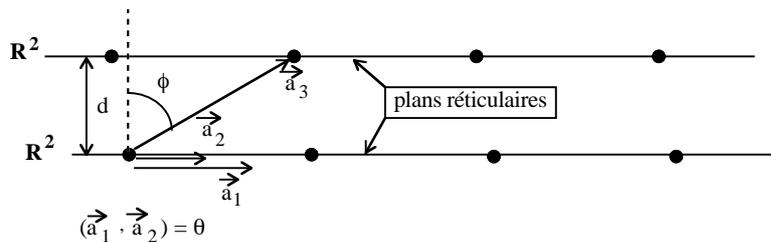


Figure 8.11: La distance d entre deux plans adjacents d'une même famille contenant \vec{a}_1 et \vec{a}_2 est simplement reliée au module de \vec{b}_3 (voir (8.112)).

Un plan réticulaire est un sous-réseau de dimension 2, où on peut définir deux vecteurs primitifs notés \vec{a}_1 et \vec{a}_2 . Par définition, le vecteur \vec{b}_3 est orthogonal à \vec{a}_1 et \vec{a}_2 , c'est donc une normale à ce plan. Prenons pour \vec{a}_3 le vecteur indiqué sur la figure 8.11. Si θ est l'angle entre \vec{a}_1 et \vec{a}_2 et ϕ l'angle entre \vec{a}_3 et la normale au plan, on a :

$$\|\vec{b}_3\| = 2\pi \frac{|a_1 a_2 \sin \theta|}{|\vec{a}_3 \cdot (\vec{a}_1 \times \vec{a}_2)|} = 2\pi \frac{a_1 a_2 |\sin \theta|}{a_1 a_2 a_3 |\sin \theta \cos \phi|} . \quad (8.111)$$

Comme $a_3 |\cos \phi| = d$, distance entre deux plans adjacents, il vient finalement :

$$\|\vec{b}_3\| = \frac{2\pi}{d} . \quad (8.112)$$

Ainsi, une famille de plans étant précisée, il existe des vecteurs du réseau réciproque orthogonaux à cette famille et le plus petit d'entre eux²⁹ a une norme s'exprimant simplement, par (8.112), en fonction de la distance entre deux plans consécutifs. Clairement, l'argument précédent ne privilégie pas un certain vecteur primitif de $\vec{\mathcal{B}}$ (celui noté \vec{b}_3 ci-dessus) et vaut pour l'ensemble du réseau réciproque : tout en conservant les \vec{a}_i précédents (donc les mêmes \vec{b}_j), mais en choisissant une autre famille de plans, on montrerait de la même façon qu'il existe un vecteur \vec{K} orthogonal à cette famille et ayant *a priori* trois composantes non-nulles sur ces \vec{b}_j .

La relation étroite entre une famille de plans réticulaires et les vecteurs du réseau réciproque permet de définir commodément l'orientation de la famille considérée. Comme il est usuel de définir un plan par sa normale, on peut se donner un certain vecteur du réseau réciproque pour qualifier une famille de plans. On en vient ainsi à ce que l'on appelle les indices de Miller d'un plan : ce sont les composantes du plus petit vecteur de $\vec{\mathcal{B}}$ normal à ce plan, dont la longueur est de ce fait reliée à la distance entre deux plans réticulaires voisins selon (8.112) ; on les note traditionnellement (h, k, l) et la distance correspondante est notée d_{hkl} . Ainsi, un plan dont les indices de Miller sont (h, k, l) est un plan normal au vecteur $\vec{K} = h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3$. Les indices de Miller sont donc des entiers et comme il s'agit du plus petit vecteur réciproque, les trois nombres h, k et l sont premiers entre eux (par exemple, ils ne sauraient être tous pairs). Bien sûr, le triplet (h, k, l) dépend du choix des vecteurs primitifs \vec{b}_j .

Les indices de Miller ont une interprétation géométrique simple dans le réseau direct, laquelle sert d'ailleurs parfois de définition. Etant donné (h, k, l) , l'équation du plan perpendiculaire au vecteur réciproque

²⁹ \vec{b}_3 est bien le plus petit vecteur réciproque normal aux plans considérés.

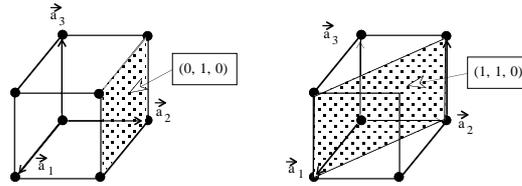


Figure 8.12: Exemples d'indices de Miller pour un réseau cubique simple

correspondant, \vec{K} , est $\vec{K} \cdot \vec{r} = C$, avec $\vec{K} = h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3$ et C une constante. Ce plan coupe les axes portant les \vec{a}_i en trois points d'abscisses x_1 , x_2 et x_3 . Ces points, appartenant au plan, satisfont son équation et on a donc :

$$\vec{K} \cdot (x_i \vec{a}_i) = C \quad (i = 1, 2, 3) . \quad (8.113)$$

Comme $\vec{K} \cdot \vec{a}_i = (h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3) \cdot \vec{a}_i$ et compte tenu de (8.77), on trouve :

$$2\pi x_1 h = 2\pi x_2 k = 2\pi x_3 l = C , \quad (8.114)$$

soit :

$$x_1 = \frac{C}{2\pi h} \quad x_2 = \frac{C}{2\pi k} \quad x_3 = \frac{C}{2\pi l} . \quad (8.115)$$

Les coordonnées des points d'intersection des plans d'une famille sont donc inversement proportionnelles aux indices de Miller (voir figure 8.12).

8.5 Diffraction par un réseau

En raison de l'ordre géométrique à longue distance présent dans un cristal, il existe un très grand nombre N de diffuseurs qui vont pouvoir *coopérer* et produire des figures de diffusion extrêmement sélectives en direction, la longueur d'onde du rayonnement étant fixée. On parle alors traditionnellement de *diffraction* pour désigner cette sélectivité étant entendu que, au niveau de chaque diffuseur, il s'agit toujours de la simple diffusion élastique de la lumière au sens rappelé ci-dessus. C'est l'énormité de N allié à l'existence d'un ordre à portée infinie qui change qualitativement les phénomènes observés quand on passe d'un milieu désordonné (ou présentant un ordre à courte distance) à un cristal. En parlant schématiquement, on passe d'une réponse *floue* dans le premier cas à une réponse par tout ou rien dans le second.

8.5.1 Diffusion des rayons X et des neutrons

La diffusion des rayons X et des neutrons³⁰ est un outil de tout premier plan pour étudier la structure de la matière, qu'elle soit en phase diluée ou condensée. Le principe en est simple : on envoie sur le système à étudier une "sonde" externe parfaitement caractérisée (énergie ou longueur d'onde, polarisation³¹). L'interaction sonde - système est complexe en général ; pour simplifier, seules les interactions *élastiques* seront considérées dans la suite³². Dans ce cadre, la "lumière" issue du milieu a la même fréquence (énergie) que l'onde incidente ; son vecteur d'onde \vec{k}_f a donc le même module que celui du faisceau incident, \vec{k}_i :

$$\|\vec{k}_f\| = \|\vec{k}_i\| . \quad (8.116)$$

En tout cas, l'interaction sonde - cible est supposée faible, puisque l'on veut une photographie de la matière telle qu'elle est, c'est-à-dire perturbée aussi peu que possible.

³⁰et des électrons pour les surfaces.

³¹Dans le cas des neutrons, la polarisation est liée au spin ($\pm\hbar/2$) des neutrons incidents ; un faisceau non polarisé est un faisceau dans lequel l'orientation du spin est aléatoire.

³²Bien sûr, les expériences de diffusion inélastique sont aussi d'usage courant et apportent d'autres informations.

Rayons X

Dans le cas des rayons X, l'énergie des photons est de quelques dizaines de keV, très supérieure aux énergies intra-atomiques ou intra-moléculaires, qui sont de l'ordre de quelques eV au plus. Le rayonnement électromagnétique incident, essentiellement par son champ électrique (de module \mathcal{E}_0), met en vibration forcée, à la même fréquence, les électrons du milieu (qui lui apparaissent comme presque libres) : classiquement, c'est ce que l'on appelle généralement la diffusion Thomson³³. Pour une onde incidente non polarisée, la diffusion Thomson est caractérisée à haute fréquence par une section efficace indépendante de la fréquence, $\sigma_T = (8\pi/3)r_e^2$, où r_e est appelé historiquement le "rayon classique" de l'électron ($r_e = e^2/(4\pi\epsilon_0 mc^2) \approx 3 \times 10^{-15}$ m)³⁴ ; r_e est la longueur la plus simple que l'on peut former avec la vitesse de la lumière, la masse et la charge de l'électron. La vibration forcée donne aux électrons une accélération non-nulle et, comme toute particule chargée accélérée rayonne, ceux-ci émettent un rayonnement électromagnétique, évidemment de même fréquence puisqu'il s'agit d'une vibration forcée. La puissance rayonnée vers l'extérieur par une charge accélérée, \mathcal{P} , est donnée par la formule classique de Larmor :

$$\mathcal{P} = \frac{2}{3} \frac{e'^2 \dot{v}^2}{c^3} \quad \left(e'^2 = \frac{e^2}{4\pi\epsilon_0} \right) . \quad (8.117)$$

En ne retenant que les effets majoritaires, l'accélération d'un électron, \dot{v} , se déduit immédiatement du champ électrique de l'onde qui le met en mouvement³⁵. Pour un électron accéléré bien localisé dans un domaine très petit par rapport à la longueur d'onde incidente, le champ électrique qu'il rayonne à une grande distance R a pour amplitude :

$$\mathcal{E} = \frac{r_e}{R} \mathcal{E}_0 , \quad (8.118)$$

\mathcal{E}_0 étant l'amplitude du champ externe. En réalité, pour les atomes et les molécules, les électrons sont répartis sur des distances de l'ordre de l'Å ; le caractère diffus de la densité électronique produit un premier type de déphasage spatial et donne lieu à des effets d'interférences dont il faut tenir compte³⁶. Ceci introduit un facteur multiplicatif, usuellement noté f et appelé facteur de structure (intra-atomique) ; c'est essentiellement la transformée de Fourier de la densité électronique, calculée pour le transfert de moment $\vec{q} = \vec{k}_f - \vec{k}_i$; $f(\vec{q} = 0)$ est égal à Z , nombre total d'électrons appartenant au centre diffuseur.

En raisonnant d'emblée avec les composantes de Fourier, il convient de trouver les déphasages relativement à un point arbitraire. Par rapport à l'origine $\vec{r} = 0$, le déphasage de la lumière émise par un seul centre diffuseur de facteur de structure f_1 situé au point \vec{r}_1 est $f_1 e^{i(\vec{k}_f - \vec{k}_i) \cdot \vec{r}_1}$; l'amplitude (complexe) du champ électrique émis par vibration forcée est, à la distance R , donnée par³⁷ :

$$\mathcal{E}_1 = \mathcal{E}_0 f_1 \frac{r_e}{R} e^{i(\vec{k}_f - \vec{k}_i) \cdot \vec{r}_1} . \quad (8.119)$$

Pour un ensemble de N diffuseurs situés aux points \vec{r}_n , $n = 1, 2, \dots, N$, l'amplitude diffusée est la somme :

$$\mathcal{E} = \mathcal{E}_0 \frac{r_e}{R} \sum_{n=1}^N f_n e^{i(\vec{k}_f - \vec{k}_i) \cdot \vec{r}_n} . \quad (8.120)$$

Ceci suppose que le champ est polarisé rectilignement, mettant tous les oscillateurs en vibration forcée le long de la même direction. Le carré du champ électrique rayonné vaut donc :

$$\mathcal{E}^2 = \left(\mathcal{E}_0 \frac{r_e}{R} \right)^2 \sum_{n,m=1}^N f_n f_m^* e^{i(\vec{k}_f - \vec{k}_i) \cdot (\vec{r}_n - \vec{r}_m)} . \quad (8.121)$$

³³Il existe également une diffusion inélastique, appelée diffusion Compton, qui peut donner de précieux renseignements sur la distribution des vitesses au sein du système.

³⁴La section efficace Thomson est donc microscopique à l'échelle atomique – c'est une surface dont le rayon est de l'ordre des dimensions nucléaires. Au contraire, la section efficace de diffusion *résonnante* est, elle, gigantesque par rapport à la taille atomique puisqu'elle vaut en gros λ_0^2 , où λ_0 est la longueur d'onde de résonance, qui est dans le domaine optique ou UV pour les transitions électroniques ordinaires des atomes et des molécules.

³⁵Pour une onde polarisée, l'électron est mis en vibration le long de la direction fixée par la polarisation électrique.

³⁶Dans l'optique ou le visible, ce déphasage est négligeable pour les transitions de type dipolaire électrique.

³⁷On omet tous les facteurs temporels $e^{i\omega t}$ qui, pour les processus élastiques, seront communs à toutes les ondes diffusées. Le calcul détaillé du déphasage, faisant apparaître le transfert de moment, est rappelé plus loin (section 8.5, figure 8.13).

Enfin, si la source n'est pas polarisée, il convient de moyenner sur les angles, ce qui revient à moyenner sur toutes les directions possibles les petites "antennes" électroniques. Comme toujours, on suppose que l'absence de polarisation de l'onde incidente se traduit par une équiprobabilité des directions de vibration de son champ électrique. \vec{k}_f et \vec{k}_i étant fixés une fois pour toutes et si θ désigne l'angle de diffusion, $\theta = (\vec{k}_i, \vec{k}_f)$, on trouve que le carré du module du champ moyen est :

$$\bar{\mathcal{E}}^2 = \frac{1}{2} (1 + \cos^2 \theta) \left(\mathcal{E}_0 \frac{r_e}{R} \right)^2 \sum_{n,m=1}^N f_n f_m^* e^{i(\vec{k}_f - \vec{k}_i) \cdot (\vec{r}_n - \vec{r}_m)} . \quad (8.122)$$

En introduisant finalement l'angle $\alpha_{nm} = (\vec{k}_f - \vec{k}_i, \vec{r}_n - \vec{r}_m)$, en désignant par λ la longueur d'onde incidente et en posant $r_{nm} = \|\vec{r}_n - \vec{r}_m\|$, l'intensité diffusée I peut se mettre sous la forme ($\|\vec{q}\| = 2 \|\vec{k}_i\| \sin(\theta/2)$) :

$$I = I_0 \sum_{n,m=1}^N f_n f_m^* e^{iK r_{nm} \cos \alpha_{nm}} , \quad K = \frac{4\pi}{\lambda} \sin \frac{\theta}{2} . \quad (8.123)$$

Le cas de diffuseurs immobiles situés aux nœuds d'un réseau sera traité dans la section suivante. Dans un milieu désordonné (un fluide par exemple, ou un verre), l'intensité mesurée correspond en fait à la moyenne des orientations de $\vec{r}_n - \vec{r}_m$. L'exponentielle portant les déphasages spatiaux dans (8.123) doit donc être moyennée sur l'angle solide $d\Omega = 2\pi \sin \alpha_{nm} d\alpha_{nm}$; en admettant un désordre complet, pour toute distance r_{nm} donnée, toutes les orientations possibles sont équiprobables ; dans ces conditions, cette moyenne s'écrit :

$$\int \frac{d\Omega}{4\pi} e^{iK r_{nm} \cos \alpha_{nm}} , \quad (8.124)$$

et se calcule aisément :

$$\frac{1}{4\pi} \int_0^\pi 2\pi \sin \alpha d\alpha e^{iK r_{nm} \cos \alpha} = \frac{\sin K r_{nm}}{K r_{nm}} . \quad (8.125)$$

L'intensité moyenne est ainsi donnée par :

$$\bar{I} = I_0 \sum_{n,m=1}^N f_n f_m^* \frac{\sin K r_{nm}}{K r_{nm}} . \quad (8.126)$$

Ceci constitue la formule de Debye, à laquelle on peut donner une forme plus agréable. On commence par isoler les termes carrés $n = m$ ($\sum_{n,m=1}^N = \sum_{n=1}^N + \sum_{n \neq m=1}^N$) ; dans l'hypothèse d'un échantillon de taille macroscopique (où l'on peut oublier les effets de surface), tout se passe comme s'il existait une quasi-invariance par translation : il en résulte que la somme double est extensive en N . Enfin, en prenant des diffuseurs monoatomiques tous identiques ($f_n = f \forall n$), (8.126) prend finalement la forme :

$$\bar{I} = N I_0 |f|^2 \left(1 + \sum_n \frac{\sin K r_n}{K r_n} \right) . \quad (8.127)$$

La somme discrète peut toujours s'écrire sous forme intégrale. En effet, s'il s'agit de se borner aux propriétés à grande échelle, on peut définir une densité ρ : c'est, pour un volume mésoscopique δV , le nombre de particules dans ce volume "élémentaire" divisé par δV et alors une description pseudo-continue peut être utilisée ; si l'aspect granulaire de la matière doit être conservé, il suffit d'introduire des fonctions de Dirac pour représenter les densités atomiques.

Notant $G(r)$ le nombre de particules par unité de volume situées à la distance r de l'origine³⁸, l'intensité moyenne (8.127) s'écrit comme suit dans la description continue³⁹ :

$$\bar{I} = N I_0 |f|^2 \left(1 + \int_0^{\mathcal{R}} 4\pi r^2 dr G(r) \frac{\sin Kr}{Kr} \right) , \quad (8.128)$$

³⁸La fonction G est supposée à symétrie sphérique. C'est en fait le nombre de particules situé à la distance r de l'une d'entre elles, c'est donc la fonction de corrélation de paire.

³⁹ $\int_{0 \leq r \leq \mathcal{R}} d^3r e^{i\vec{K} \cdot \vec{r}} = \int_0^{\mathcal{R}} 4\pi r^2 dr (\sin Kr)/(Kr)$

où \mathcal{R} représente le rayon de l'échantillon, supposé sphérique. Pour extraire la composante représentant la diffusion triviale vers l'avant ($\vec{q} = \vec{0}$), il suffit d'ajouter et de retrancher la densité moyenne ρ :

$$\bar{I} = N I_0 |f|^2 \left[1 + \int_0^{\mathcal{R}} 4\pi r^2 dr [G(r) - \rho] \frac{\sin Kr}{Kr} + \rho \int_0^{\mathcal{R}} 4\pi r^2 dr \frac{\sin Kr}{Kr} \right]. \quad (8.129)$$

Pour un échantillon macroscopique (où l'on peut oublier les effets de surface), la dernière intégrale peut être remplacée par $(2\pi)^3 \delta(\vec{K})$; ce terme représente la diffusion vers l'avant, sans intérêt car ne contenant aucune information relative au système étudié. Au contraire, cette dernière est contenue dans les deux premiers termes du crochet ; on définit ainsi précisément, en posant $G(r) = \rho g(r)$:

$$S(K) = 1 + \rho \int_0^{+\infty} 4\pi r^2 dr [g(r) - 1] \frac{\sin Kr}{Kr}, \quad (8.130)$$

qui est réellement la signature du système (on voit apparaître la fonction $h(r) = g(r) - 1$, qui tend vers zéro aux grands r , une condition nécessaire pour pouvoir lui associer une transformée de Fourier). En jouant avec la symétrie sphérique, on voit que $S(\vec{K})$ (qui, en fait, ne dépend ici que de $\|\vec{K}\|$) est aussi :

$$S(\vec{K}) = 1 + \rho \int d^3r [g(r) - 1] e^{i\vec{K}\cdot\vec{r}}, \quad (8.131)$$

ce qui permet d'effectuer à vue l'inversion de Fourier pour obtenir :

$$g(r) = 1 + \frac{1}{8\pi^3\rho} \int d^3K [S(\vec{K}) - 1] e^{-i\vec{K}\cdot\vec{r}} \quad (8.132)$$

et en déduire finalement la relation d'intérêt pratique :

$$g(r) - 1 \equiv h(r) = \frac{1}{2\pi^2\rho r} \int_0^{+\infty} dK [S(K) - 1] K \sin Kr. \quad (8.133)$$

En effet, en mesurant $S(K)$ avec une très bonne résolution en K (et en explorant un grand intervalle en K), on peut reconstruire $g(r)$ par la transformation intégrale (8.133). Cette relation joue donc un rôle majeur pour toutes les expériences de diffusion X.

Neutrons

La diffusion des neutrons⁴⁰ relève de la même description. La longueur d'onde à prendre en compte est la longueur d'onde associée par la relation de de Broglie $\lambda = \frac{h}{Mv} = \frac{h}{\sqrt{2ME}}$; pour les neutrons ceci s'exprime commodément comme suit :

$$\lambda_{\text{\AA}} \simeq \frac{0.286}{\sqrt{E_{\text{eV}}}}. \quad (8.134)$$

Pour un neutron thermique, $E \approx (k_B T)_{T=300\text{K}} \approx 25$ meV, de sorte que $\lambda \approx 1.8$ \AA. C'est dire que ce sont les neutrons thermiques qu'il convient d'utiliser pour sonder la matière ordinaire, qu'elle soit à l'état fluide ou solide. Bien sûr, un faisceau neutronique est en général non monochromatique et la distribution du module de la vitesse (une distribution de Maxwell) reflète l'état thermique du modérateur ; si celui-ci est à une température de l'ordre de 100 K, le maximum de la maxwellienne survient précisément pour E de l'ordre de 25 meV.

Les neutrons incidents et les constituants du fluide peuvent interagir principalement de deux façons, étant entendu que le neutron est insensible à l'interaction électromagnétique⁴¹. La première, relevant de l'interaction forte, est à très courte portée (la dimension nucléaire) et peut être considérée comme étant de contact⁴² pour

⁴⁰ Les techniques de diffusion neutronique sont nettement plus récentes que celles utilisant les rayons X ; elles ont commencé à se développer dans les années 50, lorsque des sources de neutrons suffisamment intenses sont devenues opérationnelles. Le formalisme théorique de base est principalement dû à Fermi et Schwinger.

⁴¹ en mettant de côté des effets extrêmement fins impliquant la structure de charge du neutron.

⁴² Une interaction est dite de *contact* si elle est de portée nulle (les objets n'interagissent que s'ils sont au contact l'un de l'autre) ; on la représente alors par une fonction de Dirac.

la diffusion sur des atomes ou des molécules. Le neutron, grâce à son moment magnétique associé à son spin ($S = \frac{1}{2}$ en unités \hbar), peut en outre se coupler aux électrons non appariés, auquel cas la portée de l'interaction effective est d'ordre atomique (l'échelle de longueur des fonctions d'onde de ces électrons).

En ce qui concerne le couplage de contact résultant de l'interaction forte, on l'écrit usuellement :

$$V_c(\vec{r}) = \frac{2\pi\hbar^2}{M} b \delta(\vec{r} - \vec{R}) . \quad (8.135)$$

C'est ce que l'on appelle le pseudo-potential de Fermi. \vec{r} et \vec{R} sont respectivement les positions du neutron et du noyau diffuseur ; b est la longueur de diffusion⁴³, qui varie d'un noyau à l'autre (et donc en particulier d'un isotope à l'autre) et dépend également du spin nucléaire. De la forme (8.135), il résulte que pour un diffuseur unique les sections efficaces différentielle et totale sont (voir (8.169)) :

$$\frac{d\sigma}{d\Omega} = b^2 , \quad \sigma = 4\pi b^2 . \quad (8.136)$$

b rentre dans ces expressions de la même façon que le rayon d'une sphère dure pour un problème classique de diffusion⁴⁴.

En ce qui concerne le couplage magnétique, notamment avec les spins des électrons célibataires, l'interaction est principalement de la forme dipolaire magnétique :

$$V_{\text{magn}} = -\gamma\mu_N \vec{\sigma} \cdot \vec{B} , \quad (8.137)$$

où $\gamma = -1.91$ est le facteur gyromagnétique du neutron, $\mu_N = |e| \hbar / (2M) > 0$ est le magnéton nucléaire et \vec{B} le champ magnétique produit par les électrons non-appariés – dont l'expression est loin d'être simple [37] : elle contient non seulement le champ dipolaire dû aux spins électroniques (couplage spin-spin) mais également le champ magnétique créé par le mouvement orbital des électrons (couplage spin-orbite). $\vec{\sigma}$ est l'opérateur vectoriel construit avec les trois matrices de Pauli.

D'une façon ou d'une autre, la diffusion élastique des neutrons, avec ses spécificités, donne accès, tout comme la diffusion X, à la fonction de corrélation $g(r)$, par l'intermédiaire de $S(K)$ (par exemple, par la relation (8.132)). Il est à noter que la diffusion *inélastique* (diffusion Brillouin), avec variation du nombre de phonons, donne des renseignements sur la loi de dispersion de ces derniers, principalement sur la vitesse du son au sein du milieu⁴⁵.

8.5.2 Conditions de von Laue et de Bragg

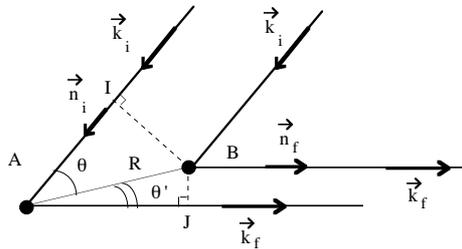


Figure 8.13: Déphasage pour deux diffuseurs séparés de la distance R .

Le lien entre sélectivité et effets coopératifs de diffuseurs ordonnés est facile à saisir. Soit \vec{k}_i le vecteur d'onde incident, \vec{k}_f le vecteur d'onde diffusé élastiquement. L'onde diffusée par l'atome situé au nœud \vec{R} d'un

⁴³Dans \mathbb{R}^3 , la fonction $\delta(\vec{r})$ est homogène à L^{-3} ; il en résulte que b a bien la dimension d'une longueur.

⁴⁴au facteur 4 près lié à la spécificité quantique (voir par exemple [8], complément C_{VIII}).

⁴⁵Si $\omega(\vec{k})$ est la loi de dispersion des phonons acoustiques, la vitesse du son, usuellement notée c , est simplement reliée à $(\vec{\nabla}\omega)_{\vec{k}=0}$ (pour un milieu isotrope, $c = \|(\vec{\nabla}\omega)_{\vec{k}=0}\|$).

réseau de Bravais est caractérisé par un déphasage $\vec{q} \cdot \vec{R}$ où $\vec{q} = \vec{k}_f - \vec{k}_i$ est le transfert de moment. En effet (voir figure 8.13), la différence de chemin pour les ondes diffusées par A et B est IA+AJ, soit :

$$R \cos \theta + R \cos \theta' \equiv \vec{R} \cdot (-\vec{n}_i) + \vec{R} \cdot \vec{n}_f = \vec{R} \cdot (\vec{n}_f - \vec{n}_i) \quad (\vec{R} = \vec{AB}) . \quad (8.138)$$

Pour deux diffuseurs séparés par \vec{R} le déphasage est donc :

$$\frac{2\pi}{\lambda} \vec{R} \cdot (\vec{n}_f - \vec{n}_i) \equiv \vec{R} \cdot (\vec{k}_f - \vec{k}_i) \equiv \vec{R} \cdot \vec{q} \quad \left(\lambda = \frac{2\pi}{\|\vec{k}_i\|} = \frac{2\pi}{\|\vec{k}_f\|} \right) . \quad (8.139)$$

Il en résulte que l'amplitude de l'onde diffusée par un réseau de Bravais \mathcal{B} est de la forme⁴⁶ :

$$\Re A_0 \sum_{\vec{R} \in \mathcal{B}} e^{-i(\omega t - \vec{q} \cdot \vec{R})} = \Re A_0 e^{-i\omega t} \sum_{\vec{R}} e^{i\vec{q} \cdot \vec{R}} . \quad (8.140)$$

La sommation sur \vec{R} est⁴⁷ :

$$S \equiv \sum_{\vec{R}} e^{i\vec{q} \cdot \vec{R}} = \sum_{\{n_i\} \in \mathbb{Z}^3} \prod_{i=1}^3 e^{in_i \vec{q} \cdot \vec{a}_i} = \prod_{i=1}^3 \sum_{n_i \in \mathbb{Z}} e^{in_i \vec{q} \cdot \vec{a}_i} = \prod_{i=1}^3 \left[2\pi \sum_{p_i \in \mathbb{Z}} \delta(\vec{q} \cdot \vec{a}_i - p_i 2\pi) \right] . \quad (8.142)$$

L'apparition des fonctions de Dirac révèle la sélectivité de la diffusion coopérative d'une infinité dénombrable

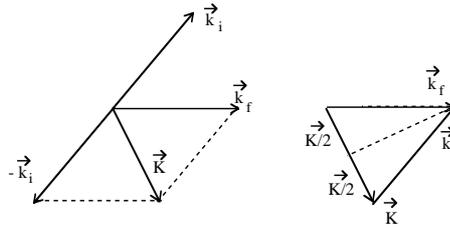


Figure 8.14: Illustration géométrique de la condition de von Laue.

d'objets ordonnés dans l'espace. $\vec{q} = \vec{k}_f - \vec{k}_i$ étant un vecteur de l'espace réciproque⁴⁸, il est naturel de le représenter (décomposer) sur les vecteurs primitifs \vec{b}_j du réseau réciproque :

$$\vec{q} = \sum_{j=1}^3 \mu_j \vec{b}_j , \quad (8.143)$$

où les μ_j sont des réels quelconques. Alors :

$$\vec{q} \cdot \vec{a}_i = \sum_{j=1}^3 \mu_j \vec{b}_j \cdot \vec{a}_i = 2\pi \mu_i , \quad (8.144)$$

de sorte que la somme S définie en (8.142) est :

$$S = \prod_{i=1}^3 \left[2\pi \sum_{p_i \in \mathbb{Z}} \delta(2\pi \mu_i - p_i 2\pi) \right] = \prod_{i=1}^3 \left[\sum_{p_i = -\infty}^{+\infty} \delta(\mu_i - p_i) \right] . \quad (8.145)$$

Ceci, compte tenu de la définition (8.143) des μ_j , montre que la somme S est nulle sauf si le transfert de moment \vec{q} est précisément un vecteur du réseau réciproque :

$$\exists \vec{K} \in \vec{\mathcal{B}} : \quad \vec{K} = \vec{k}_f - \vec{k}_i . \quad (8.146)$$

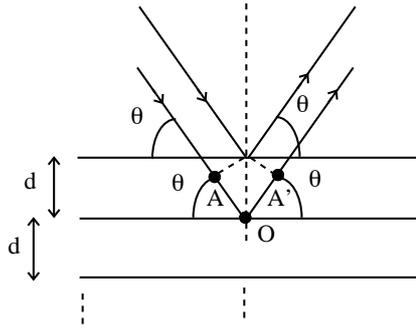


Figure 8.15: Illustration géométrique de la condition de Bragg.

Cette condition, dite condition de von Laue, a une traduction géométrique simple ; compte tenu du fait que les deux vecteurs d'onde \vec{k}_i et \vec{k}_f ont même module, $-\vec{k}_i$ et \vec{k}_f forment un losange dont \vec{K} est une diagonale ; ou encore : le vecteur formant la base d'un triangle isocèle construit sur \vec{k}_i et \vec{k}_f est un vecteur \vec{K} du réseau réciproque (voir figure 8.14).

Notons que la condition de von Laue ne fait aucune hypothèse particulière (autre qu'élastique) sur la nature de la diffusion de la lumière et résulte uniquement de l'existence d'un ordre à longue portée, matérialisé par le réseau de Bravais. Elle est équivalente à la condition dite de Bragg, qui repose au contraire sur une hypothèse supplémentaire, celle de la réflexion *spéculaire* sur les plans réticulaires (voir figure 8.15), chaque plan se comportant exactement comme un miroir. La différence de marche entre deux plans consécutifs est $AO + OA' = 2d \sin \theta$. L'interférence entre ces deux plans est donc constructive si :

$$2d \sin \theta = \text{entier} \times \text{longueur d'onde} . \quad (8.147)$$

Pour un ensemble semi-infini de plans, la sélectivité en direction est absolue et il n'y aura réflexion que pour les angles θ satisfaisant :

$$2d \sin \theta = n \lambda \quad (n \in \mathbb{N}) . \quad (8.148)$$

Cette relation est appelée condition de diffraction de Bragg⁴⁹.

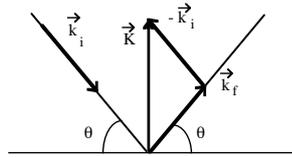


Figure 8.16: Equivalence entre les conditions de von Laue et de Bragg.

L'équivalence entre les deux conditions de von Laue et de Bragg se voit comme suit. En partant du schéma de Bragg (réflexion spéculaire sur un plan), on peut construire le vecteur $\vec{k}_f - \vec{k}_i$, qui est perpendiculaire au plan choisi, et a pour longueur (voir figure (8.16)) :

$$2 k_i \sin \theta = 2k_i \frac{n\lambda}{2d} = n \frac{2\pi}{d} . \quad (8.149)$$

Mais $(2\pi/d)$ est précisément la longueur du plus petit vecteur réciproque perpendiculaire à la famille de plans considérée, et, par ailleurs, $\vec{k}_f - \vec{k}_i \in \tilde{\mathcal{B}}$ est la condition de von Laue – ce qui achève de montrer l'équivalence.

⁴⁶ \Re désigne la partie réelle.

⁴⁷On utilise trois fois le résultat :

$$\sum_{n=-\infty}^{+\infty} e^{inx} = 2\pi \sum_{p=-\infty}^{+\infty} \delta(x - p2\pi) . \quad (8.141)$$

⁴⁸Il s'agit bien de l'espace réciproque, pas du réseau réciproque $\tilde{\mathcal{B}}$, où celui-ci est plongé. Les composantes de \vec{q} sont pour l'instant des nombres quelconques, pas entiers.

⁴⁹Elle montre en particulier qu'il ne peut y avoir réflexion de Bragg que si $\lambda/(2d) < 1$.

Que l'on raisonne d'une façon ou d'une autre, il existe donc une condition stricte pour que la diffraction se produise, que l'on peut traduire géométriquement par la construction dite d'Ewald. Étant donné le vecteur \vec{k}_i , on fixe son extrémité sur un nœud quelconque, O, du réseau réciproque et on trace une sphère centrée sur l'extrémité de ce vecteur, de rayon k_i ; par construction, tous les vecteurs \vec{k}_f allant du centre de la sphère à n'importe quel point de sa surface satisfont la condition (8.146). Il en résulte que tous les points de $\tilde{\mathcal{B}}$ se trouvant sur la surface de la sphère donnent lieu à la diffraction (voir figure 8.17). En général, il n'y a aucune coïncidence, ce qui traduit le fait que pour \vec{k}_i quelconque (en module et en orientation par rapport au cristal), il n'y a pas de diffraction de Bragg. \vec{k}_i étant fixé, il convient donc d'orienter convenablement le cristal pour observer un pic de Bragg.

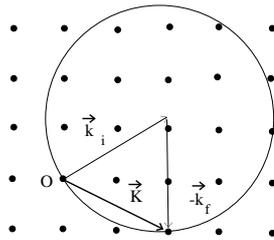


Figure 8.17: Construction d'Ewald.

Toutefois, plusieurs méthodes permettent d'obtenir des pics de diffraction sans prendre de précautions particulières :

- on utilise une source de rayons X polychromatique, comprise entre deux longueurs d'onde λ_1 et λ_2 . Alors, en traçant les deux sphères d'Ewald correspondantes, on voit que tous les nœuds du réseau situés *entre* les deux sphères donneront une diffusion sélective (voir figure 8.18)
- La méthode du cristal tournant. Maintenant, la source monochromatique étant fixe, on fait tourner le cristal lentement après l'avoir entouré d'un film sensible aux rayons X. Ainsi, la sphère d'Ewald tourne, balaie le réseau réciproque et rencontre donc un certain nombre de nœuds.
- La méthode des poudres (dite de Debye - Scherrer), qui utilise non un monocristal, mais une poudre polycristalline dont les grains, quoique petits (en fait, mésoscopiques), contiennent encore un nombre gigantesque de points du réseau direct, donnant une hypersélectivité, même avec les plus grandes résolutions spectrales.

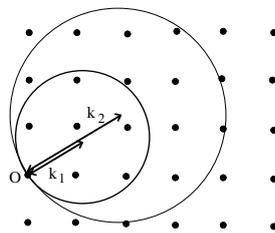


Figure 8.18: Utilisation d'une source polychromatique : tous les points de $\tilde{\mathcal{B}}$ situés entre les deux sphères donnent lieu à une diffraction de Bragg.

8.5.3 Facteurs de structure (géométrique et atomique)

Jusqu'à présent, seule la diffusion coopérative d'un ensemble ordonné d'objets ponctuels identiques a été considérée. Autrement dit, on a supposé que chaque maille primitive contient un objet à symétrie sphérique assez petit pour que sa dimension linéaire soit non-pertinente, autorisant à l'assimiler à un point. Par ailleurs, rien

n'a été dit sur l'objet qui se tient en chaque nœud, mais il est clair que l'amplitude de l'onde diffusée est proportionnelle à un certain facteur f mesurant l'aptitude de cet objet à diffuser plus ou moins le rayonnement incident.

En pratique, les solides cristallisés sont construits le plus souvent sur des réseaux avec base, formée de n objets (toujours ponctuels pour l'instant) assemblés suivant une géométrie connue (par exemple, un cubique simple avec une molécule diatomique à chaque nœud, et orientée d'une façon connue). Dès lors, l'interférence constructive des mailles primitives équivalentes par translation est modulée en amplitude par des déphasages additionnels de l'onde sur le volume contenant les objets non-structurés constituant la base. Ainsi, pour une diffraction associée au vecteur du réseau réciproque \vec{K} , il apparaît en outre un déphasage $\vec{K} \cdot (\vec{d}_i - \vec{d}_j)$ pour tout couple de diffuseurs situés aux points définis par les vecteurs \vec{d}_i et \vec{d}_j , représentant leur position \vec{d} relativement à une maille primitive. La lumière diffusée par une maille primitive a pour amplitude la somme des amplitudes diffusées par chaque élément du motif (base) ; supposant toujours pour l'instant que tous les diffuseurs sont physiquement identiques, chaque facteur $e^{i\vec{q} \cdot \vec{R}}$ de la somme S (8.142) est alors remplacé par :

$$\sum_{j=1}^n e^{i\vec{q} \cdot (\vec{R} + \vec{d}_j)} . \quad (8.150)$$

Il en résulte que S devient :

$$S \equiv \sum_{\vec{R}} e^{i\vec{q} \cdot \vec{R}} \rightsquigarrow S' = \sum_{\vec{R}} \sum_{j=1}^n e^{i\vec{q} \cdot (\vec{R} + \vec{d}_j)} . \quad (8.151)$$

Par définition, toutes les mailles primitives se valent et contiennent donc la même somme \sum_j . Celle-ci se factorise, d'où :

$$S' = \sum_{\vec{R}} \sum_{j=1}^n e^{i\vec{q} \cdot (\vec{R} + \vec{d}_j)} = \sum_{j=1}^n e^{i\vec{q} \cdot \vec{d}_j} \sum_{\vec{R}} e^{i\vec{q} \cdot \vec{R}} \equiv S_{\vec{q}} S , \quad (8.152)$$

de sorte que l'effet du motif se traduit par l'introduction d'un facteur global d'amplitude $S_{\vec{K}}$:

$$S_{\vec{K}} = \sum_{j=1}^n e^{i\vec{K} \cdot \vec{d}_j} . \quad (8.153)$$

La quantité $S_{\vec{K}}$ est connue sous le nom de facteur de diffusion géométrique, puisqu'elle ne dépend que de la géométrie de la base associée à une maille primitive. En général $S_{\vec{K}} \in \mathbb{C}$.

A titre d'exemple, soit un réseau cubique centré, dont le volume de la maille primitive est $v = a^3/2$. On peut évidemment raisonner avec son réseau réciproque $\tilde{\mathcal{B}}$, qui est CFC. Il est toutefois presque plus commode de considérer le réseau \mathcal{B} comme un cubique simple muni de la base diatomique : un atome à l'origine du cube, un autre au centre du cube, de coordonnées $(\frac{a}{2}, \frac{a}{2}, \frac{a}{2})$. Ceci étant fait, on voit que le facteur de forme est :

$$S_{\vec{K}} = 1 + e^{i\frac{a}{2} \vec{K} \cdot (\vec{i} + \vec{j} + \vec{k})} . \quad (8.154)$$

Dans cette optique, le réseau réciproque est un cubique simple de côté $2\pi/a$, et on a :

$$\forall \vec{K} \in \tilde{\mathcal{B}}_C \quad \vec{K} = \frac{2\pi}{a} (n_1 \vec{i} + n_2 \vec{j} + n_3 \vec{k}) , \quad (8.155)$$

d'où :

$$\vec{K} \cdot (\vec{i} + \vec{j} + \vec{k}) = \frac{2\pi}{a} (n_1 + n_2 + n_3) . \quad (8.156)$$

Finalement, d'après (8.154) :

$$S_{\vec{K}} = 1 + e^{i\pi (n_1 + n_2 + n_3)} = \begin{cases} 0 & \text{si } n_1 + n_2 + n_3 \text{ est impair} \\ 2 & \text{si } n_1 + n_2 + n_3 \text{ est pair} \end{cases} . \quad (8.157)$$

Ceci constitue une sorte de règle de sélection – d'origine purement géométrique – aussi appelée règle d'extinction. Ici, $S_{\vec{K}}$ est réel.

Bien évidemment, le spectre de diffraction ne dépend pas de la convention choisie pour le raisonnement : on obtient les mêmes résultats physiques en raisonnant directement avec le CC sans base. Dans cette optique, le réseau réciproque est un CFC dont le côté⁵⁰ de la maille conventionnelle est $\frac{4\pi}{a}$, et tout vecteur réciproque de \vec{B}_{CFC} est de la forme :

$$\vec{K} = \frac{4\pi}{a} \left(m_1 \frac{\vec{i} + \vec{j}}{2} + m_2 \frac{\vec{j} + \vec{k}}{2} + m_3 \frac{\vec{k} + \vec{i}}{2} \right) = \frac{2\pi}{a} \left[(m_1 + m_3)\vec{i} + (m_1 + m_2)\vec{j} + (m_2 + m_3)\vec{k} \right]. \quad (8.158)$$

La comparaison des deux expressions de \vec{K} (8.155) et (8.158) montre que l'on a :

$$n_1 = m_1 + m_3, \quad n_2 = m_1 + m_2, \quad n_3 = m_2 + m_3, \quad (8.159)$$

de sorte que :

$$n_1 + n_2 + n_3 = 2m_1 + 2m_2 + 2m_3. \quad (8.160)$$

Cette somme est bien un nombre pair, comme le prescrit la règle de sélection additionnelle (8.157) obtenue quand on raisonne avec le réseau cubique simple muni de la base diatomique restituant le CC. Ceci montre bien que le choix d'une maille conventionnelle (par opposition à une maille primitive) impose la prise en compte du facteur géométrique associé à la base ainsi introduite *de facto*.

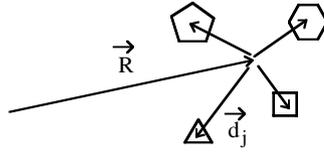


Figure 8.19: Notations utilisées dans (8.162).

Enfin, le plus souvent, les différents objets constituant la base ne sont pas tous identiques physiquement. Chaque onde réémise par un diffuseur contient un facteur d'amplitude spécifique à ce diffuseur, noté f et appelé facteur de forme atomique, qui dépend notamment du transfert de moment $\vec{q} = \vec{K}$. Si tous les objets de la base sont les mêmes, il s'agit d'un simple facteur commun à toutes les amplitudes partielles, et on peut l'absorber dans d'autres constantes multiplicatives. En revanche, si les constituants de la base sont distincts physiquement, le facteur de forme géométrique doit en tenir compte (voir figure 8.19) ; dans un tel cas, on définit un facteur de structure englobant les deux aspects⁵¹ :

$$S_{\vec{K}} = \sum_{j=1}^n f_j(\vec{K}) e^{i\vec{K} \cdot \vec{d}_j}. \quad (8.162)$$

La dépendance de f_j vis-à-vis de \vec{K} provient du déphasage calculé sur la densité $\rho(\vec{r})$ des particules qui se couplent à la sonde à l'intérieur d'un diffuseur. Il s'agit donc essentiellement d'une transformée de Fourier⁵² :

$$f_j(\vec{K}) = \int_{\mathbb{R}^3} d^3r e^{i\vec{K} \cdot \vec{r}} \rho(\vec{r}). \quad (8.163)$$

⁵⁰La maille élémentaire du cubique centré a pour volume $v = \frac{a^3}{2}$, celui de la maille primitive du réseau réciproque CFC est $\tilde{v} = \frac{(2\pi)^3}{v} = \frac{16\pi^3}{a^3}$. Comme la maille conventionnelle du CFC possède 4 atomes, son volume est $\frac{1}{4}\tilde{a}^3$ où \tilde{a} est le côté de la maille conventionnelle du CFC ; de $\frac{1}{4}\tilde{a}^3 = \frac{16\pi^3}{a^3}$, on déduit $\tilde{a} = \frac{4\pi}{a}$: la maille conventionnelle du réseau réciproque CFC est donc bien un cube de côté $\frac{4\pi}{a}$.

⁵¹Il s'agit de la définition usuelle. Ceci étant, il serait alors plus logique, dans le cas où les atomes du motifs sont identiques, de définir $S_{\vec{K}}$ (voir eq. (8.153)) comme :

$$S_{\vec{K}} = f(\vec{K}) \sum_{j=1}^n e^{i\vec{K} \cdot \vec{d}_j}, \quad (8.161)$$

où $f(\vec{K})$ est le facteur commun à tous les éléments du motif. Les définitions utilisées ici sont strictement celles de Ashcroft et Mermin [38].

⁵²Avec cette définition, le facteur de forme atomique f_j est sans dimension.

$f_j(\vec{K})$ est appelé *facteur de forme atomique* ; sa dépendance angulaire vis-à-vis de la direction de \vec{K} donne une idée précise de la forme du diffuseur (anisotropie), sa dépendance par rapport au module de \vec{K} donne des informations sur sa dimension (son extension spatiale). Avec les rayons X, $f_j(\vec{K})$ fournit la photo de la distribution électronique au sein de l'atome.

Pour des objets de dimensions atomiques et pour des rayons X, la quantité $\vec{K} \cdot \vec{r}$ est d'ordre 1 et la dépendance en \vec{K} ne doit être *a priori* négligée. Au contraire, pour la diffusion des neutrons sur la matière ordinaire, on pourra – en ce qui concerne l'interaction de Fermi – négliger cette dépendance de f_j : l'interaction est en effet alors essentiellement de contact, examinée à l'échelle de longueur de la matière condensée usuelle et seul $f_j(\vec{K} = 0)$ est pertinent.

Remarque

La diffraction de Bragg résulte de la diffusion (élastique) coopérative de diffuseurs répartis dans l'espace suivant un ordre parfait, un fait dont rend compte le calcul du déphasage spatial pour deux ondes planes $e^{i\vec{k}_i \cdot \vec{r}}$ et $e^{i\vec{k}_f \cdot \vec{r}}$. Ce calcul est manifestement d'inspiration classique.

Il est possible de donner une version quantique du même phénomène. Essentiellement, l'ingrédient que la description classique laisse sous le tapis est la *probabilité* pour le photon (ou le neutron) d'être diffusé de \vec{k}_i en \vec{k}_f . Cette probabilité peut se calculer par la théorie des perturbations et plus particulièrement par la règle d'or de Fermi ; ceci étant obtenu, on en déduit la section efficace de diffusion (finalement, c'est la quantité mesurée), qui fait ressortir la condition de von Laue. La présentation ci-dessous donne les idées principales pour la diffusion des neutrons.

D'après la règle d'or, la probabilité (par unité de temps) de la diffusion $\vec{k}_i \rightarrow \vec{k}_f$ est :

$$\mathcal{P}_{\vec{k}_i \rightarrow \vec{k}_f} = \frac{2\pi}{\hbar} \left| \int_{L^3} d^3r \psi_{\vec{k}_f}^*(\vec{r}) V(\vec{r}) \psi_{\vec{k}_i}(\vec{r}) \right|^2 \rho_{\vec{k}_f}(E_f) , \quad (8.164)$$

où $\rho_{\vec{k}}(E)$ est la densité d'états en énergie telle que :

$$\rho_{\vec{k}}(E)dE = \left(\frac{L}{2\pi} \right)^3 d^3k , \quad d^3k = k^2 dk d\Omega . \quad (8.165)$$

Pour des neutrons (masse M , vecteur d'onde \vec{k}) et avec des conditions cycliques dans un cube de côté L , (8.165) donne :

$$\rho_{\vec{k}}(E)dE = \left(\frac{L}{2\pi} \right)^3 \frac{Mk}{\hbar^2} d\Omega . \quad (8.166)$$

Dans (8.164), les fonctions d'onde initiale et finale d'un neutron seront ⁵³ les ondes planes $L^{-3/2} e^{i\vec{k}_s \cdot \vec{r}}$ avec $s = i, f$; $V(\vec{r})$ est l'interaction entre un neutron et un seul diffuseur : c'est par exemple le pseudo-potential de Fermi donné en (8.135).

Une fois connue la probabilité de transition par unité de temps, on peut écrire la section efficace élémentaire $d\sigma$:

$$d\sigma = \frac{\mathcal{P}_{\vec{k}_i \rightarrow \vec{k}_f}}{\text{courant incident}} . \quad (8.167)$$

Pour un neutron de vitesse $v = \hbar k_i / M$ dans le volume L^3 , le courant est $L^{-3} (\hbar k_i / M)$. En combinant (8.164), (8.166) et (8.167), on trouve :

$$\frac{d\sigma}{d\Omega} = \left(\frac{m}{2\pi\hbar^2} \right)^2 \left| \int_{L^3} d^3r e^{i\vec{q} \cdot \vec{r}} V(\vec{r}) \right|^2 , \quad \vec{q} = \vec{k}_i - \vec{k}_f . \quad (8.168)$$

Si maintenant on fait le choix selon Fermi, (8.135), pour un noyau situé au site \vec{R} , on trouve le résultat très simple :

$$\frac{d\sigma}{d\Omega} = \left| b e^{i\vec{q} \cdot \vec{R}} \right|^2 = b^2 . \quad (8.169)$$

⁵³En toute rigueur, l'état final est un état de diffusion du potentiel $V(\vec{r})$.

La longueur de diffusion de Fermi donne directement la section efficace, et joue exactement le même rôle que le rayon d'une sphère dure dans une collision classique. La section efficace totale est $\sigma = 4\pi b^2$.

Pour des diffuseurs tous identiques ordonnés suivant le réseau \mathcal{B} , $V(\vec{r})$ est :

$$V(\vec{r}) = \frac{2\pi\hbar^2}{M} b \sum_{\vec{R} \in \mathcal{B}} \delta(\vec{r} - \vec{R}) . \quad (8.170)$$

La section efficace est alors :

$$\frac{d\sigma}{d\Omega} = \left| b \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{q} \cdot \vec{R}} \right|^2 \equiv b^2 |S|^2 . \quad (8.171)$$

Pour un réseau avec une base formée de n noyaux de longueurs de diffusion b_j , on a :

$$\frac{d\sigma}{d\Omega} = \left| \sum_{\vec{R} \in \mathcal{B}} \sum_{j=1}^n b_j e^{i\vec{q} \cdot (\vec{R} + \vec{d}_j)} \right|^2 = \left| \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{q} \cdot \vec{R}} \right|^2 \left| \sum_{j=1}^n b_j e^{i\vec{q} \cdot \vec{d}_j} \right|^2 . \quad (8.172)$$

Le facteur de diffusion géométrique dû au motif est maintenant :

$$\mathcal{S}_{\vec{q}} = \sum_{j=1}^n b_j e^{i\vec{q} \cdot \vec{d}_j} , \quad (8.173)$$

et a ici la dimension d'une longueur ; les b_j jouent dans (8.173) le même rôle que les f_j dans (8.162) et n'apportent pas de dépendance en \vec{q} puisque le potentiel de Fermi est de contact : pour cette interaction, ce qui tient lieu de facteur de forme "atomique" est indépendant du transfert de moment.

Chapitre 9

Electrons dans un cristal

9.1 Préliminaires

Avant d'en venir à la description quantitative du mouvement d'une particule (un électron) dans un champ de force périodique dérivant d'un potentiel, il convient de donner quelques éléments justifiant la pertinence de cette problématique pour l'étude des solides cristallisés.

Un solide, ordonné ou non, est un système fort complexe, même si l'on s'en tient aux interactions élémentaires les plus simples, celle de Coulomb par exemple. En effet, un solide contient des noyaux et des électrons, une telle distinction n'étant pas d'ailleurs la plus naturelle : on a vu que, selon les cas et la nature physique du solide, il était plus judicieux de le considérer comme un assemblage d'ions (solides ioniques ou métalliques) ou de molécules (solides moléculaires ou covalents).

En tout état de cause, la toute première approximation est évidemment celle de Born et Oppenheimer. La séparation entre noyaux (ou cœurs) et électrons étant faite, on en vient en outre à distinguer, le cas échéant, plusieurs sortes d'électrons : d'une part ceux qui sont dans les couches atomiques profondes, quasiment insensibles à ce qui se passe lors de la formation du solide (électrons de cœur) et ceux de la couche externe (complète ou incomplète) qui sont au contraire aux premières loges pour encaisser les altérations survenant lors de la formation du solide (électrons de valence¹). On peut considérer que cette distinction repose essentiellement sur un simple critère énergétique : les électrons de valence ont une énergie de liaison comparable, par nature, à celle d'une liaison chimique (1 eV ou une fraction d'eV), alors que les électrons internes sont très fortement liés (pour un élément de masse atomique même modeste, l'énergie de la couche (1s) est très vite² de l'ordre du keV).

À l'issue de ce dégrossissage, et en se plaçant dans la première étape de l'approximation de Born et Oppenheimer³, on reste *a priori* avec des ions et les électrons de valence, ces derniers étant soumis au potentiel plus ou moins fort des premiers. En plus du potentiel périodique du réseau, les électrons sont évidemment en interaction mutuelle, au moins par leur répulsion électrostatique.

On se doute – il suffit de se souvenir du cas de l'atome d'hélium – qu'il s'agit, tel quel, d'un problème redoutable, généralement appelé Problème à N -corps, que l'on ne sait résoudre qu'approximativement, à l'exception

¹En outre, pour un métal, une partie des électrons de valence à l'état atomique se retrouvent presque libres dans le métal : ceux-ci sont alors appelés électrons de conduction, en opposition avec ceux qui restent liés aux ions, que l'on continue à appeler électrons de valence.

²Pour un métal un peu lourd, comme le tungstène ou le molybdène, l'éjection d'un électron 1s donne lieu à réarrangement interne produisant les raies X notées K_α , K_β , etc.

³La deuxième étape consiste à étudier le mouvement des cœurs, qui donne lieu aux vibrations du solide et, dans l'hypothèse des petites oscillations, conduit à des oscillateurs harmoniques dont la quantification introduit naturellement les phonons du réseau – voir chapitre 10.

de quelques modèles simples se bornant à singer la répulsion électronique⁴. Dans une approche élémentaire, suivie ici, on décide d'oublier purement et simplement la répulsion de Coulomb entre électrons, ce qui, en se souvenant à nouveau du cas de l'hélium, peut paraître très grossier. En réalité, et de façon assez surprenante, nombreux sont les solides où l'approximation subséquente permet d'expliquer un très grand nombre de propriétés (pas toutes, évidemment) et, en particulier, d'énoncer les critères permettant de savoir pourquoi il existe des isolants et des conducteurs. La justification de ce petit miracle n'est pas aisée, mais il est possible d'imaginer que, en raison de la forte densité d'électrons, des phénomènes d'écrantage viennent réduire considérablement la portée et l'importance de la répulsion de Coulomb⁵ entre des électrons qui, de surcroît, disposent d'un vaste espace (le solide) et sont, du fait de cette délocalisation, relativement éloignés les uns des autres. Pour ces raisons, l'interaction *effective* est loin de ressembler à la répulsion de Coulomb nue.

Tout ceci étant admis, on reste – en ce qui concerne la première étape de Born - Oppenheimer – avec quelques électrons presque indépendants par atome, chacun d'entre eux étant soumis à un potentiel statique périodique, celui des ions immobiles situés aux nœuds d'un réseau de Bravais. Les électrons étant déclarés indépendants, une étape préliminaire consiste à étudier le mouvement de l'un d'entre eux. L'état (mécanique ou thermodynamique) d'une assemblée de N électrons s'en déduira par simple produit tensoriel des espaces d'états à une particule.

En définitive, dans le cadre ainsi précisé, l'équation aux valeurs propres à résoudre est celle associée au Hamiltonien d'un unique électron dans un potentiel périodique :

$$H = \frac{\vec{p}^2}{2m} + V(\vec{r}) , \quad V(\vec{r} + \vec{R}) = V(\vec{r}) \quad \forall \vec{R} \in \mathcal{B} . \quad (9.1)$$

On se doute que ceci n'est que le point de départ de nouvelles approximations, puisque la classe de potentiels $V(\vec{r})$ exactement solubles est fort réduite⁶. Avant d'en venir à quelques schémas approximatifs de portée universelle (voir plus loin), il est important d'énoncer des résultats exacts, ne reposant que sur l'existence de la symétrie de translation discrète du réseau, traduite par l'invariance de l'énergie potentielle V dans toute translation d'un vecteur $\vec{R} \in \mathcal{B}$.

9.2 Théorème de Bloch

La symétrie de translation caractéristique de V entraîne que le Hamiltonien (9.1) commute avec tous les opérateurs de translation discrète $\mathcal{T}(\vec{R})$ associés chacun à une translation géométrique $T(\vec{R})$. De ce seul fait résulte le théorème de Bloch, dont l'énoncé est le suivant⁷ :

Pour un Hamiltonien périodique du type (9.1), toute fonction (état) propre $\psi(\vec{r})$ peut être mise sous la forme d'une onde plane modulée en amplitude par une fonction $u(\vec{r})$ ayant la symétrie de périodicité du réseau :

$$\forall \psi(\vec{r}), H\psi(\vec{r}) = \varepsilon\psi(\vec{r}), \quad \exists u(\vec{r}), \psi(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u(\vec{r}), \quad u(\vec{r}) = u(\vec{r} + \vec{R}) \quad \forall \vec{R} \in \mathcal{B} . \quad (9.2)$$

⁴Le plus célèbre d'entre eux est sans doute le modèle de Hubbard, dont la version sur réseau unidimensionnel (rigide) a été résolue en 1968 par Lieb et Wu. Si l'énoncé du problème est simple, sa résolution ne l'est pas.

⁵On a vu que la densité électronique est globalement caractérisée par la longueur r_s , définissant le rayon de la sphère occupée en moyenne par un électron. Cette longueur est de quelques unités atomiques pour un métal. Au contraire, quand $r_s \gg a_0$, les électrons sont quasiment immobiles, au sens où l'énergie cinétique est très petite devant l'énergie potentielle ; Wigner avait d'ailleurs prévu théoriquement que, dans cette situation, les électrons forment un réseau triangulaire (solide de Wigner). Ceci a été observé à la fin des années '80, par dépôt d'électrons à la surface de l'hélium liquide et par mesure de la modification des ondes capillaires.

⁶S'agissant d'exemple de potentiel périodique, celui de Kronig - Penney (un peigne de Dirac) est un grand classique.

⁷Des arguments similaires peuvent se transposer pour des Hamiltoniens périodiques en temps ; le même théorème est alors connu en Mathématiques sous le nom de *Théorème de Floquet*.

Par ailleurs, si les paramètres de maille du réseau tendent vers zéro (potentiel constant !), (9.2) dit que les fonctions propres d'une particule libre sont des ondes planes $e^{i\vec{k}\cdot\vec{r}}$: c'est l'invariance galiléenne.

Bien sûr, chaque fonction propre et chaque énergie propre doivent être indexées par un (des) nombre(s) quantique(s), collectivement notés n , associés aux observables qui, avec H , forment un ECO. En outre, l'apparition du vecteur \vec{k} – paramètre homogène à un vecteur d'onde dont la signification sera donnée plus loin –, conduit, quand il le faut, à écrire (9.2) plus précisément (notations traditionnelles) :

$$H\psi_{n\vec{k}}(\vec{r}) = \varepsilon_n(\vec{k}) \psi_{n\vec{k}}(\vec{r}) , \quad \psi_{n\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r}) , \quad u_{n\vec{k}}(\vec{r}) = u_{n\vec{k}}(\vec{r} + \vec{R}) . \quad (9.3)$$

L'affirmation contenue dans le théorème de Bloch peut d'ailleurs être énoncée autrement. En effet, on a la suite d'égalités :

$$\psi_{n\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot(\vec{r} + \vec{R})} u_{n\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot\vec{R}} e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r}) , \quad (9.4)$$

soit finalement :

$$\psi_{n\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot\vec{R}} \psi_{n\vec{k}}(\vec{r}) . \quad (9.5)$$

On utilisera selon les cas la forme (9.3) ou la forme (9.5).

Ce théorème peut être établi de plusieurs façons ; deux démonstrations sont données dans la suite, chacune d'entre elles ayant ses vertus et son utilité pour la référence ultérieure. Auparavant, il convient de noter que les deux relations (9.3) et (9.5) traduisent en fait une propriété de symétrie, ici de translation : le réseau se superpose à lui-même dans n'importe quelle translation de vecteur \vec{R} , de sorte que le module carré de la fonction d'onde doit être invariant :

$$|\psi_{n\vec{k}}(\vec{r} + \vec{R})|^2 = |\psi_{n\vec{k}}(\vec{r})|^2 . \quad (9.6)$$

Bien sûr – tout comme dans le cas de l'invariance de jauge –, la fonction elle-même n'est pas invariante et acquiert au plus un facteur de phase sous l'opération de symétrie⁸. C'est ce que dit, et précise, l'équation (9.5).

1^{ère} démonstration

Elle introduit explicitement les opérateurs $\mathcal{T}(\vec{R})$, définis précisément comme suit⁹ :

$$\forall \phi(\vec{r}), \mathcal{T}(\vec{R}) \phi(\vec{r}) = \phi(\vec{r} - \vec{R}) , \quad (9.8)$$

où l'on reconnaît la définition usuelle de l'opérateur agissant dans l'espace des états, associé à une transformation géométrique effectuée dans l'espace réel \mathbb{R}^3 , ici une translation. Comme il s'agit d'opérations de symétrie, tous les $\mathcal{T}(\vec{R})$ commutent avec H donné par (9.1) :

$$\forall \vec{R} \in \mathcal{B} : [H, \mathcal{T}(\vec{R})] = 0 ; \quad (9.9)$$

le sens physique de cette relation est clair : elle signifie qu'un état propre $\psi(\vec{r})$ de H étant donné, son transformé $\mathcal{T}(\vec{R})\psi(\vec{r})$ est aussi état propre avec la même énergie¹⁰ (la moindre des choses). En effet, (9.9) dit que, pour toute fonction quelconque ϕ , on a $H\mathcal{T}(\vec{R})\phi = \mathcal{T}(\vec{R})H\phi$; choisissons pour ϕ un état propre de H , ψ , satisfaisant $H\psi = \varepsilon\psi$. Il vient alors¹¹ :

$$H(\mathcal{T}\psi) = \mathcal{T}H\psi = \mathcal{T}\varepsilon\psi = \varepsilon(\mathcal{T}\psi) ; \quad (9.10)$$

⁸Ici la phase est constante (elle est indépendante de la position \vec{r}) – dans le cas d'une transformation de jauge, au contraire, la phase est variable dans l'espace. Elle garde toutefois toute sa pertinence : $\psi_{\vec{k}}$ et $\psi_{\vec{k}'}$ représentent deux états physiques distincts, tant que la différence $\vec{k} - \vec{k}'$ n'est pas un vecteur du réseau réciproque (voir plus loin)

⁹La définition (9.8) de la fonction transformée par l'opérateur de translation est la plus naturelle, et c'est elle qui est retenue pour la suite. On trouve parfois dans la littérature la définition "duale" alternative :

$$\mathcal{T}(\vec{R}) \phi(\vec{r}) = \phi(\vec{r} + \vec{R}) . \quad (9.7)$$

¹⁰On retiendra qu'un sous-espace dégénéré est *globalement* invariant dans les opérations de symétrie.

Pour un groupe de symétrie non abélien, les états propres de H ne sont pas forcément propres des opérateurs de symétrie : penser au groupe des rotations dans \mathbb{R}^3 dans le cas d'un champ central ; les harmoniques sphériques, propres du moment cinétique orbital, ne sont pas propres des opérateurs de rotation (L_{\pm} échange les Y_{lm} les unes dans les autres, et les L_u sont les générateurs des rotations). Toutefois, L_u ne modifie pas la valeur du nombre quantique l , puisque l'énergie ne change pas par rotation.

¹¹les parenthèses sont redondantes algébriquement, mais sont utiles pour faire comprendre les choses.

ceci montre que la fonction $\mathcal{T}\psi$ est propre avec la valeur propre ε , tout comme l'est ψ .

De toute évidence, et en s'appuyant sur la géométrie élémentaire, l'ensemble des $\mathcal{T}(\vec{R})$ peut être muni d'une structure de groupe, au moyen de la relation de définition qui s'impose :

$$\mathcal{T}(\vec{R})\mathcal{T}(\vec{R}') = \mathcal{T}(\vec{R} + \vec{R}') . \quad (9.11)$$

L'inverse est défini comme :

$$\mathcal{T}^{-1}(\vec{R}) = \mathcal{T}(-\vec{R}) . \quad (9.12)$$

Ce groupe a ici la puissance de \mathbb{N} (translations *discrètes*) ; il est commutatif (abélien), puisque la somme $\vec{R} + \vec{R}'$ est commutative :

$$\mathcal{T}(\vec{R})\mathcal{T}(\vec{R}') = \mathcal{T}(\vec{R} + \vec{R}') = \mathcal{T}(\vec{R}' + \vec{R}) = \mathcal{T}(\vec{R}')\mathcal{T}(\vec{R}) \iff [\mathcal{T}(\vec{R}), \mathcal{T}(\vec{R}')] = 0 \quad (9.13)$$

Il en résulte que tous les \mathcal{T} ont des vecteurs propres communs. Par ailleurs, les $\mathcal{T}(\vec{R})$ sont des opérateurs unitaires, puisqu'ils forment un sous-ensemble discret des translations continues ; parmi ces dernières existent les translations infinitésimales, qui ne sauraient être anti-unitaires¹². En conséquence :

$$\mathcal{T}^\dagger(\vec{R}) = \mathcal{T}^{-1}(\vec{R}) = \mathcal{T}(-\vec{R}) . \quad (9.14)$$

Comme H et $\mathcal{T}(\vec{R})$ commutent, ils ont des vecteurs propres en commun. Soit donc une fonction propre¹³ $\psi(\vec{r})$ telle que les deux équations suivantes sont satisfaites :

$$H\psi(\vec{r}) = \varepsilon\psi(\vec{r}) , \quad (9.15)$$

$$\mathcal{T}(\vec{R})\psi(\vec{r}) = \tau(\vec{R})\psi(\vec{r}) \quad \forall \vec{R} \in \mathcal{B} , \quad (9.16)$$

où, tout naturellement, la valeur propre τ de $\mathcal{T}(\vec{R})$ dépend de \vec{R} . L'équation (9.15) n'est pas explicitement utile pour la suite : elle rappelle seulement qu'une fonction telle que ψ , commune à H et aux $\mathcal{T}(\vec{R})$, *existe*.

En conséquence de (9.16) et (9.11), on a :

$$\mathcal{T}(\vec{R})\mathcal{T}(\vec{R}')\psi(\vec{r}) = \tau(\vec{R})\tau(\vec{R}')\psi(\vec{r}) \quad (9.17)$$

et :

$$\mathcal{T}(\vec{R} + \vec{R}')\psi(\vec{r}) = \tau(\vec{R} + \vec{R}')\psi(\vec{r}) . \quad (9.18)$$

Les équations (9.17) et (9.18) sont vraies $\forall \psi$ également propre de H , donc pour une base complète. Il en résulte :

$$\tau(\vec{R} + \vec{R}') = \tau(\vec{R})\tau(\vec{R}') . \quad (9.19)$$

τ est donc une fonction exponentielle – d'ailleurs, en vertu de l'unitarité des \mathcal{T} toutes leurs valeurs propres sont de la forme $e^{i \times \text{phase réelle}}$, ce que l'on choisit d'écrire plus précisément, pour la commodité :

$$\tau(\vec{R}) = e^{-i\Phi(\vec{R})} , \quad (9.20)$$

où Φ est une fonction à valeurs réelles ; (9.19) s'écrit alors :

$$\Phi(\vec{R} + \vec{R}') = \Phi(\vec{R}) + \Phi(\vec{R}') ; \quad (9.21)$$

cette relation dit que Φ est une forme linéaire de \vec{R} . $\forall \vec{R}$, une telle forme linéaire peut toujours s'écrire en produit scalaire¹⁴ ; $\Phi(\vec{R})$ est donc de la forme $\vec{k} \cdot \vec{R}$, où \vec{k} est quelconque et, en définitive :

$$\tau(\vec{R}) = e^{-i\vec{k} \cdot \vec{R}} . \quad (9.22)$$

Compte tenu de ce résultat, la relation (9.16) s'écrit maintenant, $\forall \psi$ propre de H et des \mathcal{T} :

$$\mathcal{T}(\vec{R})\psi(\vec{r}) = e^{-i\vec{k} \cdot \vec{R}}\psi(\vec{r}) \iff \psi(\vec{r} - \vec{R}) = e^{-i\vec{k} \cdot \vec{R}}\psi(\vec{r}) . \quad (9.23)$$

Ceci n'est autre que la formulation (9.5) du théorème de Bloch, avec le vecteur $-\vec{R}$.

¹²Un théorème dû à Wigner affirme que l'on peut associer à toute opération de symétrie un opérateur qui est soit linéaire unitaire (produit scalaire invariant), soit anti-linéaire unitaire (produit scalaire changé en son complexe conjugué).

En outre, les générateurs des translations sont proportionnels à l'impulsion \vec{p} , qui est hermitique. Pour une translation finie de vecteur \vec{r}_0 , on a $\mathcal{T}(\vec{r}_0) = e^{(1/i\hbar)\vec{p} \cdot \vec{r}_0}$. Le développement de l'exponentielle, avec $\vec{p} = -i\hbar \vec{\nabla}$ montre que pour toute fonction entière $f(\vec{r})$ on a $\mathcal{T}(\vec{r}_0)f(\vec{r}) = f(\vec{r} - \vec{r}_0)$, en conformité avec (9.8).

¹³Pour la simplicité de l'écriture, on omet les indices ($n\vec{k}$) apparaissant dans les énoncés (9.3) et (9.5) du théorème de Bloch.

¹⁴C'est ainsi, par exemple, que l'on peut définir l'espace dual d'un espace vectoriel.

2^{ème} démonstration

La démonstration précédente est en un sens cantonnée à l'espace réel. Au contraire, celle qui suit se place d'emblée dans l'espace réciproque, exploite les propriétés de la transformée de Fourier du potentiel périodique $V(\vec{r})$ et détermine en fait la fonction propre $\psi(\vec{r})$ par sa transformée de Fourier $\phi(\vec{q})$.

Ici arrive une petite difficulté technique : la fonction ψ est étendue sur tout le cristal (en principe infini) et donc n'a pas de transformée de Fourier au sens usuel. Pour contourner cette difficulté, on introduit provisoirement un cristal de taille finie, construit pour la commodité sur un (très grand) nombre de mailles primitives. L'espace direct étant dès lors *fini*, la transformée de Fourier de ψ est une intégrale sur un domaine fini (mais "grand"), de sorte que la question de son existence ne se pose pas. Tout naturellement, on a envie de prendre des conditions aux limites franches, en disant que ψ s'annule strictement en-dehors de la grande boîte (l'électron ne peut sortir de la boîte) ; en fait, ces conditions se révèlent peu commodes, techniquement parlant¹⁵. Pour cette raison, on choisit des conditions *cycliques*, dites de Born - von Karman, partant de l'idée que, le système physique étant par hypothèse (implicite) bien conditionné, le détail de ce qui se passe à la frontière de la grande boîte est sans importance physique (les éventuelles différences entre un jeu de conditions aux limites et un autre se traduisent par des termes s'annulant dans la limite d'un système infiniment étendu – l'analogie ici de la "limite thermodynamique").

Un jeu de vecteurs primitifs de \mathcal{B} , \vec{a}_i ($i = 1, 2, 3$), étant effectivement choisi, ces conditions identifient le point¹⁶ \vec{r} et le point $\vec{r} + N_i \vec{a}_i$ pour chaque direction de \mathcal{B} ($i = 1, 2, 3, N_i \gg 1$). Ces deux points étant confondus (ce sont donc finalement les mêmes, l'espace étant replié sur lui-même), il en résulte :

$$f(\vec{r} + N_i \vec{a}_i) = f(\vec{r}) \quad \forall \text{ fonction } f \quad (i = 1, 2, 3; N_i \gg 1) . \quad (9.25)$$

Ceci étant admis, on voit que le vecteur \vec{k} apparaissant dans (9.5) est maintenant quantifié à toute petite échelle. En effet, en prenant pour f une onde plane, on a :

$$e^{i\vec{k} \cdot (\vec{r} + N_i \vec{a}_i)} = e^{i\vec{k} \cdot \vec{r}} , \quad (9.26)$$

soit :

$$e^{iN_i \vec{k} \cdot \vec{a}_i} = 1 . \quad (9.27)$$

Cette relation contraint les valeurs de \vec{k} . En développant \vec{k} sur des vecteurs primitifs \vec{b}_j du réseau réciproque, $\vec{k} = \kappa_1 \vec{b}_1 + \kappa_2 \vec{b}_2 + \kappa_3 \vec{b}_3$, et avec $\vec{a}_i \cdot \vec{b}_j = 2\pi \delta_{ij}$, on voit que les composantes κ_j satisfont $e^{2\pi i N_j \kappa_j} = 1$, ce qui impose :

$$\kappa_j = \text{entier relatif} \times \frac{1}{N_j} . \quad (9.28)$$

Le vecteur \vec{k} admet donc¹⁷ la décomposition :

$$\vec{k} = \sum_{j=1}^3 \frac{m_j}{N_j} \vec{b}_j \quad (m_j \in \mathbb{Z}) . \quad (9.29)$$

Comme toujours, l'application de conditions aux limites produit la quantification (les composantes de \vec{k} sont des multiples entiers de quantités finies bien définies) ; le vecteur d'onde \vec{k} est maintenant quantifié suivant une

¹⁵Pour une particule enfermée dans une grande boîte mais libre par ailleurs, les fonctions propres sont des sinus (penser au puits infini à une dimension dont l'une des extrémités est à l'origine). Quand on replie l'espace à la Born - von Karman, les fonctions propres sont des ondes planes, beaucoup plus agréables à manipuler.

¹⁶À une dimension d'espace, le repliement transforme un segment ouvert de grande longueur en cercle de grand rayon.

¹⁷Noter que la cyclicité impose aussi :

$$\mathcal{T}(\vec{K} = N_i \vec{a}_i) = \mathbf{1} . \quad (9.24)$$

¹⁷À ce stade, \vec{k} décrit (presque) tout \mathbb{R}^3 (c'est-à-dire \mathbb{R}^3 discrétisé) – il n'y a pas d'autre restriction pour les k_j , donc pas de restriction sur les m_j relativement aux N_j . On verra par la suite que deux états caractérisés par \vec{k} et $\vec{k} + \vec{K}$ ($\vec{K} \in \mathcal{B}$) représentent en fait le même état physique, ce qui permettra d'épuiser tous les états physiques distincts en cantonnant \vec{k} à une maille primitive de \mathcal{B} , soit en se bornant (par exemple) à $1 \leq m_j \leq N_j$.

grille de dimensions linéaires $\frac{1}{N_j}$; chaque volume élémentaire est une image en miniature d'une maille primitive du réseau réciproque¹⁸ et a pour volume $\delta^3 k$:

$$\delta^3 k = \frac{1}{N_1 N_2 N_3} |\vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)| \equiv \frac{1}{N} |(\vec{b}_1, \vec{b}_2, \vec{b}_3)| = \frac{1}{N} \frac{(2\pi)^3}{v}, \quad (9.30)$$

v désignant toujours le volume d'une maille primitive de \mathcal{B} . En désignant¹⁹ par V le volume du cristal :

$$\delta^3 k = \frac{(2\pi)^3}{V}. \quad (9.31)$$

Le nombre des “gros points” ainsi définis situés à l'intérieur d'une maille primitive du réseau réciproque est égal à N ($= N_1 N_2 N_3$).

Maintenant, toute fonction $\psi(\vec{r})$ satisfaisant ces conditions cycliques peut être développée sur des ondes planes satisfaisant elles-mêmes ces conditions. On peut donc écrire *a priori* :

$$\psi(\vec{r}) = \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q} \cdot \vec{r}}, \quad (9.32)$$

où $\phi(\vec{q})$ est la composante de Fourier de $\psi(\vec{r})$ pour la valeur \vec{q} et où la sommation (discrète) court sur tous les \vec{q} de la forme :

$$\vec{q} = \sum_{j=1}^3 \frac{m_j}{N_j} \vec{b}_j \quad (m_j \in \mathbb{Z}) \iff e^{i N_i \vec{q} \cdot \vec{a}_i} = 1. \quad (9.33)$$

m_j étant dans tout \mathbb{Z} , \vec{q} n'est pas borné et décrit “tout” \mathbb{R}^3 (et pas seulement une maille de $\tilde{\mathcal{B}}$), cet espace étant simplement ponctué par les petits volumes résultant des conditions cycliques ; en quelque sorte, la somme $\sum_{\vec{q}}$ dans (9.32) est un précurseur de l'intégrale (c'est presque une intégrale²⁰). Il n'y a pas de doute que l'expression (9.32) fournit une fonction satisfaisant les conditions cycliques²¹ ; avec (9.33), on peut écrire en effet :

$$\psi(\vec{r} + N_i \vec{a}_i) = \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q} \cdot (\vec{r} + N_i \vec{a}_i)} = \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q} \cdot \vec{r}} = \psi(\vec{r}). \quad (9.34)$$

L'objectif final est de trouver $\psi(\vec{r})$: selon (9.32), il est équivalent de trouver ses composantes de Fourier $\phi(\vec{q})$. Pour cela, exploitons maintenant la périodicité de $V(\vec{r})$. De ce fait (et pour un potentiel $V(\vec{r})$ suffisamment régulier), le développement suivant existe :

$$V(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \mathcal{V}(\vec{K}) e^{-i\vec{K} \cdot \vec{r}}, \quad (9.35)$$

où la somme court maintenant sur le réseau réciproque. En effet, comme $e^{i\vec{K} \cdot (\vec{r} + \vec{R})} = e^{i\vec{K} \cdot \vec{r}}$, le second membre est bien l'expression d'une fonction périodique dans l'espace réel – ce n'est d'ailleurs rien d'autre qu'une série de Fourier spatiale. Il existe évidemment une relation inverse :

$$\mathcal{V}(\vec{K}) = \frac{1}{v} \int_{\text{maille élémentaire}} V(\vec{r}) e^{i\vec{K} \cdot \vec{r}} d^3 r, \quad (9.36)$$

l'intégration étant restreinte à une maille élémentaire de \mathcal{B} . En effet, on a :

$$\int_{\text{maille élémentaire}} V(\vec{r}) e^{i\vec{K} \cdot \vec{r}} d^3 r = \sum_{\vec{K}' \in \tilde{\mathcal{B}}} \mathcal{V}(\vec{K}') \int_{\text{maille élémentaire}} e^{i(\vec{K} - \vec{K}') \cdot \vec{r}} d^3 r. \quad (9.37)$$

¹⁸On peut dire que c'est un “gros point”.

¹⁹Télescopage de notations : ne pas confondre le volume du cristal avec le potentiel $V(\vec{r})$!

²⁰On peut même la considérer comme une somme de Darboux.

²¹Il est moins évident que toute fonction satisfaisant ces conditions peut ainsi être développée.

Par ailleurs $\vec{K} - \vec{K}'$ est de la forme $\sum_j m_j \vec{b}_j$ et $\vec{r} = \sum_i x_i \vec{a}_i$ avec $x_i \in [0, 1]$ puisque \vec{r} est restreint à une maille élémentaire ; l'intégrale dans (9.37) est donc :

$$\int_{\text{maille élémentaire}} e^{2\pi i \sum_j m_j x_j} d^3 r \quad (9.38)$$

et vaut 0 si dès qu'un seul m_j est différent de zéro, et v si tous les m_j sont nuls ; en définitive :

$$\int_{\text{maille élémentaire}} e^{i(\vec{K}-\vec{K}')\cdot\vec{r}} d^3 r = v \delta_{\vec{K} \vec{K}'} , \quad (9.39)$$

où $\delta_{\vec{K} \vec{K}'}$ est le symbole de Kronecker. Transposées aux variables espace - vecteur d'onde, les relations (9.35), (9.36) et (9.39) ne sont rien d'autre que ($n, n' \in \mathbb{Z}$) :

$$f(t) = \sum_{n \in \mathbb{Z}} f_n e^{-i n \omega t} , \quad f_n = \frac{\omega}{2\pi} \int_{\text{période}} f_n e^{i n \omega t} dt , \quad \int_{\text{période}} e^{i(n-n')\omega t} dt = 2\pi \delta_{nn'} , \quad (9.40)$$

pour le couple énergie - pulsation dans le cas d'une fonction de période $T = \frac{2\pi}{\omega}$. Notons qu'avec la définition (9.36), V et \mathcal{V} ont la même dimension physique, celle d'une énergie.

Avant de continuer, quelques remarques donnant lieu à des simplifications dans la suite :

- On peut toujours choisir

$$\mathcal{V}(\vec{K} = 0) = 0 , \quad (9.41)$$

puisque $V(\vec{r})$ est défini à une constante additive près. Adopter cette relation, c'est fixer à zéro la moyenne spatiale sur une maille primitive du potentiel périodique (cette convention sera utile pour l'approximation des électrons presque libres ; au stade actuel, elle n'est pas nécessaire).

- $V(\vec{r}) \in \mathbb{R} \iff \mathcal{V}(\vec{K}) = \mathcal{V}(-\vec{K})^*$
- si le cristal a la symétrie d'inversion d'espace ($V(-\vec{r}) = V(\vec{r})$), alors $\mathcal{V}(\vec{K})$ est aussi réel puisque l'intégrale $\int d^3 r V(\vec{r}) \sin(\vec{K}\cdot\vec{r})$, possédant un intégrand impair, est nulle. En pareil cas :

$$\mathcal{V}(-\vec{K}) = \mathcal{V}(\vec{K}) = \mathcal{V}(\vec{K})^* . \quad (9.42)$$

En reportant les développements (9.32) et (9.35) dans l'équation aux valeurs propres écrite en représentation- q :

$$\left[\frac{\vec{p}^2}{2m} + V(\vec{r}) \right] \psi(\vec{r}) = \varepsilon \psi(\vec{r}) , \quad (9.43)$$

on obtient deux types de termes ($\vec{p} = -i\hbar\vec{\nabla}_{\vec{r}}$) :

$$\frac{\vec{p}^2}{2m} \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q}\cdot\vec{r}} = -\frac{\hbar^2}{2m} \sum_{\vec{q}} (i\vec{q})^2 \phi(\vec{q}) e^{i\vec{q}\cdot\vec{r}} , \quad (9.44)$$

et :

$$V(\vec{r}) \sum_{\vec{q}} \phi(\vec{q}) e^{i\vec{q}\cdot\vec{r}} = \sum_{\vec{q}} \sum_{\vec{K}} \mathcal{V}(\vec{K}) \phi(\vec{q}) e^{i\vec{q}\cdot\vec{r}} e^{-i\vec{K}\cdot\vec{r}} = \sum_{\vec{q}'} \sum_{\vec{K} \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}) \phi(\vec{q}' + \vec{K}) e^{i\vec{q}'\cdot\vec{r}} . \quad (9.45)$$

En identifiant maintenant dans les deux membres de l'équation aux valeurs propres (9.43) le coefficient de $e^{i\vec{q}\cdot\vec{r}}$, on obtient :

$$+\frac{\hbar^2 \vec{q}^2}{2m} \phi(\vec{q}) + \sum_{\vec{K} \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}) \phi(\vec{q} + \vec{K}) = \varepsilon \phi(\vec{q}) \iff [\varepsilon - \varepsilon^{(0)}(\vec{q})] \phi(\vec{q}) = \sum_{\vec{K} \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}) \phi(\vec{q} + \vec{K}) . \quad (9.46)$$

où on a posé :

$$\varepsilon^{(0)}(\vec{q}) = \frac{\hbar^2 \vec{q}^2}{2m} . \quad (9.47)$$

Par inspection, on voit que le système linéaire²² (9.46) couple exclusivement entre elles les composantes $\phi(\vec{q})$ ne différant que d'un vecteur \vec{K} appartenant à $\tilde{\mathcal{B}}$. Autrement dit, \vec{q} étant confiné dans une maille primitive de $\tilde{\mathcal{B}}$, seuls les vecteurs de la forme $\vec{q} + \vec{K}$, \vec{K} quelconque²³ dans $\tilde{\mathcal{B}}$, apparaissent dans la sommation de l'expression (9.32). En désignant par \vec{k} une valeur particulière pour \vec{q} , on obtient une fonction ψ paramétrée par \vec{k} , que l'on note $\psi_{\vec{k}}(\vec{r})$ et dont le développement est de la forme :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}) e^{i(\vec{k} + \vec{K}) \cdot \vec{r}} = e^{i\vec{k} \cdot \vec{r}} \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}) e^{i\vec{K} \cdot \vec{r}} . \quad (9.48)$$

À la réflexion, il n'est pas surprenant que la somme sur \vec{q} ($= \vec{k}$) donnant le développement de Fourier de $\psi_{\vec{k}}$ ne contienne que des vecteurs d'onde équivalents par translation de $\tilde{\mathcal{B}}$. \vec{k} est un bon nombre quantique et H , ou l'un de ses avatars, ne peut avoir d'éléments de matrice entre deux états associés à deux \vec{k} inéquivalents, donnant deux valeurs propres distinctes pour un même (et quelconque) opérateur de translation.

Maintenant, la somme située le plus à droite dans (9.48) est visiblement une fonction de \vec{r} ayant la symétrie de translation du réseau direct, puisque $e^{i\vec{K} \cdot (\vec{R} + \vec{r})} = e^{i\vec{K} \cdot \vec{r}}$: elle n'est rien d'autre que la fonction $u_{\vec{k}}$ introduite dans la formulation (9.2) du théorème de Bloch ; (9.48) peut ainsi se récrire :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}}(\vec{r}) , \quad u_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}) e^{i\vec{K} \cdot \vec{r}} = u_{\vec{k}}(\vec{r} + \vec{R}) \quad \forall \vec{R} \in \mathcal{B} ; \quad (9.49)$$

ainsi s'achève la deuxième démonstration de ce théorème. Par ailleurs, l'écriture de droite montre que :

$$u_{\vec{k} + \vec{K}_0}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}_0 + \vec{K}) e^{i\vec{K} \cdot \vec{r}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}') e^{i(\vec{K}' - \vec{K}_0) \cdot \vec{r}} = e^{-i\vec{K}_0 \cdot \vec{r}} u_{\vec{k}}(\vec{r}) \quad \forall \vec{K}_0 \in \tilde{\mathcal{B}} . \quad (9.50)$$

9.3 Premières conséquences du théorème de Bloch

9.3.1 Première zone de Brillouin

Quand on adopte les conditions cycliques de Born - von Karman, les valeurs de \vec{k} sont de la forme (9.33) :

$$\vec{k} = \sum_{j=1}^3 \frac{n_j}{N_j} \vec{b}_j \quad (n_j \text{ entier}) . \quad (9.51)$$

Le cristal est donc de taille finie, boîte parallélépipédique construite sur les trois vecteurs $N_i \vec{a}_i$ ($N_i \gg 1$). Ceci étant, il n'y a que $N = N_1 N_2 N_3$ valeurs distinctes possibles pour \vec{k} . En effet, si n_j est plus grand que N_j , ceci veut dire qu'il existe deux entiers p_j et r_j tels que :

$$n_j = p_j N_j + r_j \quad (r_j < N_j) . \quad (9.52)$$

Reporté dans (9.51), ceci montre que \vec{k} est de la forme :

$$\vec{k} = \sum_{j=1}^3 \frac{p_j N_j + r_j}{N_j} \vec{b}_j = \sum_{j=1}^3 \left(p_j + \frac{r_j}{N_j} \right) \vec{b}_j \equiv \vec{K} + \vec{k}_1 , \quad (9.53)$$

\vec{K} désignant toujours un vecteur du réseau réciproque. Comparons maintenant $\psi_{\vec{k}}$ et $\psi_{\vec{k}_1}$. Comme $e^{i\vec{K} \cdot \vec{R}} = 1$, on a :

$$\psi_{\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k} \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) = e^{i(\vec{k}_1 + \vec{K}) \cdot \vec{R}} \psi_{\vec{k}}(\vec{R}) = e^{i\vec{k}_1 \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) . \quad (9.54)$$

²²Il s'agit bien d'un *système* : il existe une telle équation pour chaque valeur de \vec{q} choisie sur la petite grille plongée dans \mathbb{R}^3 définie plus haut.

²³à l'exception de $\vec{K} = 0$ en vertu de (9.41).

Par ailleurs, on a toujours :

$$\psi_{\vec{k}_1}(\vec{r} + \vec{R}) = e^{i\vec{k}_1 \cdot \vec{R}} \psi_{\vec{k}_1}(\vec{r}) , \quad (9.55)$$

de sorte que $\psi_{\vec{k}}$ et $\psi_{\vec{k}_1}$ satisfont la même relation de déphasage donnée par le théorème de Bloch : c'est ceci qui permet de les identifier. D'ailleurs, le développement (9.48) montre clairement que $\psi_{\vec{k}}$ et $\psi_{\vec{k} + \vec{K}_0}$ sont identiques ; il suffit de en effet faire un changement de variable muette de sommation pour montrer que les deux expressions de $\psi_{\vec{k}}$ et $\psi_{\vec{k} + \vec{K}_0}$ coïncident strictement :

$$\psi_{\vec{k} + \vec{K}_0}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}_0 + \vec{K}) e^{i(\vec{k} + \vec{K}_0 + \vec{K}) \cdot \vec{r}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}} \phi(\vec{k} + \vec{K}') e^{i(\vec{k} + \vec{K}') \cdot \vec{r}} \equiv \psi_{\vec{k}}(\vec{r}) . \quad (9.56)$$

Ceci se voit encore comme suit ; en vertu de (9.50), on a encore :

$$\psi_{\vec{k} + \vec{K}_0}(\vec{r}) = e^{i(\vec{k} + \vec{K}_0) \cdot \vec{r}} u_{\vec{k} + \vec{K}_0}(\vec{r}) = e^{i(\vec{k} + \vec{K}_0) \cdot \vec{r}} e^{-i\vec{K}_0 \cdot \vec{r}} u_{\vec{k}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}}(\vec{r}) \equiv \psi_{\vec{k}}(\vec{r}) . \quad (9.57)$$

Ainsi, $\psi_{\vec{k}}$ et $\psi_{\vec{k} + \vec{K}}$, $\vec{K} \in \tilde{\mathcal{B}}$ représentent un seul et même état physique : il n'y a pas lieu de distinguer ces deux fonctions²⁴. Cette propriété de périodicité de $\psi_{\vec{k}}$ se reporte évidemment sur la valeur propre associée $\varepsilon(\vec{k})$. On retiendra donc :

$$\psi_{\vec{k} + \vec{K}}(\vec{r}) = \psi_{\vec{k}}(\vec{r}) , \quad \varepsilon(\vec{k} + \vec{K}) = \varepsilon(\vec{k}) \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.58)$$

D'où le résultat important : le vecteur d'onde \vec{k} apparaissant dans les fonctions de Bloch peut être toujours choisi à l'intérieur d'une maille primitive du réseau réciproque. L'usage est de choisir la maille de Wigner - Seitz de $\tilde{\mathcal{B}}$, traditionnellement alors appelée 1^{ère} zone de Brillouin, en abrégé BZ1.

9.3.2 Équation pour la fonction $u_{\vec{k}}$

Toute fonction propre d'un Hamiltonien périodique dans l'espace peut donc être cherchée sous la forme d'une fonction de Bloch :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}}(\vec{r}) , \quad (9.59)$$

avec :

$$u_{\vec{k}}(\vec{r} + \vec{R}) = u_{\vec{k}}(\vec{r}) . \quad (9.60)$$

Tout revient donc à trouver la fonction $u_{\vec{k}}(\vec{r})$; il est facile d'écrire l'équation qu'elle doit satisfaire. Partant de :

$$\left[-\frac{\hbar^2}{2m} \Delta + V(\vec{r}) \right] \psi_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) \psi_{\vec{k}}(\vec{r}) , \quad (9.61)$$

et injectant la forme (9.59), on trouve :

$$-\frac{\hbar^2}{2m} \left[(i\vec{k})^2 u_{\vec{k}} + 2i\vec{k} \cdot \vec{\nabla} u_{\vec{k}} + \Delta u_{\vec{k}} \right] + V(\vec{r}) u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) u_{\vec{k}}(\vec{r}) , \quad (9.62)$$

soit :

$$\frac{\hbar^2}{2m} (-i\vec{\nabla} + \vec{k})^2 u_{\vec{k}}(\vec{r}) + V(\vec{r}) u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) u_{\vec{k}}(\vec{r}) , \quad (9.63)$$

ou encore, puisque $-i\vec{\nabla} = \hbar^{-1}\vec{p}$:

$$\frac{1}{2m} (\vec{p} + \hbar\vec{k})^2 u_{\vec{k}}(\vec{r}) + V(\vec{r}) u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) u_{\vec{k}}(\vec{r}) , \quad (9.64)$$

En vertu de la périodicité de $u_{\vec{k}}$, (9.60), il suffit de résoudre (9.63) à l'intérieur d'une maille primitive de \mathcal{B} . Comme toujours, il convient de préciser des conditions aux limites sur la frontière du domaine de résolution, faute de quoi, l'équation aux dérivées partielles possédant une infinité de solutions, le problème serait mal posé physiquement. On sait que toute fonction propre doit être continue et à dérivée continue tant que le potentiel

²⁴Tout comme pour une fonction 2π -périodique $f(\theta)$, les deux fonctions $f(\theta)$ et $f(\theta + n2\pi)$ ($n \in \mathbb{Z}$) sont en fait une seule et même fonction.

ne présente pas de saut d'amplitude infinie, ce que l'on suppose. Il en résulte qu'ici les conditions aux limites sont :

$$u_{\vec{k}} \text{ et } \vec{\nabla} u_{\vec{k}} \text{ continues sur la frontière de la maille primitive .} \quad (9.65)$$

Comme toujours, l'application de conditions aux limites produit la quantification des valeurs propres ε , désormais affublées de l'indice discret n , tout comme les fonctions propres. Une notation précise est donc :

$$H_u(\vec{r}, \vec{p}; \vec{k}) u_{n\vec{k}}(\vec{r}) = \varepsilon_n(\vec{k}) u_{n\vec{k}}(\vec{r}) , \quad (9.66)$$

avec, d'après (9.64) :

$$H_u(\vec{r}, \vec{p}; \vec{k}) = \frac{1}{2m} (\vec{p} + \hbar\vec{k})^2 + V(\vec{r}) . \quad (9.67)$$

Les valeurs propres $\varepsilon_n(\vec{k})$ dépendent évidemment de \vec{k} ; pour chaque valeur de n , l'ensemble des valeurs de \vec{k} formant un ensemble quasi-continu, $\varepsilon_n(\vec{k})$ varie quasi-continûment et définit une *bande d'énergie* s'étendant de $\inf_{\vec{k}} \varepsilon_n$ à $\sup_{\vec{k}} \varepsilon_n$.

La caractéristique essentielle de cette bande est sa loi de dispersion, définie par la variation de ε_n avec \vec{k} . L'existence de ces bandes d'énergie est clairement une *conséquence* directe de la symétrie de translation. Noter un fait majeur : la fonction $\varepsilon_n(\vec{k})$ est ce qu'elle est et, *a priori*, peut être fort différente de la fonction associée à l'électron libre $\varepsilon^{(0)}(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m}$. On verra que, même quand le potentiel de réseau $V(\vec{r})$ est petit, la fonction $\varepsilon_n(\vec{k})$ possède des propriétés qui fait la démarque *qualitativement* de $\varepsilon^{(0)}(\vec{k})$ (en un sens précisé ultérieurement, elle a des "trous").

Par ailleurs, comme on a été conduit à identifier deux états \vec{k} et $\vec{k} + \vec{K}$ ne différant que par un vecteur de $\tilde{\mathcal{B}}$, les valeurs de $\varepsilon_n(\vec{k})$ et de $\varepsilon_n(\vec{k} + \vec{K})$ coïncident ; $\psi_{\vec{k}}$ et $\psi_{\vec{k} + \vec{K}}$ sont les mêmes, les valeurs propres sont donc aussi les mêmes. Les fonctions $\varepsilon_n(\vec{k})$ sont donc invariantes dans toute translation du réseau réciproque :

$$\varepsilon_n(\vec{k} + \vec{K}) = \varepsilon_n(\vec{k}) \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.68)$$

9.3.3 Impulsion de l'électron dans le cristal

Le vecteur $\hbar\vec{k}$ apparaissant dans le théorème de Bloch ressemble indéniablement à une impulsion au sens usuel. Il ne s'agit cependant pas de la vraie impulsion de l'électron, quoique la terminologie traditionnelle l'appelle "impulsion de l'électron dans le cristal". En effet, la véritable impulsion, notée \vec{p} comme d'habitude, est représentée par l'opérateur $-i\hbar\vec{\nabla}$; ce n'est pas une constante du mouvement ($[H, \vec{p}] \neq 0$), puisque l'électron est soumis à une force donnée par l'opposé du gradient du potentiel périodique. Au contraire, \vec{k} est un bon "nombre" quantique, qui, en tant que tel, étiquette correctement les états propres d'un électron.

Plus précisément, l'équation de Heisenberg pour \vec{p} est²⁵ :

$$i\hbar \frac{d}{dt} \vec{p} = [\vec{p}, H] = -i\hbar \vec{\nabla} V \quad \iff \quad \frac{d}{dt} \vec{p} = [\vec{p}, H] = -\vec{\nabla} V . \quad (9.70)$$

D'ailleurs, $\psi_{n\vec{k}}$ n'est pas propre de \vec{p} :

$$\vec{p} \psi_{n\vec{k}} = -i\hbar \vec{\nabla} e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r}) = -i\hbar [i\vec{k} u_{n\vec{k}} + (\vec{\nabla} u_{n\vec{k}})] e^{i\vec{k}\cdot\vec{r}} = \hbar\vec{k} \psi_{n\vec{k}} - i\hbar e^{i\vec{k}\cdot\vec{r}} (\vec{\nabla} e^{-i\vec{k}\cdot\vec{r}} \psi_{n\vec{k}}) . \quad (9.71)$$

Le second terme n'a aucune raison d'être proportionnel à $\psi_{n\vec{k}}$ – il ne disparaît que si $u_{n\vec{k}}$ est constante, c'est-à-dire pour un électron *libre*. Enfin, comme le vecteur \vec{k} est finalement défini à un vecteur \vec{K} près (puisque, par exemple, $\varepsilon_n(\vec{k})$ a la symétrie de translation décrite par (9.68)), il existe un arbitraire pour \vec{k} ; à lui seul, cet arbitraire démontre le fait que $\hbar\vec{k}$ ne peut être une impulsion physique au sens usuel.

²⁵En laissant tomber l'indice H pour simplifier les notations ; avec les notations précises :

$$\frac{d}{dt} \vec{p}_H = -(\vec{\nabla} V)_H . \quad (9.69)$$

9.3.4 Vitesse d'un électron de Bloch

On appelle désormais électron de Bloch un électron décrit par une fonction d'onde de Bloch, soit $\psi_{n\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r})$. On va maintenant établir la relation suivante, remarquable et très simple, entre la vitesse moyenne d'un tel électron et l'énergie propre associée à $\psi_{n\vec{k}}$:

$$\langle \vec{v} \rangle_{n\vec{k}} \equiv \langle \psi_{n\vec{k}} | \frac{\vec{p}}{m} | \psi_{n\vec{k}} \rangle = \hbar^{-1} \vec{\nabla}_{\vec{k}} \varepsilon_n \equiv \vec{\nabla}_{\hbar\vec{k}} \varepsilon_n . \quad (9.72)$$

Pour établir ce résultat, on utilise le théorème de Hellmann - Feynman, qui stipule que pour un Hamiltonien quelconque $H(\lambda)$ dépendant d'un paramètre λ et avec :

$$H(\lambda) \psi(\lambda) = E(\lambda) \psi(\lambda) \quad \langle \psi(\lambda) | \psi(\lambda) \rangle = 1 , \quad (9.73)$$

on a :

$$\frac{dE}{d\lambda} = \langle \psi(\lambda) | \frac{\partial H}{\partial \lambda} | \psi(\lambda) \rangle . \quad (9.74)$$

En effet :

$$\begin{aligned} \frac{d}{d\lambda} \langle \psi(\lambda) | H(\lambda) | \psi(\lambda) \rangle &= \langle \frac{\partial \psi}{\partial \lambda} | H | \psi(\lambda) \rangle + \langle \psi(\lambda) | \frac{\partial H}{\partial \lambda} | \psi(\lambda) \rangle + \langle \psi(\lambda) | H(\lambda) | \frac{\partial \psi}{\partial \lambda} \rangle \\ &= E(\lambda) \left(\langle \frac{\partial \psi}{\partial \lambda} | \psi(\lambda) \rangle + \langle \psi(\lambda) | \frac{\partial \psi}{\partial \lambda} \rangle \right) + \langle \psi(\lambda) | \frac{\partial H}{\partial \lambda} | \psi(\lambda) \rangle . \end{aligned} \quad (9.75)$$

La parenthèse est nulle en vertu de la normalisation de ψ , et (9.74) en résulte.

Ici, le Hamiltonien pour la fonction $u_{\vec{k}}$ est H_u donné par l'expression (9.67) ; il dépend du vecteur \vec{k} . Le théorème de Hellmann - Feynman s'écrit ici :

$$\vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k}) = \langle u_{n\vec{k}} | \vec{\nabla}_{\vec{k}} H_u | u_{n\vec{k}} \rangle = \langle u_{n\vec{k}} | \frac{\hbar}{m} (\vec{p} + \hbar\vec{k}) | u_{n\vec{k}} \rangle = \frac{\hbar^2}{m} \langle u_{n\vec{k}} | -i\vec{\nabla}_{\vec{r}} + \vec{k} | u_{n\vec{k}} \rangle . \quad (9.76)$$

On a :

$$\begin{aligned} (-i\vec{\nabla}_{\vec{r}} + \vec{k}) u_{n\vec{k}} &= (-i\vec{\nabla}_{\vec{r}} + \vec{k}) e^{-i\vec{k}\cdot\vec{r}} \psi_{n\vec{k}}(\vec{r}) = (-i)(-i\vec{k}) e^{-i\vec{k}\cdot\vec{r}} \psi_{n\vec{k}} - i e^{-i\vec{k}\cdot\vec{r}} \vec{\nabla}_{\vec{r}} \psi_{n\vec{k}} + \vec{k} e^{-i\vec{k}\cdot\vec{r}} \psi_{n\vec{k}} \\ &= -i e^{-i\vec{k}\cdot\vec{r}} \vec{\nabla}_{\vec{r}} \psi_{n\vec{k}} . \end{aligned} \quad (9.77)$$

D'où :

$$\vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k}) = \frac{\hbar^2}{m} (-i) \langle u_{n\vec{k}} | e^{-i\vec{k}\cdot\vec{r}} \vec{\nabla}_{\vec{r}} | \psi_{n\vec{k}} \rangle = -i \frac{\hbar^2}{m} \langle \psi_{n\vec{k}} | \vec{\nabla}_{\vec{r}} | \psi_{n\vec{k}} \rangle = \frac{\hbar}{m} \langle \psi_{n\vec{k}} | \vec{p} | \psi_{n\vec{k}} \rangle , \quad (9.78)$$

soit la relation annoncée :

$$\langle \psi_{n\vec{k}} | \frac{\vec{p}}{m} | \psi_{n\vec{k}} \rangle = \frac{1}{\hbar} \vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k}) . \quad (9.79)$$

On remarque au passage l'analogie avec une équation de Hamilton :

$$\langle \vec{v} \rangle = \frac{\partial \varepsilon_n}{\partial(\hbar\vec{k})} \quad \longleftrightarrow \quad \dot{q} = \frac{\partial H}{\partial p} , \quad (9.80)$$

la pseudo-impulsion $\hbar\vec{k}$ jouant le rôle d'un moment conjugué et $\varepsilon_n(\vec{k})$ celui d'un Hamiltonien. Comme dans tout état propre – qui produit un état stationnaire par simple adjonction du facteur de phase $e^{-i\hbar^{-1}\varepsilon_n(\vec{k})t}$ –, la vitesse d'un électron de Bloch est constante dans le temps. L'hypothèse d'un réseau d'ions fixes est évidemment essentielle pour ce point : on devine que si l'on introduit les vibrations du réseau, les états de Bloch ne sont plus stationnaires et, alors, la vitesse moyenne d'un électron peut de fait varier – et d'ailleurs se dégrader – par collisions (création - annihilation de phonons).

9.3.5 Symétrie de la fonction $\varepsilon(\vec{k})$

Pour simplifier, on omet ci-dessous l'indice n et on considère exclusivement une bande d'énergie $\varepsilon(\vec{k})$. Il s'agit de montrer que l'on a :

$$\varepsilon(-\vec{k}) = \varepsilon(\vec{k}) , \quad \vec{\nabla}\varepsilon\left(\frac{\vec{K}}{2}\right) = 0 \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.81)$$

dans l'hypothèse où la fonction $\varepsilon(\vec{k})$ est partout deux fois dérivable.

La fonction de Bloch satisfait :

$$H \psi_{\vec{k}} = \varepsilon(\vec{k}) \psi_{\vec{k}} . \quad (9.82)$$

On sait que, lorsque l'on ne prend pas en compte les degrés de spin, l'opération de renversement du temps se traduit par la conjugaison complexe de la fonction d'onde. Par ailleurs, lorsque le Hamiltonien est invariant par renversement du temps, une fonction propre ψ et sa transformée ψ^* correspondent à la même valeur propre E . Dans le contexte actuel, ces résultats généraux s'expriment comme suit.

H étant supposé invariant par renversement du temps²⁶, une fonction de Bloch $\psi_{\vec{k}}(\vec{r})$ et sa complexe conjuguée $\psi_{\vec{k}}^*(\vec{r})$ ont la même énergie $\varepsilon(\vec{k})$. Par ailleurs, on a :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}}(\vec{r}) \quad \psi_{\vec{k}}^*(\vec{r}) = e^{-i\vec{k}\cdot\vec{r}} u_{\vec{k}}^*(\vec{r}) . \quad (9.83)$$

$\psi_{\vec{k}}^*(\vec{r})$ est donc une fonction de Bloch caractérisée par le préfacteur $e^{-i\vec{k}\cdot\vec{r}} \equiv e^{i(-\vec{k})\cdot\vec{r}}$, elle est donc du genre $u_{-\vec{k}}(\vec{r})e^{i(-\vec{k})\cdot\vec{r}}$. La comparaison avec (9.83) permet l'identification :

$$u_{-\vec{k}}(\vec{r}) = u_{\vec{k}}^*(\vec{r}) . \quad (9.84)$$

Il en résulte que :

$$\psi_{\vec{k}}^*(\vec{r}) = \psi_{-\vec{k}}(\vec{r}) . \quad (9.85)$$

Ces deux fonctions, transformées l'une de l'autre par renversement du temps, ont la même énergie ; mais $\psi_{-\vec{k}}(\vec{r})$ satisfait :

$$H \psi_{-\vec{k}} = \varepsilon(-\vec{k}) \psi_{-\vec{k}} . \quad (9.86)$$

d'où il résulte :

$$\varepsilon(-\vec{k}) = \varepsilon(\vec{k}) . \quad (9.87)$$

$\varepsilon_{\vec{k}}$ étant une fonction paire, sa dérivée est une fonction impaire :

$$\vec{\nabla}\varepsilon(-\vec{k}) = -\vec{\nabla}\varepsilon(\vec{k}) . \quad (9.88)$$

En outre, $\varepsilon(\vec{k})$ est une fonction périodique en \vec{K} , tout comme son gradient ; en particulier :

$$\varepsilon(\vec{k} - \vec{K}) = \varepsilon(\vec{k}) \quad \text{et} \quad \vec{\nabla}\varepsilon(\vec{k} - \vec{K}) = \vec{\nabla}\varepsilon(\vec{k}) . \quad (9.89)$$

Pour les vecteurs d'onde tels que $\vec{k} - \vec{K} = -\vec{k}$, soit $\vec{k} = \frac{\vec{K}}{2}$ (ce sont certains points remarquables de la frontière de la maille de Wigner - Seitz), les équations (9.88) et (9.89) donnent simultanément²⁷ :

$$\vec{\nabla}\varepsilon\left(-\frac{\vec{K}}{2}\right) = -\vec{\nabla}\varepsilon\left(\frac{\vec{K}}{2}\right) \quad \text{et} \quad \vec{\nabla}\varepsilon\left(-\frac{\vec{K}}{2}\right) = \vec{\nabla}\varepsilon\left(\frac{\vec{K}}{2}\right) \quad \iff \quad \vec{\nabla}\varepsilon\left(\frac{\vec{K}}{2}\right) = 0 . \quad (9.90)$$

$\varepsilon(\vec{k})$ est *extremum* en ces points. De plus, $\varepsilon(\vec{k})$ est évidemment stationnaire en $\vec{k} = 0$, par parité.

²⁶C'est toujours le cas quand $H = \frac{\vec{p}^2}{2m} + V(\vec{r})$, une expression qui exclut la présence d'un champ magnétique. L'invariance de H par renversement du temps vient du fait que H est *quadratique* par rapport à la vitesse. Au contraire, un champ magnétique implique forcément une boucle de courant (en termes plus savants : un moment cinétique, de spin ou orbital), qui s'inverse dans le changement $t \rightarrow -t$.

²⁷La discussion en cours exclut les cas où $\varepsilon(\vec{k})$ présente des points anguleux.

9.3.6 Surface de Fermi

La surface de Fermi est définie et existe dans tout modèle à électrons indépendants – son observation expérimentale dans un métal est d'ailleurs usuellement considérée comme un test certifiant la pertinence d'une telle approximation²⁸ –, quel que soit le potentiel à un corps ressenti par un électron dans le modèle effectif ou renormalisé considéré. Dans ces conditions, la construction de l'état fondamental d'un ensemble de N électrons s'effectue facilement, en conformité avec principe de Pauli : on prend les états électroniques un par un, par ordre d'énergie croissante, et on y place deux électrons de spins opposés, jusqu'à épuisement du stock d'électrons. L'énergie maximum ainsi atteinte s'appelle énergie de Fermi, ε_F . Si les niveaux sont repérés par un vecteur d'onde \vec{k} , les niveaux occupés de plus haute énergie sont tels que :

$$\varepsilon(\vec{k}) = \varepsilon_F . \quad (9.91)$$

Cette relation définit une surface dans l'espace \vec{k} , appelée surface de Fermi.

La situation la plus simple est celle d'électrons libres. Alors, la loi de dispersion est :

$$\varepsilon^{(0)}(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m} \quad (9.92)$$

de sorte que la surface de Fermi a pour équation :

$$\frac{\hbar^2 \vec{k}^2}{2m} = \varepsilon_F . \quad (9.93)$$

C'est donc une sphère, de rayon noté k_F . La valeur de k_F s'obtient comme suit. On admet des conditions cycliques, associant un petit cube de côté $2\pi/L$ à tout vecteur \vec{k} . En tenant compte du spin, et s'agissant de caser N électrons, on peut écrire :

$$2 \frac{(4\pi/3) k_F^3}{(2\pi/L)^3} = N \iff k_F = \left(3\pi^2 \frac{N}{L^3} \right)^{1/3} \equiv (3\pi^2 \rho)^{1/3} . \quad (9.94)$$

Pour un électron soumis à un potentiel périodique, la surface de Fermi est définie par considération des plus hautes bandes occupées. Dans le cas le plus simple, il n'y a qu'une seule plus haute bande occupée²⁹, alors la surface de Fermi est encore définie par l'équation (9.91), mais $\varepsilon(\vec{k})$ est maintenant la loi de dispersion relative à cette bande. Toutefois, comme il ne faut compter qu'une fois chaque état et que deux états ne différant que par un vecteur \vec{K} de $\tilde{\mathcal{B}}$ sont identiques, \vec{k} doit être restreint à une maille primitive de $\tilde{\mathcal{B}}$, la première zone de Brillouin par exemple. Il est donc utile de préciser :

$$\varepsilon(\vec{k}) = \varepsilon_F \quad \forall \vec{k} \in \text{BZ1} . \quad (9.95)$$

La valeur de ε_F est déterminée en intégrant la densité d'états $\mathcal{D}(E)$ (définie plus loin) jusqu'à ε_F et en écrivant que cette intégrale est égale au nombre d'électrons à caser. Précisément, pour un cristal à d dimensions, on écrit :

$$\int_0^{\varepsilon_F} dE 2 \left(\frac{L}{2\pi} \right)^d \int_{\mathbb{R}^d} d^d k \delta(\varepsilon(\vec{k}) - E) = N_e , \quad (9.96)$$

où N_e est le nombre d'électrons à répartir dans la bande ayant la dispersion $\varepsilon(\vec{k})$; de façon équivalente :

$$\int_0^{\varepsilon_F} dE \int_{\mathbb{R}^d} d^d k \delta(\varepsilon(\vec{k}) - E) = 2^{d-1} \pi^d \rho_e , \quad (9.97)$$

où ρ_e est la densité électronique.

Notons enfin que, compte tenu de la périodicité de $\varepsilon(\vec{k})$ dans $\tilde{\mathcal{B}}$, la surface de Fermi complète s'obtient par les translations de $\tilde{\mathcal{B}}$ de l'objet défini par (9.95) ; cette construction est utile pour l'analyse des problèmes de transport et permet, notamment, de distinguer le caractère ouvert ou fermé des orbites (trajectoires) semi-classiques.

²⁸ A tort ou à raison, l'image de particules indépendantes fait penser à un fluide ; un ensemble de tels fermions est appelé liquide de Fermi. On dit souvent que l'observation d'une surface de Fermi est la signature d'un liquide de Fermi.

²⁹ Quand il y a recouvrement en énergie entre les bandes, il peut y avoir plusieurs bandes incomplètement remplies chacune jusqu'à la même énergie ε_F (penser aux vases communicants).

9.4 Applications choisies du théorème de Bloch

Le théorème de Bloch énonce un résultat exact sur la nature des modes propres d'une particule dans un potentiel périodique³⁰, et montre notamment que les valeurs propres $\varepsilon_n(\vec{k})$ sont des fonctions (périodiques dans \vec{B}) du vecteur \vec{k} : ce sont les lois de dispersion de la particule dans le cristal, donnant la variation de l'énergie en fonction de l'impulsion, plus précisément de ce qui tient lieu d'impulsion – voir les mises en garde terminologiques du chapitre précédent. Il s'agit maintenant de présenter deux applications importantes de ce théorème, qui définissent deux cadres d'approximation adaptés à deux situations physiques extrêmes, identifiées par les caractères de ces lois de dispersion :

1. l'approximation des électrons presque libres, correspondant au cas où, dans le solide, les électrons sont seulement faiblement perturbés par le potentiel de réseau par comparaison à leur état libre. Il en résulte que la loi de dispersion ressemble à la parabole $\varepsilon^{(0)}(\vec{k}) = \hbar^2 \frac{\vec{k}^2}{2m}$; elle est donc fortement dispersée et on verra qu'elle est constituée d'arcs séparés par des petits *gaps*. Ceci correspond à des bandes larges et à une forte délocalisation spatiale des électrons (onde plane faiblement modulée). Cette situation prévaut dans les métaux et les semi-conducteurs
2. à l'opposé, l'approximation des liaisons fortes décrit le cas où chaque électron ressent, à l'ordre zéro, le potentiel localisé d'un atome (ou ion) du réseau. À cet ordre, le spectre d'énergie est celui d'un atome : des niveaux discrets, bien séparés. En fait, chaque électron peut se déplacer d'un atome à l'autre, ce qui donne une petite dispersion aux niveaux issus des niveaux atomiques. Il en résulte que les bandes d'énergie sont étroites et séparées par des grands *gaps*. Cette situation – lorsque la plus haute bande est pleine – est typique des isolants.

9.4.1 L'approximation des électrons presque libres

On adopte donc ici l'hypothèse suivant laquelle le potentiel de réseau $V(\vec{r})$ constitue une petite perturbation sur les états propres libres :

$$\psi_{\vec{k}}^{(0)}(\vec{r}) = C e^{i\vec{k}\cdot\vec{r}} \quad \varepsilon^{(0)}(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m}, \quad (9.98)$$

ce qui autorise d'emblée à envisager un traitement de type perturbatif, avec les précautions et les hypothèses usuelles.

Ces états propres à l'ordre zéro sont qualitativement différents des vrais états propres de l'électron dans le cristal, puisque ceux-ci doivent avoir la symétrie exigée par le théorème de Bloch et ses conséquences en raison de la symétrie de translation discrète – ce n'est pas le cas des états libres (9.98), qui sont caractéristiques d'une symétrie de translation continue, celle d'une particule libre (une onde plane telle que (9.98) est formellement une fonction de Bloch – $u_{\vec{k}}$ étant une constante (ce qui est bien une fonction périodique. . .) –, mais \vec{k} et $\vec{k} + \vec{K}$ ne représentent pas le même état physique, au contraire des “vrais” états de Bloch) ; d'ailleurs, dans (9.98), $\hbar\vec{k}$ est visiblement la *vraie* impulsion de l'électron.

Pour tout dire, on rencontre ici un bon exemple du fait que la symétrie fonctionne par tout ou rien et que des considérations *quantitatives* – du genre de celles que l'on énonce pour mettre en évidence un petit paramètre par exemple – appartiennent à un autre plan. Que le potentiel de réseau soit faible ou intense ne change rien au fait que la symétrie de translation devient discrète même si $V(\vec{r})$ est infiniment petit³¹. Il s'agit bien d'une *brisure* de symétrie, puisque la symétrie est devenue discrète sous l'effet du réseau de pas fini. Le groupe des translations discrètes est visiblement un sous-groupe des translations continues ; on retrouve ce dernier en faisant tendre vers zéro le pas du réseau.

³⁰Pour une discussion détaillée du potentiel périodique, voir [8], complément D_{III} et [7], ch. 8.

³¹On a toujours intérêt à faire jouer la symétrie le plus tôt possible. Notamment, l'introduction explicite de vecteurs ayant la symétrie de la perturbation produit des matrices diagonales par blocs. Il arrive même, quand la brisure de symétrie est assez forte, qu'à l'intérieur d'un sous-espace dégénéré, la projection de l'opérateur de perturbation dans ce sous-espace est représentée par une matrice complètement diagonale ; la simple lecture de ces éléments diagonaux donne alors immédiatement, sans aucun calcul supplémentaire, les corrections d'énergie au 1^{er} ordre.

Comme on l'a vu précédemment, l'existence du potentiel de réseau entraîne que les fonctions propres et les valeurs propres sont périodiques dans $\tilde{\mathcal{B}}$:

$$\psi_{n\vec{k}-\vec{K}}(\vec{r}) = \psi_{n\vec{k}}(\vec{r}) , \quad \varepsilon_n(\vec{k}-\vec{K}) = \varepsilon_n(\vec{k}) \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.99)$$

Ceci signifie qu'il suffit de les connaître à l'intérieur d'une maille primitive de $\tilde{\mathcal{B}}$ et, dans la suite, on pourra toujours, si besoin est, choisir le vecteur \vec{k} dans la première zone de Brillouin³². Compte tenu de l'hypothèse sur la petitesse de V , on va utiliser une méthode de perturbation. Il convient donc de préciser complètement le problème, tel qu'il se pose à l'ordre zéro, sans hésiter d'ailleurs à le reformuler dans la perspective du but recherché. Par ailleurs, on sait que la technique de perturbation, avant d'être mise en œuvre, exige l'analyse préalable de la dégénérescence du problème non-perturbé.

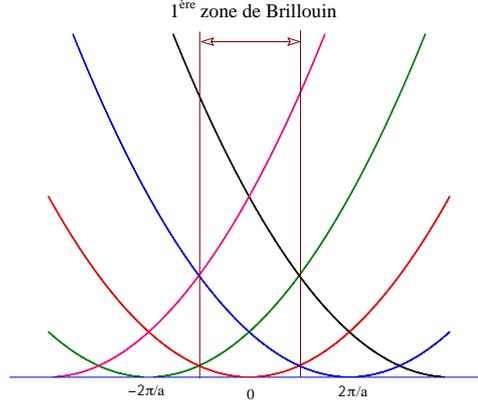


Figure 9.1: Schéma en zones réduites pour un réseau \mathcal{B} unidimensionnel de pas a . Les différentes paraboles représentent les $\varepsilon_{\vec{K}}^{(0)}(\vec{k})$ (voir (9.101)). La première zone de Brillouin s'étend entre $\pm \frac{\pi}{a}$.

Les solutions exactes sont périodiques dans $\tilde{\mathcal{B}}$, ainsi que l'exprime (9.99). D'un autre côté, comme V reste en toute hypothèse petit, on s'attend nécessairement à ce que l'on ait toujours :

$$\varepsilon(\vec{k}) \approx \varepsilon^{(0)}(\vec{k}) . \quad (9.100)$$

La bonne façon de formuler le problème à l'ordre zéro est donc de prolonger par périodisation, conformément à (9.99), tous les arcs de la parabole libre, faisant dès lors de la loi de dispersion libre une fonction elle aussi infiniment multivaluée. Ceci se fait en translatant tous les arcs de la parabole libre qui sont en-dehors de cette zone, et on obtient alors le schéma dit en zones réduites (voir figure 9.1). Dans cette représentation, $\forall \vec{k} \in \text{BZ1}$, l'énergie d'un électron en fonction de \vec{k} est infiniment multivaluée³³. Avec la même notation que précédemment, $\varepsilon^{(0)}(\vec{k})$, le prolongement périodique revient donc à définir une infinité de lois paraboliques $\varepsilon_{\vec{K}}^{(0)}$ telles que :

$$\varepsilon_{\vec{K}}^{(0)}(\vec{k}) = \varepsilon^{(0)}(\vec{k}-\vec{K}) \equiv \frac{\hbar^2}{2m}(\vec{k}-\vec{K})^2 \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.101)$$

Il y aura dégénérescence stricte à chaque fois que ces paraboles se croisent, c'est-à-dire lorsque :

$$\varepsilon_{\vec{K}_1}^{(0)}(\vec{k}) = \varepsilon_{\vec{K}_2}^{(0)}(\vec{k}) \iff \varepsilon^{(0)}(\vec{k}-\vec{K}_1) = \varepsilon^{(0)}(\vec{k}) , \quad (9.102)$$

où $\vec{K} = \vec{K}_2 - \vec{K}_1$ est encore un vecteur de $\tilde{\mathcal{B}}$. Comme $\varepsilon^{(0)}(\vec{k})$ ne dépend que du module des vecteurs d'onde, cette égalité (9.102) est satisfaite à chaque fois que³⁴ :

$$\|\vec{k}\| = \|\vec{k}-\vec{K}\| = \|\vec{K}-\vec{k}\| . \quad (9.103)$$

³²Dès qu'il existe un potentiel périodique, \vec{k} n'est défini physiquement que *modulo* $\vec{K} \in \tilde{\mathcal{B}}$, tout comme un angle est défini à 2π près ; on peut dire que l'équivalent de BZ1 est alors le segment $[-\pi, +\pi]$ pour une fonction 2π -périodique.

³³On peut se demander ce que devient la périodisation de $\varepsilon_{\vec{K}}^{(0)}$ quand on en revient au cas strictement libre. De fait, la limite de l'électron libre se retrouve comme suit : il suffit d'imaginer que l'on fait tendre vers zéro la maille primitive du réseau, ce qui à la limite produit un potentiel $V(\vec{r})$ prenant la même valeur *partout*. Dans cette limite, la longueur de tout vecteur $\vec{K} \in \tilde{\mathcal{B}}$ tend vers l'infini et on récupère *in fine* l'unique parabole qui est la limite de l'arc de plus basse énergie dans le schéma en zones réduites.

³⁴En l'absence de champ magnétique, $\varepsilon^{(0)}$ est aussi une fonction paire de \vec{k} ; toutefois, il n'y a pas lieu de considérer la dégénérescence $\vec{k} \rightarrow -\vec{k}$ correspondante : \vec{k} est un bon nombre quantique et les éléments de matrice $\langle \psi_{\vec{k}} | H | \psi_{-\vec{k}} \rangle$ sont nuls.

On peut dire les choses autrement. Pour un électron libre, la condition de dégénérescence est :

$$\varepsilon^{(0)}(\vec{k}) = \varepsilon^{(0)}(\vec{k}') . \quad (9.104)$$

Cette égalité est vraie si \vec{k} et \vec{k}' ont même module, en particulier si $\vec{k} = \vec{k}'$. Mais, dès qu'un potentiel périodique est présent, l'égalité des vecteurs d'onde – ou l'égalité des modules – doit être comprise *modulo* \vec{K} . Dans cette optique, (9.104) est équivalente à (9.103).

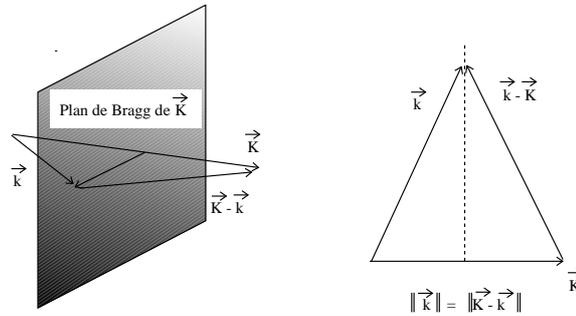


Figure 9.2: Schéma de définition du plan de Bragg d'un vecteur $\vec{K} \in \tilde{\mathcal{B}}$.

L'équation (9.103) signifie que, \vec{K} étant fixé, l'extrémité de \vec{k} est dans le plan médiateur de \vec{K} , autrement dit le plan de Bragg associé à \vec{K} . Ainsi, il apparaît une dégénérescence stricte apparaît lorsque \vec{k} appartient à un plan de Bragg³⁵ ; en pareil cas, on pourra toujours poser :

$$\vec{k} = \frac{\vec{K}}{2} + \vec{q} , \quad \vec{q} \perp \vec{K} . \quad (9.105)$$

Soit alors un vecteur primitif \vec{b}_0 définissant une certaine direction de $\tilde{\mathcal{B}}$, on peut distinguer deux cas :

- ou bien $\vec{K} = \text{entier pair} \times \vec{b}_0 \equiv 2n_0\vec{b}_0$. Dans ces conditions $\frac{\vec{K}}{2} = \vec{K}_0 \in \tilde{\mathcal{B}}$ et, après la bonne translation, on se retrouve à l'origine de $\tilde{\mathcal{B}}$. On parle alors de dégénérescence en centre de zone
- ou bien $\vec{K} = \text{entier impair} \times \vec{b}_0 \equiv (2n_0 + 1)\vec{b}_0$. Alors, $\frac{\vec{K}}{2} = \vec{K}_0 + \frac{1}{2}\vec{b}_0 \notin \tilde{\mathcal{B}}$ et la dégénérescence stricte se produit en un certain point de la frontière de la maille de Wigner - Seitz³⁶. On parle alors de dégénérescence en bord de zone.

Comme le paramètre \vec{k} varie continûment³⁷, il ne serait pas satisfaisant d'utiliser presque partout dans la zone la théorie de perturbation relative au cas non-dégénéré, sauf en certains points isolés où on appliquerait le formalisme du cas dégénéré ; une telle procédure conduirait d'ailleurs à des difficultés techniques puisque les dénominateurs d'énergie peuvent devenir graduellement de plus en plus petits (voir aussi (9.106)). Il faut donc construire un schéma de calcul plus subtil que celui énoncé par la théorie des perturbations "académique"³⁸.

On sait que le critère convenable pour qu'une théorie de perturbation fonctionne bien est de rassembler – en les traitant sur un même pied d'égalité – tous les états d'ordre zéro $\psi_n^{(0)}$ dont les énergies $E_n^{(0)}$ ont des différences au plus comparables aux éléments de matrice de la perturbation entre ces états. Ici, le potentiel de réseau (la perturbation) est supposé faible, c'est-à-dire que ses éléments de matrice sont réputés petits par

³⁵sans préjuger d'autres coïncidences par ailleurs, résultant d'une haute symétrie ponctuelle.

³⁶définie comme le polyèdre dont les faces sont les plans de Bragg de tous les vecteurs réciproques joignant l'origine de $\tilde{\mathcal{B}}$ aux nœuds premiers voisins et communément appelée 1^{ère} zone de Brillouin. L'intérieur de la 1^{ère} zone de Brillouin est l'ensemble des points de $\tilde{\mathcal{B}}$ qui sont plus près du nœud origine $\vec{K} = 0$ que de tous les autres nœuds (c'est la maille de Wigner - Seitz du réseau réciproque).

³⁷À une petite quantification près imposée par les conditions cycliques.

³⁸On rencontre une situation analogue pour un atome soumis à un champ magnétique \mathcal{B} ; alors les valeurs propres dépendent continûment de \mathcal{B} et le traitement de perturbation ordinaire doit être redéfini en présence d'un croisement de niveaux : on ne peut pas s'en tenir à la stricte recette énoncée par la théorie de perturbations dans sa version élémentaire.

rapport à une énergie typique ; compte tenu des paramètres physiques disponibles, cette énergie ne peut être que³⁹ $\frac{\hbar^2}{ma^2}$. Au total, le bon critère s'énonce par une condition du type :

$$|\varepsilon_{\vec{K}}^{(0)}(\vec{k}) - \varepsilon_{\vec{K}'}^{(0)}(\vec{k})| \lesssim |V| \ll \frac{\hbar^2}{ma^2} , \quad (9.106)$$

et le traitement de perturbation doit considérer *simultanément* tous les états d'ordre zéro satisfaisant (9.106).

V étant suffisamment faible, les arguments précédents permettent de délimiter deux situations, \vec{k} étant dans BZ1 :

1. \vec{k} est tel que toutes les branches $\varepsilon_{\vec{K}}^{(0)}(\vec{k})$ sont bien isolées les unes des autres. Alors, une théorie de perturbation non-dégénérée est acceptable
2. \vec{k} est tel que plusieurs branches $\varepsilon_{\vec{K}_i}^{(0)}(\vec{k})$ se trouvent au voisinage les unes des autres, l'étendue du voisinage se mesurant par l'ordre de grandeur du potentiel périodique conformément à (9.106). Notamment, ceci se produit, comme on vient de le voir, lorsque \vec{k} est près du centre ou au voisinage du plan de Bragg de $\vec{K} = \vec{K}_i - \vec{K}_j$ (par exemple sur la frontière de BZ1, quand $\vec{K}_1 = 0$, et \vec{K}_2 joint l'origine à un nœud premier voisin dans $\tilde{\mathcal{B}}$, par exemple). S'il y a g telles branches, le moins que l'on puisse faire est alors de considérer *simultanément* les g fonctions $\psi_{\vec{k}-\vec{K}_i}^{(0)}$, ou plus précisément le sous-espace $g \times g$ qu'elles engendrent. S'y prendre ainsi c'est finalement étendre la recette de perturbation dégénérée dans le *voisinage* de la dégénérescence $\|\vec{k} - \vec{K}_i\| = \|\vec{k} - \vec{K}_j\|$ et non plus strictement au point de dégénérescence.

Dans le cas le plus simple, on a $g = 2$; c'est la situation lorsque \vec{k} est non seulement dans le plan de $\vec{K} = \vec{K}_i - \vec{K}_j$ mais encore pas trop éloigné de la direction du vecteur \vec{K} lui-même. Si l'écart grandit, \vec{k} peut se rapprocher de la droite d'intersection de deux plans de Bragg et alors visiblement, $g > 2$.

Examinons maintenant comment les deux cas se présentent naturellement, en revenant au problème exact et en le représentant sur une base d'ondes planes (les fonctions à l'ordre zéro) – ceci revient à se placer dans l'espace de Fourier. Une fonction d'onde de Bloch s'écrit :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}}(\vec{r}) , \quad (9.107)$$

et comme la fonction $u_{\vec{k}}(\vec{r})$ a la périodicité du réseau, on peut la décomposer en série de Fourier sur le réseau réciproque :

$$u_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} U_{\vec{k}}(\vec{K}) e^{-i\vec{K}\cdot\vec{r}} . \quad (9.108)$$

Il en résulte que l'on peut écrire :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} U_{\vec{k}}(\vec{K}) e^{i(\vec{k}-\vec{K})\cdot\vec{r}} . \quad (9.109)$$

On sait que deux états de Bloch \vec{k} et $\vec{k} + \vec{K}_0$, $\vec{K}_0 \in \tilde{\mathcal{B}}$, sont en fait un seul et même état, d'où il résulte que la fonction $U_{\vec{k}}(\vec{K})$ ne dépend que de la différence $\vec{k} - \vec{K}$. En effet, (9.109) où on remplace \vec{k} par $\vec{k} + \vec{K}_0$ donne :

$$\psi_{\vec{k}+\vec{K}_0}(\vec{r}) = \sum_{\vec{K} \in \tilde{\mathcal{B}}} U_{\vec{k}+\vec{K}_0}(\vec{K}) e^{i(\vec{k}+\vec{K}_0-\vec{K})\cdot\vec{r}} ; \quad (9.110)$$

on fait le changement de variable muette $\vec{K} - \vec{K}_0 = \vec{K}'$ dans la sommation, d'où :

$$\psi_{\vec{k}+\vec{K}_0}(\vec{r}) = \sum_{\vec{K}' \in \tilde{\mathcal{B}}} U_{\vec{k}+\vec{K}_0}(\vec{K}' + \vec{K}_0) e^{i(\vec{k}-\vec{K}')\cdot\vec{r}} \equiv \sum_{\vec{K} \in \tilde{\mathcal{B}}} U_{\vec{k}+\vec{K}_0}(\vec{K} + \vec{K}_0) e^{i(\vec{k}-\vec{K})\cdot\vec{r}} ; \quad (9.111)$$

comme $\psi_{\vec{k}} = \psi_{\vec{k}+\vec{K}_0}$, l'identification terme à terme de (9.109) et (9.111) donne :

$$U_{\vec{k}+\vec{K}_0}(\vec{K} + \vec{K}_0) = U_{\vec{k}}(\vec{K}) . \quad (9.112)$$

³⁹ a est l'ordre de grandeur de la dimension linéaire d'une maille primitive.

Ceci signifie que la fonction $U_{\vec{k}}(\vec{K})$ est invariante dans toute translation de $\vec{\mathcal{B}}$ et ne dépend donc que de $\vec{k} - \vec{K}$. En posant désormais $U_{\vec{k}}(\vec{K}) = c_{\vec{k}-\vec{K}}$, (9.109) prend la forme :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{K} \in \vec{\mathcal{B}}} c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} . \quad (9.113)$$

$\psi_{\vec{k}}(\vec{r})$ est ainsi un développement en ondes planes, chacune de ces dernières satisfaisant le théorème de Bloch – une fonction constante dans l'espace est évidemment périodique en \mathcal{B} – et ses conséquences (9.99). En termes de *kets*, on peut écrire :

$$|\psi_{\vec{k}}\rangle = \sum_{\vec{K} \in \vec{\mathcal{B}}} c_{\vec{k}-\vec{K}} |f_{\vec{k}-\vec{K}}\rangle , \quad (9.114)$$

où :

$$\langle \vec{r} | f_{\vec{k}-\vec{K}} \rangle = e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} \equiv \psi_{\vec{k}-\vec{K}}^{(0)}(\vec{r}) \quad (9.115)$$

sont les ondes planes non-perturbées, solutions du problème à l'ordre zéro. Les $c_{\vec{k}-\vec{K}}$ sont donc les coefficients cherchés, que l'on imagine – dans un traitement perturbatif – être des séries entières en V .

Par ailleurs, la fonction $\psi_{\vec{k}}$ satisfait :

$$-\frac{\hbar^2}{2m} \Delta \psi_{\vec{k}}(\vec{r}) + V(\vec{r}) \psi_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k}) \psi_{\vec{k}}(\vec{r}) . \quad (9.116)$$

D'autre part, $\mathcal{V}(\vec{K})$ désignant toujours les composantes de Fourier⁴⁰ du potentiel périodique $V(\vec{r})$:

$$V(\vec{r}) = \sum_{\vec{K} \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}) e^{-i\vec{K} \cdot \vec{r}} \iff \mathcal{V}(\vec{K}) = \frac{1}{v} \int_{\text{maille élémentaire de } \mathcal{B}} d^3r V(\vec{r}) e^{i\vec{K} \cdot \vec{r}} . \quad (9.117)$$

En reportant les développements (9.113) et (9.117) dans (9.116), on trouve :

$$\begin{aligned} - \sum_{\vec{K} \in \vec{\mathcal{B}}} \frac{\hbar^2}{2m} [i(\vec{k} - \vec{K})]^2 c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} + \sum_{\vec{K}' \in \vec{\mathcal{B}}, \vec{K}'' \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}') c_{\vec{k}-\vec{K}''} e^{-i\vec{K}' \cdot \vec{r}} e^{i(\vec{k}-\vec{K}'') \cdot \vec{r}} \\ = \sum_{\vec{K} \in \vec{\mathcal{B}}} \varepsilon(\vec{k}) c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} , \end{aligned} \quad (9.118)$$

soit, en posant $\vec{K}' + \vec{K}'' = \vec{K}$ dans la somme double :

$$\sum_{\vec{K} \in \vec{\mathcal{B}}} [\varepsilon^{(0)}(\vec{k} - \vec{K}) - \varepsilon(\vec{k})] c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} + \sum_{\vec{K}, \vec{K}' \in \vec{\mathcal{B}}^2} \mathcal{V}(\vec{K}') c_{\vec{k}-\vec{K}+\vec{K}'} e^{i(\vec{k}-\vec{K}) \cdot \vec{r}} = 0 . \quad (9.119)$$

Par identification des termes en $e^{-i\vec{K} \cdot \vec{r}}$ et après simplification par $e^{i\vec{k} \cdot \vec{r}}$, on en déduit le *système* d'équations pour les coefficients de Fourier de $\psi_{\vec{k}}$:

$$[\varepsilon^{(0)}(\vec{k} - \vec{K}) - \varepsilon(\vec{k})] c_{\vec{k}-\vec{K}} + \sum_{\vec{K}' \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K}') c_{\vec{k}-\vec{K}+\vec{K}'} = 0 , \quad \forall \vec{K} \in \vec{\mathcal{B}} , \quad (9.120)$$

ou, de façon équivalente :

$$[\varepsilon^{(0)}(\vec{k} - \vec{K}) - \varepsilon(\vec{k})] c_{\vec{k}-\vec{K}} + \sum_{\vec{K}'' \in \vec{\mathcal{B}}} \mathcal{V}(\vec{K} - \vec{K}'') c_{\vec{k}-\vec{K}''} = 0 , \quad \forall \vec{K} \in \vec{\mathcal{B}} . \quad (9.121)$$

En choisissant convenablement le zéro d'énergie potentielle – à la valeur moyenne de $V(\vec{r})$ sur une maille primitive –, la composante uniforme $\mathcal{V}(\vec{K}) = 0$ est nulle ; dans ces conditions, la somme dans (9.121) court explicitement sur les valeurs de $\vec{K}'' \neq \vec{K}$. Par ailleurs, ce choix entraîne qu'il n'y a pas de correction d'énergie

⁴⁰ $\mathcal{V}(\vec{K})$ est homogène à une énergie.

au premier ordre en l'absence de dégénérescence. En effet, en pareil cas, la correction au premier ordre est simplement la valeur moyenne de la perturbation prise sur un état d'ordre zéro, soit ici :

$$\int_{\mathbb{R}^3} d^3r e^{-i(\vec{k}-\vec{K})\cdot\vec{r}} V(\vec{r}) e^{i(\vec{k}-\vec{K})\cdot\vec{r}} = \int_{\mathbb{R}^3} d^3r V(\vec{r}) = 0 . \quad (9.122)$$

Jusqu'à présent, aucune approximation n'a été faite : le système (9.121) est exact, il est tout juste une reformulation de l'équation aux valeurs propres (9.116) dans l'espace de Fourier. Inexploitable en tant que tel dans le cas général, on en effectue maintenant l'analyse perturbative.

Le système (9.121) s'écrit aussi :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}} \mathcal{V}(\vec{K} - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.123)$$

Si $V \equiv 0$, il se réduit à :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = 0 , \quad \forall \vec{K} \in \tilde{\mathcal{B}} . \quad (9.124)$$

L'équation (9.124) dit que $c_{\vec{k}-\vec{K}}$ est forcément nul tant que $\varepsilon(\vec{k}) \neq \varepsilon^{(0)}(\vec{k} - \vec{K})$ et inversement : c'est seulement lorsque $\varepsilon(\vec{k})$ est égal à $\varepsilon^{(0)}(\vec{k} - \vec{K})$, que $c_{\vec{k}-\vec{K}}$ n'est pas nul et est, à ce stade, indéterminé ; c'est une constante non nulle faute de quoi la solution serait identiquement nulle, condition rencontrée à propos de toute équation aux fonctions propres. En tant que constante à la limite $V \rightarrow 0$, c'est donc un terme d'ordre $\mathcal{O}(V^0)$.

Dans le cas 1., par définition, toutes les énergies $\varepsilon^{(0)}(\vec{k} - \vec{K})$ sont, pour le \vec{k} choisi, bien isolées les unes des autres. Considérons-en une, celle ayant $\vec{K} = \vec{K}_1$ et récrivons le système (9.123) en exhibant à part la ligne $\vec{K} = \vec{K}_1$:

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_1)] c_{\vec{k}-\vec{K}_1} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}_1} \mathcal{V}(\vec{K}_1 - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad (9.125)$$

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}} \mathcal{V}(\vec{K} - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad \forall \vec{K} \in \tilde{\mathcal{B}}, \vec{K} \neq \vec{K}_1 . \quad (9.126)$$

Quand $V \rightarrow 0$, un seul crochet $[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})]$ figurant aux premiers membres du système (9.123) s'annule, celui de (9.125), et la valeur propre perturbée $\varepsilon(\vec{k})$ tend vers l'unique valeur isolée $\varepsilon^{(0)}(\vec{k} - \vec{K}_1)$ considérée. Bien sûr, tous les autres premiers membres de (9.126) s'annulent aussi, mais c'est alors parce que, le crochet n'étant pas nul à la limite, le coefficient $c_{\vec{k}-\vec{K}}$ lui-même tend vers zéro $\forall \vec{K} \neq \vec{K}_1$. Dans ces conditions, chaque $c_{\vec{k}-\vec{K}}$, $\vec{K} \neq \vec{K}_1$ est au moins d'ordre 1 en V ; il en résulte que la somme au second membre de (9.125) – qui ne contient que les $c_{\vec{k}-\vec{K}'}$, $\vec{K}' \neq \vec{K}_1$ – est au moins du second ordre⁴¹. En définitive, la différence d'énergie $\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_1)$ est du second ordre⁴² et la correction résultant du petit potentiel périodique est très petite. La fonction d'onde diffère de la fonction d'onde d'ordre zéro par des termes du premier ordre : l'ordre zéro est la série (9.108) réduite alors à un seul terme $e^{i(\vec{k}-\vec{K})\cdot\vec{r}}$, c'est bien une onde plane. Au total, il n'y a pas de changement qualitatif notable.

Dans le cas 2., toujours avec \vec{k} fixé, il existe par hypothèse un certain nombre, g , d'énergies quasi dégénérées associées à des vecteurs \vec{K}_i distincts :

$$|\varepsilon^{(0)}(\vec{k} - \vec{K}_i) - \varepsilon^{(0)}(\vec{k} - \vec{K}_j)| \lesssim |V| \quad (i, j = 1, 2, \dots, g) . \quad (9.127)$$

Mettons à part toutes les lignes où ces \vec{K}_i sont au premier membre de (9.123) :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}_i} \mathcal{V}(\vec{K}_i - \vec{K}') c_{\vec{k}-\vec{K}'} , \quad \vec{K}_i = \vec{K}_1, \vec{K}_2, \dots, \vec{K}_g , \quad (9.128)$$

⁴¹Comme toujours quand il s'agit de perturbation ordinaire (non-singulière), on fait l'hypothèse suivant laquelle tous les développements sont en puissances entières de la perturbation, ici V .

⁴²Bien sûr, l'absence de correction d'énergie au 1^{er} ordre est une conséquence du calage de $V(\vec{r})$ assurant que $\mathcal{V}(\vec{0}) = 0$.

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}} \mathcal{V}(\vec{K} - \vec{K}') c_{\vec{k}-\vec{K}'}, \quad \forall \vec{K} \in \tilde{\mathcal{B}}, \quad \vec{K} \neq \vec{K}_i, \quad i = 1, 2, \dots, g. \quad (9.129)$$

Le même argument que celui développé pour le cas 1. montre alors que les g crochets $[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)]$ des premiers membres de (9.128) s'annulent à la limite $V = 0$, donc les g coefficients $c_{\vec{k}-\vec{K}_i}$ sont tous d'ordre $\mathcal{O}(V^0)$. Tous les autres sont au moins d'ordre $\mathcal{O}(V)$: tout comme dans le cas 1., ils apparaissent au premier membre de (9.129), multipliés par le crochet de la différence des énergies, qui ne s'annule pas, cependant que le second membre contient des termes d'ordre $\mathcal{O}(V)$ (les g termes du genre $\mathcal{V}c_{\vec{k}-\vec{K}_i}$ - et aussi des termes d'ordre supérieur) et donc s'annule à la limite.

Alors, en extrayant du système (9.123) les seules équations où ces \vec{K}_i figurent au premier membre (il y en a g), il vient :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}' \in \tilde{\mathcal{B}}, \vec{K}' \neq \vec{K}_i} \mathcal{V}(\vec{K}_i - \vec{K}') c_{\vec{k}-\vec{K}'}, \quad \forall \vec{K} \in \{\vec{K}_i\}. \quad (9.130)$$

Au second membre de (9.130), figurent $g-1$ coefficients d'ordre zéro, $c_{\vec{k}-\vec{K}_j}$, $j \neq i$, tous les autres étant d'ordre $\mathcal{O}(V)$. Le système (9.130) tronqué aux g lignes pertinentes peut donc s'écrire :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}_j \in \{\vec{K}_i\}, j \neq i} \mathcal{V}(\vec{K}_i - \vec{K}_j) c_{\vec{k}-\vec{K}_j} + \mathcal{O}(V^2) \quad (i = 1, 2, \dots, g), \quad (9.131)$$

où tous les coefficients $c_{\vec{k}-\vec{K}_j}$ explicitement écrits sont $\mathcal{O}(V^0)$. Ceci étant, (9.131) montre alors que la correction d'énergie est une somme finie de termes d'ordre $\mathcal{O}(V)$, elle est donc aussi d'ordre $\mathcal{O}(V)$. Si on décide de rester à l'ordre le plus bas non trivial, il suffit de considérer seulement la somme explicitée au second membre de (9.131), et on est réduit à devoir résoudre :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)] c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}_j \in \{\vec{K}_i\}, j \neq i} \mathcal{V}(\vec{K}_i - \vec{K}_j) c_{\vec{k}-\vec{K}_j} \quad (i = 1, 2, \dots, g). \quad (9.132)$$

La somme au second membre contient au total $g-1$ termes ; l'omission des autres termes est le pendant, au voisinage d'une dégénérescence, de ce que prescrit la recette de perturbation dégénérée : elle précise de ne considérer que la *projection* de l'opérateur de perturbation à l'intérieur de chaque sous-espace dégénéré, en omettant tous les couplages non-diagonaux entre deux sous-espaces dégénérés distincts. En définitive, on obtient un système linéaire $g \times g$ pour les g coefficients $c_{\vec{k}-\vec{K}_i}$ relatifs aux g énergies quasi-dégénérées : c'est bien une version plus habile de la méthode de perturbations standard.

Afin de fixer les idées dans ce cas 2., supposons qu'il n'existe que deux valeurs de \vec{K} , \vec{K}_1 et \vec{K}_2 , telles que l'égalité $\varepsilon^{(0)}(\vec{k} - \vec{K}_1) = \varepsilon^{(0)}(\vec{k} - \vec{K}_2)$ peut être satisfaite ; comme tout est périodique dans le réseau réciproque, on peut toujours supposer que l'un des deux \vec{K}_i est nul, par exemple \vec{K}_1 , et poser alors $\vec{K}_2 = \vec{K}$ pour simplifier les écritures. Dès lors, on se trouve dans le cas où \vec{k} est au voisinage d'un seul plan de Bragg, précisément celui de \vec{K} . Le système simplifié (9.132), se réduit alors au couple d'équations :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k})] c_{\vec{k}} - \mathcal{V}(-\vec{K}) c_{\vec{k}-\vec{K}} = 0, \quad (9.133)$$

$$-\mathcal{V}(\vec{K}) c_{\vec{k}} + [\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})] c_{\vec{k}-\vec{K}} = 0. \quad (9.134)$$

Ce système linéaire homogène n'a de solution non triviale que si son déterminant est nul, ce qui fournit les deux valeurs propres approchées :

$$\varepsilon_{\pm}(\vec{k}) = \Sigma\varepsilon \pm \sqrt{[\Delta\varepsilon]^2 + |\mathcal{V}(\vec{K})|^2} \equiv \Sigma\varepsilon \pm h(\vec{k}), \quad (9.135)$$

où on a posé :

$$\Sigma\varepsilon = \frac{1}{2}[\varepsilon^{(0)}(\vec{k}) + \varepsilon^{(0)}(\vec{k} - \vec{K})] \quad \Delta\varepsilon = \frac{1}{2}[\varepsilon^{(0)}(\vec{k}) - \varepsilon^{(0)}(\vec{k} - \vec{K})]. \quad (9.136)$$

Cette expression donne donc la levée de dégénérescence dans le voisinage d'un point de l'espace réciproque où $\varepsilon^{(0)}(\vec{k}) = \varepsilon^{(0)}(\vec{k} - \vec{K})$; cette égalité est certainement vérifiée si :

$$\|\vec{k}\| = \|\vec{k} - \vec{K}\| \quad (9.137)$$

exprimant que l'extrémité du vecteur \vec{k} se trouve dans le plan bissecteur de \vec{K} , qui est par définition le plan de Bragg de \vec{K} , noté $\mathcal{P}_{\frac{\vec{K}}{2}}$.

Il y a donc maintenant deux solutions distinctes $\varepsilon_{\pm}(\vec{k})$, traduisant la levée de dégénérescence, laquelle est stricte quand \vec{k} est dans ce plan ; dans cette configuration, $\Delta\varepsilon$ s'annule et (9.135) donne :

$$\varepsilon_{\pm}(\vec{k} \in \mathcal{P}_{\frac{\vec{K}}{2}}) = \varepsilon^{(0)}(\vec{k}) \pm |\mathcal{V}(\vec{K})| . \quad (9.138)$$

Ainsi, les deux énergies qui coïncidaient à l'ordre zéro quand l'extrémité de \vec{k} est dans un plan de Bragg sont maintenant distinctes : le potentiel périodique a ouvert un *gap* d'énergie, la valeur propre $\varepsilon(\vec{k})$ différant nettement de $\varepsilon^{(0)}(\vec{k})$. Par continuité, il en va de même tant que \vec{k} reste au voisinage d'un tel plan. Quand l'extrémité de \vec{k} s'éloigne de tout plan de Bragg, les énergies perturbées rejoignent gentiment les $\varepsilon^{(0)}(\vec{k})$, à des termes $\propto(V^2)$ près.

En particulier, le plus petit vecteur \vec{k} de cette sorte est $\vec{k} = \frac{\vec{K}}{2}$, parallèle à une direction de $\tilde{\mathcal{B}}$ et on a alors plus précisément :

$$\varepsilon_{\pm}(\vec{k} = \frac{\vec{K}}{2}) = \varepsilon^{(0)}(\vec{k} = \frac{\vec{K}}{2}) \pm |\mathcal{V}(\vec{K})| . \quad (9.139)$$

Il existe un vecteur primitif \vec{b}_0 de $\tilde{\mathcal{B}}$ tel que $\vec{K} = n\vec{b}_0$ où $n \in \mathbb{Z}$. Si n est impair, le *gap* apparaît en bord de zone ; au contraire, si n est pair, celui-ci se forme en centre de zone. Ainsi par exemple, si on part de $\vec{k} = 0$, le premier *gap* rencontré dans la direction \vec{b}_0 est situé en $\vec{k} = 1 \times \frac{\vec{b}_0}{2}$, il est au bord de BZ1 ; le suivant est en $\vec{k} = 2 \times \frac{\vec{b}_0}{2} = \vec{b}_0$. Dans le schéma en zones réduites, il est au centre de BZ1, et ainsi de suite.

Le schéma de $\varepsilon^{(0)}(\vec{k})$ en zones réduites montre bien l'apparition d'un *gap* à chaque fois que deux courbes représentant un arc translaté $\varepsilon^{(0)}(\vec{k} - \vec{K})$ se croisent. À l'inverse, si l'on revient à la parabole libre non périodisée, on voit qu'elle se transforme en une suite d'arcs disjoints, les (petites) discontinuités survenant pour $\vec{k} = \pm\frac{\vec{K}}{2}, \pm\vec{K}, \pm 3\frac{\vec{K}}{2}, \dots$. L'amplitude des *gaps* décroît quand l'énergie augmente, partant de l'idée de $\mathcal{V}(\vec{q})$ est une fonction décroissante de $\|\vec{q}\|$ (pour $V(\vec{r})$ donné, plus \vec{q} a un grand module, plus l'intégrale de Fourier est oscillante). Par ailleurs, évidemment, les *gaps* décroissent lorsque le potentiel V décroît et, à la limite $V = 0$, on retrouve la parabole libre, dénuée de discontinuités.

Il est utile de regarder d'un peu plus près le voisinage d'un *gap*. Revenons à (9.135) en choisissant $\vec{K}_1 = 0$ (alors $\vec{K} = \vec{K}_2$) et, pour la commodité, mesurons \vec{k} relativement à $\frac{\vec{K}}{2}$:

$$\vec{k} = \frac{1}{2}\vec{K} + \vec{q} , \quad (9.140)$$

où \vec{q} est orienté de façon *quelconque* par rapport à \vec{K} (\vec{k} n'est dans le plan de Bragg que si $\vec{q} \perp \vec{K}$). Alors $\vec{k} - \vec{K} = -\frac{1}{2}\vec{K} + \vec{q}$ et :

$$\Sigma\varepsilon = \frac{1}{2} \frac{\hbar^2}{2m} \left[\left(\vec{q} + \frac{1}{2}\vec{K} \right)^2 + \left(\vec{q} - \frac{1}{2}\vec{K} \right)^2 \right] = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}) , \quad \Delta\varepsilon = \frac{\hbar^2}{m} \vec{q} \cdot \vec{K} . \quad (9.141)$$

Les valeurs propres approchées données par (9.135) s'écrivent maintenant :

$$\varepsilon_{\pm}(\vec{k}) = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}) \pm \sqrt{\left(\frac{\hbar^2}{2m}\right)^2 (\vec{K} \cdot \vec{q})^2 + |\mathcal{V}(\vec{K})|^2} . \quad (9.142)$$

Quand de surcroît l'extrémité de \vec{k} est dans le plan de Bragg (ce que l'on note abusivement $\vec{k} \in \mathcal{P}_{\frac{\vec{K}}{2}}$), alors $\vec{q} \perp \vec{K}$ et il reste :

$$\varepsilon_{\pm}(\vec{k} \in \mathcal{P}_{\frac{\vec{K}}{2}}) = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}) \pm |\mathcal{V}(\vec{K})| = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \frac{\hbar^2}{2m}(\vec{k}^2 - \vec{k} \cdot \vec{K} + \frac{\vec{K}^2}{2}) \pm |\mathcal{V}(\vec{K})| , \quad (9.143)$$

en agrément avec (9.138). De ceci il résulte :

$$\left(\vec{\nabla}_{\vec{k}} \varepsilon_{\pm}\right)_{\vec{k} \in \mathcal{P}_{\frac{\vec{K}}{2}}} = \frac{\hbar^2}{m} \left(\vec{k} - \frac{\vec{K}}{2}\right) , \quad (9.144)$$

quantité qui s'annule bien en $\vec{k} = \frac{\vec{K}}{2}$ (voir (9.90)). Le gradient de ε_{\pm} donne la direction de la normale à la surface $\varepsilon_{\pm}(\vec{k}) = C^{\text{ste}}$. On voit géométriquement que, l'extrémité de \vec{k} étant dans $\mathcal{P}_{\frac{\vec{K}}{2}}$, $\vec{k} - \frac{\vec{K}}{2}$ est dans ce plan. En d'autres termes, toute surface d'énergie constante intersecte perpendiculairement tout plan de Bragg.

Par ailleurs, quand \vec{k} est perpendiculaire à $\mathcal{P}_{\frac{\vec{K}}{2}}$ (c'est-à-dire $\vec{k} \parallel \vec{K}$ et $\vec{q} \parallel \vec{K}$) :

$$\varepsilon_{\pm}(\vec{k} \parallel \vec{K}) = \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \varepsilon^{(0)}(\vec{q}) \pm \sqrt{\left(\frac{\hbar^2 K q}{2m}\right)^2 + |\mathcal{V}(\vec{K})|^2} . \quad (9.145)$$

Cette expression permet de visualiser la variation des énergies de bande quand \vec{k} traverse un plan de Bragg tout en restant parallèle à \vec{K} ; quand \vec{k} est près de $\frac{\vec{K}}{2}$, \vec{q} est très petit et on peut faire un développement limité en q :

$$\begin{aligned} \varepsilon_{\pm}(\vec{k} \parallel \vec{K}, \vec{k} \simeq \frac{\vec{K}}{2}) &\simeq \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) + \frac{\hbar^2 q^2}{2m} \pm |\mathcal{V}(\vec{K})| \left[1 + \frac{[\hbar^2 K q / (2m)]^2}{2|\mathcal{V}(\vec{K})|^2}\right] \\ &= \varepsilon^{(0)}\left(\frac{\vec{K}}{2}\right) \pm |\mathcal{V}(\vec{K})| + \frac{\hbar^2 q^2}{2m} \left(1 \pm \frac{\hbar^2 K^2}{4m|\mathcal{V}(\vec{K})|}\right) . \end{aligned} \quad (9.146)$$

Depuis le début, on fait l'hypothèse que le couplage de réseau est faible devant les autres énergies pertinentes ; se donnant l'ordre de grandeur a du pas du réseau, il faut donc notamment que :

$$|\mathcal{V}(\vec{K})| \ll \frac{\hbar^2}{ma^2} . \quad (9.147)$$

Il en résulte que le signe du coefficient de q^2 dans (9.146) est positif pour ε_+ , négatif pour ε_- . Ceci montre que, comme attendu, les deux bandes se repoussent et ressemblent à deux paraboles de concavités opposées, séparées, là où elles sont extrémales, par l'énergie $2|\mathcal{V}(\vec{K})|$, qui est l'énergie de *gap*. Au voisinage du *gap*, tout se passe comme si l'électron situé dans la bande ε_- avait une masse *négative*. Il est d'ailleurs tentant de définir une masse effective m_{eff} comme :

$$m_{\text{eff}} = \hbar^{-2} \left[\frac{\partial^2 \varepsilon}{\partial q^2} \right]^{-1} = \frac{m}{1 \pm \frac{\hbar^2 K^2}{4m} |\mathcal{V}(\vec{K})|} , \quad (9.148)$$

où m est la masse "nue" de l'électron. Ceci vaut lorsque \vec{k} est parallèle à \vec{K} . Plus généralement, on définit un tenseur de masse effective, d'éléments :

$$m_{\text{eff } uv} = \hbar^{-2} \left[\frac{\partial^2 \varepsilon}{\partial k_u \partial k_v} \right]^{-1} \quad (u, v = x, y, z) , \quad (9.149)$$

qui joue un rôle essentiel dans la description du transport électronique dans un métal.

Une fois les valeurs propres (approchées) obtenues, il est possible de trouver les fonctions propres. Pour simplifier les notations, on suppose dans la suite que le cristal possède la symétrie d'inversion⁴³ :

$$V(-\vec{r}) = V(\vec{r}) , \quad (9.150)$$

⁴³C'est toujours vrai pour une structure cristalline sans motif, avec des objets à symétrie sphérique.

ce qui, allié au fait que V est réel, assure que $\mathcal{V}(\vec{K}) \in \mathbb{R}$. Les deux équations (9.133) et (9.134) peuvent être considérées comme la forme matricielle de l'équation aux vecteurs propres :

$$H_{\text{eff}} |\psi_{\vec{k}}\rangle = \varepsilon(\vec{k}) |\psi_{\vec{k}}\rangle \quad H_{\text{eff}} = \Sigma \varepsilon \mathbf{1} + \Delta \varepsilon \sigma_z + \mathcal{V}(\vec{K}) \sigma_x , \quad (9.151)$$

définie dans le sous-espace $\{|f_{\vec{k}}\rangle, |f_{\vec{k}-\vec{K}}\rangle\}$, avec $\langle \vec{r} | f_{\vec{q}} \rangle \equiv f_{\vec{q}}(\vec{r}) = e^{i\vec{q} \cdot \vec{r}}$. On a :

$$H_{\text{eff}} - \Sigma \varepsilon \mathbf{1} = h(\vec{k}) (\cos 2\theta \sigma_z + \sin 2\theta \sigma_x) = h(\vec{k}) e^{-i\theta \sigma_y} \sigma_z e^{+i\theta \sigma_y} . \quad (9.152)$$

où $h(\vec{k}) = \sqrt{[\Delta \varepsilon]^2 + |\mathcal{V}(\vec{K})|^2}$ (voir (9.135)) et avec :

$$\tan 2\theta = \frac{\mathcal{V}(\vec{K})}{\Delta \varepsilon} . \quad (9.153)$$

Les deux vecteurs propres $|\psi_{\vec{k}\pm}\rangle$ sont donc :

$$|\psi_{\vec{k}+}\rangle = e^{-i\theta \sigma_y} |f_{\vec{k}}\rangle \quad |\psi_{\vec{k}-}\rangle = e^{-i\theta \sigma_y} |f_{\vec{k}-\vec{K}}\rangle . \quad (9.154)$$

En effet, on a par exemple :

$$H_{\text{eff}} |\psi_{\vec{k}+}\rangle = h(\vec{k}) e^{-i\theta \sigma_y} \sigma_z e^{+i\theta \sigma_y} e^{-i\theta \sigma_y} |f_{\vec{k}}\rangle = h(\vec{k}) e^{-i\theta \sigma_y} \sigma_z |f_{\vec{k}}\rangle = +h(\vec{k}) e^{-i\theta \sigma_y} |f_{\vec{k}}\rangle \equiv +h(\vec{k}) |\psi_{\vec{k}+}\rangle , \quad (9.155)$$

et de même :

$$H_{\text{eff}} |\psi_{\vec{k}-}\rangle = h(\vec{k}) e^{-i\theta \sigma_y} \sigma_z e^{+i\theta \sigma_y} e^{-i\theta \sigma_y} |f_{\vec{k}-\vec{K}}\rangle = h(\vec{k}) e^{-i\theta \sigma_y} \sigma_z |f_{\vec{k}-\vec{K}}\rangle = -h(\vec{k}) e^{-i\theta \sigma_y} |f_{\vec{k}-\vec{K}}\rangle \equiv -h(\vec{k}) |\psi_{\vec{k}-}\rangle . \quad (9.156)$$

De (9.154) on déduit la forme explicite des fonctions de Bloch approchées $\psi_{\vec{k}\pm}$:

$$\psi_{\vec{k}+}(\vec{r}) = \cos \theta f_{\vec{k}}(\vec{r}) + \sin \theta f_{\vec{k}-\vec{K}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} (\cos \theta + \sin \theta e^{-i\vec{K} \cdot \vec{r}}) \equiv e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}+}(\vec{r}) , \quad (9.157)$$

$$\psi_{\vec{k}-}(\vec{r}) = -\sin \theta f_{\vec{k}}(\vec{r}) + \cos \theta f_{\vec{k}-\vec{K}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} (-\sin \theta + \cos \theta e^{-i\vec{K} \cdot \vec{r}}) \equiv e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}-}(\vec{r}) . \quad (9.158)$$

Chacune de ces fonctions redonne une simple onde plane dès que $\theta \simeq 0$: c'est ce qui arrive dès que $|\vec{K} \cdot (\vec{k} - \frac{\vec{K}}{2})|$ est supérieur ou de l'ordre de $m\hbar^{-2} |\mathcal{V}(\vec{K})| \ll a^{-2}$: dès que l'extrémité de \vec{k} s'éloigne un tout petit peu du plan de Bragg, les fonctions perturbées coïncident pratiquement avec les ondes planes, propres à l'ordre zéro. D'une façon générale, on trouve que $|\psi_{\vec{k}\pm}(\vec{r})|^2 = 1 \pm \sin 2\theta \cos \vec{K} \cdot \vec{r}$.

Les expressions (9.157) et (9.158) se simplifient lorsque l'extrémité de \vec{k} est précisément dans un plan de Bragg. Alors, $\Delta \varepsilon = 0$, $\theta = +\frac{\pi}{4}$ si $\mathcal{V}(\vec{K})$ est positif et :

$$\psi_{\vec{k}+}(\vec{r}) = \sqrt{2} e^{i(\vec{k} - \frac{\vec{K}}{2}) \cdot \vec{r}} \cos \frac{\vec{K} \cdot \vec{r}}{2} , \quad (9.159)$$

$$\psi_{\vec{k}-}(\vec{r}) = -i\sqrt{2} e^{i(\vec{k} - \frac{\vec{K}}{2}) \cdot \vec{r}} \sin \frac{\vec{K} \cdot \vec{r}}{2} . \quad (9.160)$$

Maintenant, les fonctions propres ne sont plus des ondes planes, mais des ondes stationnaires, combinaisons linéaires à poids égaux des deux exponentielles $e^{\pm i \frac{\vec{K}}{2} \cdot \vec{r}}$. C'est pourquoi il est usuel de dire que, en pareil cas, l'électron subit une réflexion de Bragg.

Enfin, il est visible que les $\psi_{\vec{k}\pm}(\vec{r})$, (9.157) et (9.158), satisfont le théorème de Bloch (les $u_{\vec{k}\pm}$ sont périodiques dans \mathcal{B}).

9.4.2 L'approximation des liaisons fortes

Il s'agit de la situation extrême opposée à celle que l'on vient de décrire : on va maintenant supposer que le potentiel ressenti par un électron au voisinage d'un cœur est essentiellement celui de ce dernier, le cristal pouvant alors être vu, en première approximation, comme une simple juxtaposition d'atomes, objets neutres en interaction faible. Il en résulte que les "bonnes" fonctions d'ordre zéro sont maintenant les fonctions atomiques centrées sur un atome donné, notées $\psi_{\nu \vec{R}}(\vec{r})$ (orbitales atomiques) ; ν désigne un ensemble de nombres quantiques. En gros, l'approximation consiste à supposer que, quand on s'éloigne d'un cœur, le potentiel de réseau $V(\vec{r})$, commence à différer du pur potentiel atomique quand on atteint une distance au moins égale à la portée des fonctions atomiques $\psi_{\nu \vec{R}}(\vec{r})$.

Toutes les fonctions $\psi_{\nu \vec{R}}$, ν fixé, sont en fait les répliques les unes des autres et se transforment les unes dans les autres par les translations du réseau de Bravais. Si on désigne simplement par $\psi_{\nu \vec{R}=\vec{0}}(\vec{r}) \equiv \psi_{\nu}(\vec{r})$ les fonctions centrées sur le nœud origine, on a :

$$\psi_{\nu \vec{R}}(\vec{r}) = \mathcal{T}(\vec{R}) \psi_{\nu}(\vec{r}) = \psi_{\nu}(\vec{r} - \vec{R}) . \quad (9.161)$$

À ce stade, on dispose *a priori* d'un jeu constitué par tous les états propres d'un Hamiltonien atomique (celui de l'atome situé au nœud origine $\vec{R} = \vec{0}$, H_0) et de toutes leurs répliques par le jeu des translations du réseau de Bravais. Cet ensemble est visiblement surcomplet, au sens où ses éléments ne sont pas tous linéairement indépendants les uns des autres. En effet, H_0 à lui seul offre une base complète ; en conséquence, toute fonction peut être développée sur cette base, en particulier n'importe quelle fonction du type $\psi_{\nu \vec{R} \neq \vec{0}}(\vec{r})$. Il existe donc des développements du genre :

$$\psi_{\nu \vec{R} \neq \vec{0}}(\vec{r}) = \sum_{\nu'} C_{\nu \nu'}(\vec{R}) \psi_{\nu'}(\vec{r}) . \quad (9.162)$$

avec leurs coefficients non tous identiquement nuls. Il en résulte que l'on ne doit pas retenir pour $\{\psi_{\nu}\}_{\nu}$ tous les états propres de H_0 , mais seulement un sous-ensemble. En pratique, d'ailleurs, on ne peut jamais introduire (pour des raisons techniques) qu'un petit nombre de fonctions de ce genre. En outre, l'idée physique sous-jacente à l'approximation en cours repose sur la notion d'état localisé : de ce point de vue, il serait stupide d'inclure parmi les ψ_{ν} les états non liés de H_0 – qui sont pourtant nécessaires pour assurer la complétude de la base propre de cet opérateur⁴⁴. Evacuer d'emblée ces états incongrus élimine de fait la difficulté liée à la surcomplétude.

D'ailleurs, finalement, rien n'oblige à choisir comme états localisés certains états propres de H_0 . Le caractère essentiel à préserver est la localisation autour de $\vec{R} = 0$, ce qui reste acquis quand on fait des combinaisons linéaires d'états $\psi_{\nu}(\vec{r})$: une combinaison linéaire en nombre fini d'états localisés donne clairement une fonction qui l'est encore. En définitive, l'ingrédient de base dans la suite est une (en fait plusieurs) fonction(s) notée(s) $\phi(\vec{r})$, localisée(s) dans le voisinage de $\vec{R} = \vec{0}$ et (mentalement) développable(s) sur un *petit* jeu de g fonctions atomiques ψ_{ν} localisées sur ce site :

$$\phi(\vec{r}) = \sum_{\nu=1}^g c_{\nu} \psi_{\nu}(\vec{r}) , \quad (9.163)$$

avec :

$$H_0 \psi_{\nu}(\vec{r}) = E_{\nu} \psi_{\nu}(\vec{r}) . \quad (9.164)$$

Les coefficients c_{ν} dans (9.163) sont déterminés suivant une procédure qu'il n'est pas utile de préciser ici et que l'on considèrera comme donnés. Celle-ci étant achevée, on dispose de g fonctions ϕ localisées autour du site $\vec{R} = 0$. Un label permet de les distinguer les unes des autres, le cas échéant, mais est omis dans la suite sauf nécessité. Propres du Hamiltonien atomique relatif à ce site, les énergies propres correspondantes sont bien séparées, leurs différences étant de l'ordre de l'eV.

Dans l'hypothèse des liaisons fortes, l'image physique à l'ordre zéro est celle d'un électron confiné près du site \vec{R} et placé dans une orbitale $\psi_{\nu \vec{R}}$ (ou une fonction du type ϕ ci-dessus), fortement localisée. \vec{R} variant,

⁴⁴L'inclusion de tels états délocalisés (étendus) ferait repartir clairement en direction de l'approximation des électrons presque libres.

toutes ces fonctions sont, à cet ordre, strictement équivalentes : on peut placer l'électron n'importe où, son énergie (à cet ordre) sera toujours la même⁴⁵. En réalité, l'électron a la possibilité de sauter d'un site à l'autre, ce qui se traduit par le fait que les éléments de matrice $\langle \psi_{\nu, \vec{R}} | H | \psi_{\nu', \vec{R}'} \rangle$ sont différents de zéro – quoique petits pour que le schéma de liaisons fortes ait un sens. L'existence de ce mouvement par sauts donne à l'électron une énergie cinétique, qui s'exprime par une loi de dispersion $\varepsilon(\vec{k})$. Comme l'amplitude des sauts est réputée petite, la variation de ε en \vec{k} sera faible : au total, la bande sera peu dispersée.

Outre $\phi(\vec{r})$, il convient donc d'introduire toutes les fonctions translattées :

$$\phi_{\vec{R}}(\vec{r}) = \mathcal{T}(\vec{R}) \phi(\vec{r}) = \phi(\vec{r} - \vec{R}) ; \quad (9.165)$$

notons dès maintenant que rien n'impose aux $\phi_{\vec{R}}$ d'être orthogonales entre elles ; en général :

$$\langle \phi_{\vec{R}} | \phi_{\vec{R}'} \rangle \neq 0 \quad \forall \vec{R}, \vec{R}' . \quad (9.166)$$

Quelle que soit l'image en arrière-plan (suggérant une approximation ou une autre), il s'agit toujours de trouver les états d'un électron dans un potentiel périodique $V(\vec{r})$, qui, de ce fait, satisfont le théorème de Bloch. À partir des $\phi_{\vec{R}}(\vec{r})$, il convient donc de trouver des états $\psi_{\vec{k}}(\vec{r})$ satisfaisant :

$$\mathcal{T}(\vec{R}) \psi_{\vec{k}}(\vec{r}) = e^{-i\vec{k} \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) , \quad (9.167)$$

qui est équivalent à :

$$\psi_{\vec{k}}(\vec{r} + \vec{R}) = e^{+i\vec{k} \cdot \vec{R}} \psi_{\vec{k}}(\vec{r}) . \quad (9.168)$$

Il existe des combinaisons linéaires des $\phi_{\vec{R}}(\vec{r})$ satisfaisant la relation (9.168). En effet, posons :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi_{\vec{R}'}(\vec{r}) = \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} - \vec{R}') . \quad (9.169)$$

En reportant dans (9.168), il vient :

$$\sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} + \vec{R} - \vec{R}') = e^{i\vec{k} \cdot \vec{R}} \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} - \vec{R}') . \quad (9.170)$$

En posant $\vec{R}'' = \vec{R}' - \vec{R}$ au premier membre :

$$\sum_{\vec{R}'' \in \mathcal{B}} A_{\vec{k}}(\vec{R} + \vec{R}'') \phi(\vec{r} - \vec{R}'') = e^{i\vec{k} \cdot \vec{R}} \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}') \phi(\vec{r} - \vec{R}') . \quad (9.171)$$

Les variables \vec{R}' et \vec{R}'' sont muettes, d'où :

$$A_{\vec{k}}(\vec{R} + \vec{R}') = e^{i\vec{k} \cdot \vec{R}} A_{\vec{k}}(\vec{R}') , \quad (9.172)$$

ce qui fournit la solution :

$$A_{\vec{k}}(\vec{R}) = C e^{i\vec{k} \cdot \vec{R}} \quad (C = \text{constante arbitraire}) . \quad (9.173)$$

En conclusion, étant donné un jeu de fonctions $\phi_{\vec{R}}$ déduites les unes des autres par les translations⁴⁶,

⁴⁵Le problème est donc très fortement dégénéré.

⁴⁶Noter toutefois que ce résultat ne dépend visiblement pas de l'aspect localisé éventuel des $\phi_{\vec{R}}(\vec{r})$: le point essentiel est que toutes ces fonctions se déduisent les unes des autres par translation.

toute fonction du type⁴⁷ :

$$\psi_{\vec{k}}(\vec{r}) = C \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k} \cdot \vec{R}} \phi_{\vec{R}}(\vec{r}) \quad (9.175)$$

satisfait le théorème de Bloch⁴⁸, quelle que soit la constante arbitraire C , choisie par normalisation⁴⁹. Elle peut se calculer en prenant des conditions cycliques dans une grande boîte contenant N sites au total ; en remarquant que les produits scalaires $\langle \phi_{\vec{R}} | \phi_{\vec{R}'} \rangle$ ne sont pas nuls, on trouve facilement⁵⁰ :

$$N |C|^2 \sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} \langle \phi_{\vec{R}=\vec{0}} | \phi_{\vec{R}} \rangle = 1 . \quad (9.177)$$

Si l'on dispose de g fonctions localisées de type ϕ , $\{\phi_n\}$, et donc aussi de toutes leurs translatées $\{\phi_{n\vec{R}}\}$, on pourra construire g fonctions de Bloch à \vec{k} donné :

$$\psi_{n\vec{k}}(\vec{r}) = C \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k} \cdot \vec{R}} \phi_{n\vec{R}}(\vec{r}) \quad (9.178)$$

L'usage des fonctions $\psi_{n\vec{k}}$ simplifie considérablement le problème sur un plan technique. En effet, elles prédiagonalisent H , par le fait que, satisfaisant le théorème de Bloch, elles sont fonctions propres des opérateurs de translation :

$$\mathcal{T}(\vec{R}) \psi_{n\vec{k}}(\vec{r}) = e^{-i\vec{k} \cdot \vec{R}} \psi_{n\vec{k}}(\vec{r}) . \quad (9.179)$$

En effet, H commute avec tous les opérateurs \mathcal{T} :

$$[H, \mathcal{T}(\vec{R})] = 0 \iff H\mathcal{T}(\vec{R}) = \mathcal{T}(\vec{R})H . \quad (9.180)$$

Prenons l'élément de matrice de (9.180) entre deux fonctions $\psi_{n\vec{k}}$ et $\psi_{n'\vec{k}'}$:

$$\langle \psi_{n\vec{k}} | H\mathcal{T}(\vec{R}) | \psi_{n'\vec{k}'} \rangle = \langle \psi_{n\vec{k}} | \mathcal{T}(\vec{R})H | \psi_{n'\vec{k}'} \rangle . \quad (9.181)$$

En vertu de (9.179), le premier membre est :

$$e^{-i\vec{k}' \cdot \vec{R}} \langle \psi_{n\vec{k}} | H | \psi_{n'\vec{k}'} \rangle . \quad (9.182)$$

Par définition de l'adjoint d'un opérateur, le second membre de (9.181) est :

$$(|\psi_{n\vec{k}}\rangle, \mathcal{T}(\vec{R})H|\psi_{n'\vec{k}'}\rangle) = ([\mathcal{T}(\vec{R})]^\dagger |\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) . \quad (9.183)$$

Comme les \mathcal{T} sont unitaires et que $[\mathcal{T}(\vec{R})]^{-1} = \mathcal{T}(-\vec{R})$, ceci vaut encore :

$$(\mathcal{T}(-\vec{R})|\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) = (e^{-i\vec{k} \cdot (-\vec{R})} |\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) = e^{-i\vec{k} \cdot \vec{R}} (|\psi_{n\vec{k}}\rangle, H|\psi_{n'\vec{k}'}\rangle) , \quad (9.184)$$

⁴⁷ Il est intéressant de remarquer que, finalement, la fonction de bonne symétrie (9.175) peut être obtenue directement en faisant agir l'opérateur \mathbf{T} :

$$\mathbf{T} = \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k} \cdot \vec{R}} \mathcal{T}(\vec{R}) \quad (9.174)$$

sur une fonction $\phi_{\vec{R}_0}$ quelconque vis-à-vis de la symétrie de translation. Ce fait est général : pour toute symétrie (géométrique, de permutations, etc.), la théorie des groupes permet de définir les bons projecteurs construisant automatiquement une fonction de symétrie requise. Ces opérateurs sont des combinaisons linéaires des opérateurs du groupe de symétrie, les coefficients étant (ici) les valeurs propres de ces opérateurs (plus généralement, ce sont les caractères des opérateurs, définis relativement à un type de symétrie donné). Pour un ensemble de particules identiques, la construction de la fonction ayant la bonne symétrie de permutation relève du même traitement : on applique un opérateur $Y_1 = N!^{-1} \sum_{\lambda} \varepsilon_{\lambda} P_{\lambda}$ à une fonction de symétrie *quelconque* pour obtenir la fonction de symétrie voulue.

⁴⁸ $\psi_{\vec{k}}$ satisfait le théorème de Bloch, mais n'est pas pour autant un état de Bloch exact, au sens où la fonction $u_{\vec{k}}$ doit satisfaire sa propre équation, ce qui n'est pas assuré en général avec le choix limité des fonctions "atomiques" $\phi_{\nu\vec{R}}$.

⁴⁹ Si on dispose de g fonctions sur chaque site, on peut définir g fonctions ϕ_n linéairement indépendantes, donc autant de fonctions $\psi_{n\vec{k}}$ satisfaisant le théorème de Bloch.

⁵⁰ Au total, la fonction de Bloch normalisée est, pour une seule dimension d'espace :

$$\psi_{\vec{k}}(x) = \frac{1}{\sqrt{N}} \sum_{n \in \mathbb{Z}} e^{in\vec{k}a} \phi_n(x) . \quad (9.176)$$

où l'antilinearité dans le bra a été utilisée pour écrire la dernière égalité. La quantité à droite n'est autre que $e^{-i\vec{k}\cdot\vec{R}} \langle \psi_{n\vec{k}} | H | \psi_{n'\vec{k}'} \rangle$. En rapprochant maintenant (9.182) et (9.184), il vient :

$$(e^{-i\vec{k}'\cdot\vec{R}} - e^{-i\vec{k}\cdot\vec{R}}) \langle \psi_{n\vec{k}} | H | \psi_{n'\vec{k}'} \rangle = 0 , \quad (9.185)$$

d'où il résulte que si $\vec{k}' \neq \vec{k}$, l'élément de matrice de H est nul : les $\psi_{n\vec{k}}$ assurent bien une prédiagonalisation de H . Toutefois, H possède *a priori* des éléments de matrice non-nuls entre deux états de même k mais de n différents :

$$\langle \psi_{n\vec{k}} | H | \psi_{n'\vec{k}} \rangle \neq 0 . \quad (9.186)$$

À ce stade, le problème est complètement posé dans les meilleures conditions et on est réduit à résoudre une équation aux valeurs propres dans un espace de dimension g , pour lequel la procédure numérique est la plus efficace.

Afin d'extraire les idées essentielles, supposons qu'en fait *une seule* orbitale atomique est introduite sur chaque site ($g = 1$), $\psi_{1\vec{R}}$, réplique par translation de l'orbitale ψ_1 définie sur le site à l'origine. Les définitions précédentes restent valables, avec partout $g = 1$; l'indice ν ne prend désormais qu'une seule valeur $\nu = 1$ mais est maintenu pour des raisons de clarté dans les notations. Notons aussi que la fonction ϕ définie en (9.163) se réduit à ψ_1 :

$$\phi(\vec{r}) = \psi_1(\vec{r}) , \quad H_0\psi_1 = E_1\psi_1 . \quad (9.187)$$

On sent bien que cette simplification sera justifiée si les niveaux atomiques sont bien séparés en énergie et si la modification résultant de la présence du réseau n'introduit que des éléments de matrice petits devant ces différences d'énergie. Cette simplification étant admise, et en conséquence de (9.185), H se trouve représenté par une matrice complètement diagonale sur les N fonctions $\{\psi_{\vec{k}}\}_{\vec{k}}$. Il reste dès lors à calculer l'élément diagonal $\langle \psi_{\vec{k}} | H | \psi_{\vec{k}} \rangle$, qui n'est autre, dans cette approximation, que la valeur propre $\varepsilon(\vec{k})$ cherchée. Pour calculer cet élément de matrice, on s'y prend comme suit.

On part de :

$$\psi_{\vec{k}} = C \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k}\cdot\vec{R}} \psi_1(\vec{r} - \vec{R}) , \quad (9.188)$$

d'où :

$$\varepsilon(\vec{k}) = |C|^2 \sum_{\vec{R}} \sum_{\vec{R}'} e^{-i\vec{k}\cdot\vec{R}'} e^{i\vec{k}\cdot\vec{R}} \langle \psi_{1\vec{R}'} | H | \psi_{1\vec{R}} \rangle . \quad (9.189)$$

L'élément de matrice se transforme en jouant avec l'invariance par translation de H :

$$\begin{aligned} \langle \psi_{1\vec{R}'} | H | \psi_{1\vec{R}} \rangle &= \int d^3r \psi_1^*(\vec{r} - \vec{R}') H \psi_1(\vec{r} - \vec{R}) = \int d^3r \psi_1^*(\vec{r}) H \psi_1(\vec{r} - \vec{R} + \vec{R}') \\ &\equiv \langle \psi_1 | H | \psi_{1\vec{R}-\vec{R}'} \rangle , \end{aligned} \quad (9.190)$$

d'où :

$$\varepsilon(\vec{k}) = |C|^2 \sum_{\vec{R}} \sum_{\vec{R}'} e^{i\vec{k}\cdot(\vec{R}-\vec{R}')} \langle \psi_1 | H | \psi_{1\vec{R}-\vec{R}'} \rangle = |C|^2 \sum_{\vec{R}} \sum_{\vec{R}''} e^{i\vec{k}\cdot\vec{R}''} \langle \psi_1 | H | \psi_{1\vec{R}''} \rangle . \quad (9.191)$$

La sommation sur \vec{R} fournit N fois le même terme, de sorte que :

$$\varepsilon(\vec{k}) = N|C|^2 \sum_{\vec{R}''} e^{i\vec{k}\cdot\vec{R}''} \langle \psi_1 | H | \psi_{1\vec{R}''} \rangle . \quad (9.192)$$

Pour achever le calcul, on effectue la partition suivante du Hamiltonien. Le Hamiltonien d'un électron dans le réseau est :

$$H = \frac{\vec{p}^2}{2m} + \sum_{\vec{R} \in \mathcal{B}} v(\vec{r} - \vec{R}) , \quad (9.193)$$

où $v(\vec{r})$ est l'énergie potentielle d'un électron soumis à l'attraction du seul nœud situé à l'origine⁵¹ ; en isolant le terme $\vec{R} = 0$:

$$H = \left[\frac{\vec{p}^2}{2m} + v(\vec{r}) \right] + \sum_{\vec{R} \in \mathcal{B}^*} v(\vec{r} - \vec{R}) \equiv H_0 + \Delta V , \quad (9.194)$$

⁵¹On obtient bien le potentiel du réseau en effectuant sur $v(\vec{r})$ toutes les translations du réseau de Bravais.

\mathcal{B}^* désignant le réseau privé du site à l'origine. Maintenant, l'élément de matrice du Hamiltonien apparaissant dans (9.192) est :

$$\langle \psi_1 | H_0 + \Delta V | \psi_{1\vec{R}} \rangle = E_1 \langle \psi_1 | \psi_{1\vec{R}} \rangle + \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle , \quad (9.195)$$

d'où la nouvelle forme de (9.191) compte tenu de (9.177) :

$$\varepsilon(\vec{k}) = \frac{\sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} (E_1 \langle \psi_1 | \psi_{1\vec{R}} \rangle + \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle)}{\sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \psi_{1\vec{R}} \rangle} = E_1 + \frac{\sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle}{\sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \psi_{1\vec{R}} \rangle} . \quad (9.196)$$

En raison du caractère localisé des $\psi_{1\vec{R}}(\vec{r})$, les éléments de matrice ont une dépendance grossièrement exponentielle par rapport à la distance $\|\vec{R}\|$ à l'origine. Le moins que l'on puisse faire est donc de retenir les seuls couples correspondant à des sites au plus premiers voisins entre eux. On obtient ainsi dans un premier temps :

$$\varepsilon(\vec{k}) = E_1 + \frac{\langle \psi_1 | \Delta V | \psi_1 \rangle + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle}{1 + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \psi_{1\vec{R}} \rangle} . \quad (9.197)$$

où les crochets indiquent que la sommation se fait sur les seuls premiers voisins du site origine $\vec{R} = \vec{0}$. Les produits scalaires de la somme du dénominateur sont usuellement très petits, et sont négligés le plus souvent. Dans ces conditions, on se retrouve avec :

$$\varepsilon(\vec{k}) = E_1 + \langle \psi_1 | \Delta V | \psi_1 \rangle + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle . \quad (9.198)$$

Le second terme peut être absorbé dans une redéfinition⁵² de E_1 , notée ε_0 ; avec ces diverses approximations, le résultat final retenu est :

$$\varepsilon(\vec{k}) = \varepsilon_0 + \sum_{\langle \vec{0}, \vec{R} \rangle} e^{i\vec{k}\cdot\vec{R}} \langle \psi_1 | \Delta V | \psi_{1\vec{R}} \rangle \quad (9.199)$$

c'est la dépendance en \vec{k} de la somme qui introduit la dispersion de la bande, laquelle reste faible puisque les éléments de matrice de ΔV restent petits par rapport aux énergies fixant l'ordre de grandeur de ε_0 .

À titre d'exemple, soit un réseau cubique simple ; le nœud origine a six premiers voisins situés à la distance a et tous les éléments de matrice sont égaux entre eux. En notant $w \in \mathbb{R}$ cette valeur commune, l'expression (9.199) donne :

$$\varepsilon(\vec{k}) = \varepsilon_0 + 2w (\cos k_x a + \cos k_y a + \cos k_z a) . \quad (9.200)$$

Cette expression est visiblement périodique ($\varepsilon(\vec{k} + \vec{K}) = \varepsilon(\vec{k})$). Elle définit une bande d'énergie s'étendant entre les deux valeurs⁵³ $\varepsilon_0 \pm 6|w|$; mesurée à la bonne aune, la dispersion est effectivement faible (bande étroite).

Remarque

D'une façon générale (et inversement), à partir de tout état de Bloch exact⁵⁴ $\psi_{\vec{k}}$, on peut indépendamment définir sa transformée de Fourier par :

$$\psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{k}\cdot\vec{R}} \phi_{\vec{R}}^{(W)}(\vec{r}) \iff \phi_{\vec{R}}^{(W)}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{k} \in BZ1} e^{-i\vec{k}\cdot\vec{R}} \psi_{\vec{k}}(\vec{r}) . \quad (9.201)$$

⁵²ou, pour faire pédant, une "renormalisation".

⁵³Le signe de w est sans incidence physique, comme il se doit puisqu'il s'agit d'un élément de matrice non-diagonal. Pour y voir plus clair, il suffit de regarder en détails le cas unidimensionnel (voir note 50) ; le changement de signe de w peut s'obtenir par un rephasage $(-1)^n$ de chaque ϕ_n , ce qui remplace simplement k par $k - \frac{\pi}{a} (\frac{2n}{a})$: rien ne change (physiquement), que l'on se place dans la première zone de Brillouin (k dans l'intervalle $[-\frac{\pi}{a}, \frac{\pi}{a}]$) ou dans une zone décalée de $\frac{\pi}{a}$ ($k \in [0, \frac{2\pi}{a}]$).

Très souvent, on choisit $w < 0$ pour faire commodément les petits dessins dans la première zone de Brillouin : on remplit alors la cuvette par le centre, au lieu de remplir l'"anti-cuvette" par les bords...

⁵⁴c'est-à-dire obtenu en résolvant exactement l'équation pour $u_{\vec{k}}$.

La fonction $\phi_{\vec{R}}^{(W)}$ ainsi définie porte le nom de fonction de Wannier. Contrairement aux fonctions atomiques $\psi_{\nu \vec{R}}$ ou $\phi_{\vec{R}}$ introduites ci-dessus, deux fonctions de Wannier relatives à deux sites distincts sont orthogonales :

$$\langle \phi_{\vec{R}}^{(W)} | \phi_{\vec{R}'}^{(W)} \rangle = \delta_{\vec{R} \vec{R}'} , \quad (9.202)$$

et constituent une autre base possible de représentation ; l'orthogonalité des fonctions de Wannier résulte du fait que la transformation (9.201) est unitaire, les $\psi_{\vec{k}}(\vec{r})$ étant mutuellement orthogonales. En tant que transformée de Fourier d'une fonction étendue, $\phi_{\vec{R}}^{(W)}(\vec{r})$ est une fonction présentant une forte localisation, un caractère qu'elle partage avec les orbitales atomiques $\psi_{\nu \vec{R}}$ ou les fonctions $\phi_{\vec{R}}$ introduites en (9.165).

9.5 Conducteurs, semi-conducteurs, isolants

Les deux situations extrêmes qui viennent d'être étudiées mettent en évidence un phénomène indissolublement lié à l'existence du potentiel périodique : l'ouverture de petits *gaps* en énergie par comparaison avec la distribution continue en énergie, de 0 à $+\infty$, pour un électron libre faiblement perturbé par le potentiel périodique, ou l'apparition d'une (faible) dispersion des états atomiques pour un électron fortement lié. En raison de la périodicité et du théorème de Bloch qui en résulte, en parcourant l'axe des énergies on rencontre des zones où il existe des états (bandes permises) et d'autres où il n'y en a pas (bandes interdites).

Cette propriété peut être utilement précisée à l'aide de la densité d'états, $\mathcal{D}(E)$; par définition, on a :

$$\text{Nombre d'états d'énergie comprise entre } E_0 \text{ et } E_0 + \delta E = \mathcal{D}(E_0) \delta E . \quad (9.203)$$

$\mathcal{D}(E_0) \delta E$ est une mesure de la dégénérescence et donne le nombre d'états de "même" énergie E_0 se distinguant les uns des autres par les autres nombres quantiques nécessaires pour spécifier complètement l'état quantique. Un électron dans un potentiel périodique est caractérisé par sa pseudo-impulsion $\hbar \vec{k}$, et son spin. En l'absence de champ magnétique, ou de tout effet spin-orbite, l'énergie d'un électron dans une bande donnée, $\varepsilon_n(\vec{k})$, ne dépend pas du spin. Le nombre d'états d'énergie E_0 à δE près s'obtient en sommant sur toutes les valeurs de \vec{k} donnant $\varepsilon_n(\vec{k}) = E_0$, et sur le spin. Ce dernier fournit un simple facteur 2, d'où ($\delta^3 k = (\frac{2\pi}{L})^3$) :

$$\sum_{\text{états}} \equiv 2 \sum_{\vec{k}} \rightarrow 2 \left(\frac{L}{2\pi} \right)^3 \int d^3 k . \quad (9.204)$$

En conséquence, la densité d'états de la bande d'énergie $\varepsilon_n(\vec{k})$ est donnée par⁵⁵ :

$$\mathcal{D}_n(E_0) = 2L^3 \int_{\mathbb{R}^3} \frac{d^3 k}{(2\pi)^3} \delta(\varepsilon_n(\vec{k}) - E_0) \quad ([\mathcal{D}] = E^{-1}) . \quad (9.205)$$

Autrement dit, $\mathcal{D}_n(E)$ résulte de la sommation sur tous les états possibles filtrée à l'aide de la fonction de Dirac. On sait que pour une fonction monotone⁵⁶ :

$$\delta(f(x) - f_0) = \frac{1}{|f'(x_i)|} \delta(x - x_i) . \quad (9.207)$$

À trois dimensions, cette relation se généralise en⁵⁷ :

$$\delta(f(\vec{x}) - f_0) = \frac{1}{\|\vec{\nabla} f(\vec{x}_0)\|} \delta((\vec{x} - \vec{x}_0) \cdot \vec{n}) , \quad f(\vec{x}_0) = f_0 , \quad (9.211)$$

⁵⁵ La densité d'états est parfois définie par unité de volume, ce qui revient à poser $L = 1$ dans la définition (9.205).

⁵⁶ Si la fonction n'est pas monotone, l'égalité $f(x) = f_0$ peut être satisfaite pour plusieurs valeurs x_i de x ; alors :

$$\delta(f(x) - f_0) = \sum_i \frac{1}{|f'(x_i)|} \delta(x - x_i) \quad f(x_i) = f_0 \quad \forall i . \quad (9.206)$$

⁵⁷ En effet, le développement de Taylor au voisinage de $\vec{x}_0 \in S_{f_0}$ s'écrit :

$$f(\vec{x}) = f_0 + (\vec{x} - \vec{x}_0) \cdot \vec{\nabla} f(\vec{x}_0) + \mathcal{O}((\vec{x} - \vec{x}_0)^2) . \quad (9.208)$$

où \vec{n} est le vecteur unitaire de la normale à la surface S_{f_0} d'équation $f(\vec{r}) = f_0$. Avec ces résultats, la densité (9.205) s'écrit :

$$\mathcal{D}_n(E_0) = L^3 \int_{\mathbb{R}^3} \frac{d^3k}{4\pi^3} \frac{1}{\|\vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k}_0)\|} \delta((\vec{k} - \vec{k}_0) \cdot \vec{n}) , \quad \varepsilon_n(\vec{k}_0) = E_0 . \quad (9.212)$$

La présence de la fonction de Dirac dit que seuls les points où $(\vec{k} - \vec{k}_0) \cdot \vec{n}$ est nul contribuent à l'intégrale, c'est-à-dire tous les points dont l'écart à \vec{k}_0 compté le long de la normale s'annule : il s'agit donc de tous les points dont le vecteur \vec{k} est *sur* la surface. Une expression équivalente à (9.212) pour la fonction $\mathcal{D}_n(E)$ est donc :

$$\mathcal{D}_n(E) = L^3 \int_{S_E} \frac{d^2k}{4\pi^3} \frac{1}{\|\vec{\nabla}_{\vec{k}} \varepsilon_n(\vec{k})\|} , \quad (9.213)$$

où S_E est la surface d'équation $\varepsilon_n(\vec{k}) = E$. Ces définitions se généralisent en toute dimension. Par exemple, pour un réseau plan, la densité \mathcal{D} s'obtient en effectuant un intégrale de *ligne* et non de surface, le long de la ligne d'énergie constante $\varepsilon_n(\vec{k}) = E$, \vec{k} étant maintenant un vecteur à deux composantes k_x, k_y .

À l'intérieur d'une maille primitive de $\tilde{\mathcal{B}}$, $\varepsilon_n(\vec{k})$ présente au moins un minimum et un maximum, en tant que fonction continue différentiable et périodique. En ces *extrema*, $\vec{\nabla}_{\vec{k}} \varepsilon_n$ s'annule et l'intégrand diverge. À $d = 3$, ces singularités sont intégrables et $\mathcal{D}_n(E)$ reste une fonction continue, tout en présentant des points anguleux appelés singularités de van Hove. Au contraire, à $d = 1$, quand $\vec{\nabla}_{\vec{k}} \varepsilon_n(k)$ s'annule, \mathcal{D}_n diverge logarithmiquement.

Pour une énergie donnée E , la densité totale s'obtient en faisant simplement la somme des densités des bandes :

$$\mathcal{D}(E) = \sum_n \mathcal{D}_n(E) , \quad (9.214)$$

de sorte que l'allure générale de $\mathcal{D}(E)$ est celle indiquée sur la figure 9.3, sans le souci de faire figurer les détails (et en particulier les singularités de van Hove). Bien sûr, il peut arriver – c'est fréquent – que deux bandes se recouvrent en énergie (voir fig. 9.3, à droite). En tout état de cause, le spectre des énergies à une particule d'un solide périodique n'est ni simplement connexe, ni purement ponctuel. En quelque sorte, c'est un hybride entre celui d'un électron libre et celui d'un électron lié dans un atome.

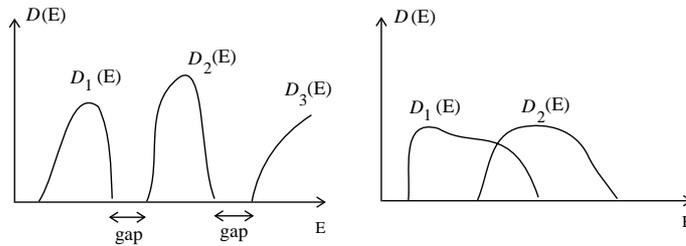


Figure 9.3: Illustration schématique de la densité d'états. À droite, l'exemple de deux bandes qui se recouvrent.

$\mathcal{D}(E)$ est tout particulièrement utile pour calculer les valeurs moyennes de grandeurs ne dépendant *que* de l'énergie ; en effet, $A(E)$ désignant une telle grandeur, sa moyenne contient des sommes du genre :

$$\mathcal{A} = \sum_{\text{états}} A(E) \quad \rightarrow \quad 2 \left(\frac{L}{2\pi} \right)^3 \int d^3k A(E) , \quad (9.215)$$

Par définition, $\vec{\nabla} f$ est perpendiculaire à la surface S_{f_0} , donc dirigé suivant la normale \vec{n} ; il en résulte que (9.208) se réécrit comme :

$$f(\vec{x}) = f_0 + \|\vec{x} - \vec{x}_0\| \|\vec{\nabla} f(\vec{x}_0)\| + \mathcal{O}((\vec{x} - \vec{x}_0)^2) , \quad (9.209)$$

ce qui donne :

$$\delta(f(\vec{x}) - f_0) = \delta(\|\vec{\nabla} f(\vec{x}_0)\|) (\vec{x} - \vec{x}_0) \cdot \vec{n} , \quad (9.210)$$

et compte tenu de (9.207), on trouve (9.211). On vérifie sans peine l'homogénéité de (9.211), sachant que $[\delta(\vec{x} \cdot \vec{n})] = L^{-1}$ (alors que $[\delta(\vec{x})] = L^{-3}$).

où le facteur 2 prend en compte le spin (ce qui suppose en outre que $A(E)$ ne dépend pas non plus du spin). L'intégration peut s'effectuer en deux temps : pour une bande de loi de dispersion $\varepsilon(\vec{k})$, on fixe la valeur de l'énergie et on intègre sur \vec{k} à E fixée :

$$\mathcal{A} = 2 \left(\frac{L}{2\pi} \right)^3 \int d^3k \delta(\varepsilon(\vec{k}) - E) A(E) . \quad (9.216)$$

Avec les règles habituelles concernant la fonction de Dirac rappelées plus haut, (9.216) devient :

$$2 \left(\frac{L}{2\pi} \right)^3 \int dE \int_{S_E} d^2k \frac{1}{\|\vec{\nabla}\varepsilon(\vec{k})\|} A(E) , \quad (9.217)$$

Au total, on a donc :

$$\mathcal{A} = \int dE \mathcal{D}(E) A(E) \equiv \int dE \tilde{A}(E) , \quad (9.218)$$

Les distinctions quantitatives des deux situations extrêmes étudiées plus haut peuvent s'illustrer comme indiqué sur la figure 9.4 : pour les électrons presque libres, les bandes sont larges et séparées par des petits *gaps* ; la situation est inverse pour un électron en liaisons fortes.

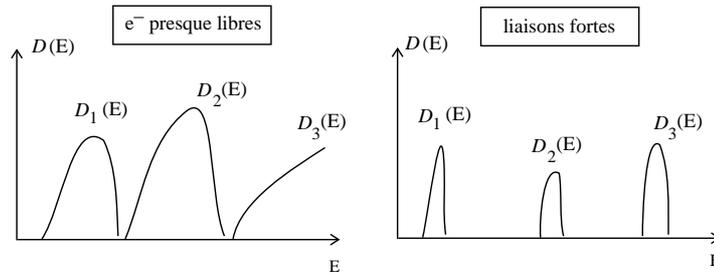


Figure 9.4: Illustration schématique des différences de la densité d'états pour les approximations des électrons presque libres (à gauche) et des liaisons fortes (à droite).

On est maintenant en mesure de comprendre – en raisonnant à température nulle⁵⁸ – l'existence de deux grandes classes de matériaux, les conducteurs et les isolants. La discussion repose crucialement sur l'hypothèse de la validité d'une théorie à une particule (ou d'une théorie de champ moyen), ce qui permet de s'appuyer sur l'existence d'états propres bien définis pour un seul électron (nu ou habillé), essentiellement corrélé aux autres par le Principe de Pauli.

Il s'agit de construire l'état du cristal – l'état fondamental –, étant donné un certain nombre d'électrons à caser, N_e . Ces derniers proviennent usuellement de la couche de valence des atomes qui s'assemblent pour former le cristal. Chaque atome fournit un petit nombre d'électrons susceptibles de participer au transport électrique, n_e ($n_e = 1$ pour un alcalin, $n_e = 8$ pour Fe, dont la configuration électronique est $[\text{Ar}] (3d)^6(4s)^2$) ; s'il y a N mailles primitives et un atome par maille (cristal monoatomique), il y a donc $N_e = n_e N$ électrons à placer dans les bandes à une particule, $\varepsilon_n(\vec{k})$, chacune d'entre elles pouvant accueillir $2N$ électrons.

Construire l'état fondamental, c'est donc remplir les états de Bloch les uns après les autres, en se déplaçant évidemment dans le sens des énergies croissantes. Pour simplifier, on supposera dans un premier temps que les bandes ne se recouvrent pas en énergie. Avec cette hypothèse, il existe une et une seule plus haute bande en énergie contenant des électrons, appelée bande de conduction⁵⁹. Deux cas sont alors possibles :

1. cette plus haute bande est pleine et contient donc $2N$ électrons. Alors, tout électron est coincé en énergie, sauf à gagner une énergie au moins égale à E_g , qu'aucun champ extérieur ne peut lui fournir. À température

⁵⁸Pour les métaux usuels, ε_F est de l'ordre de 10^4 eV ou plus, et on a toujours $k_B T \ll \varepsilon_F$; la distribution de Fermi reste donc toujours très voisine de la fonction de Heaviside $Y(\varepsilon_F - E)$.

⁵⁹l'avant-dernière bande complètement remplie s'appelle bande de valence.

nulle, le solide est donc un isolant. À température finie, les fluctuations thermiques ne peuvent rien changer tant que $k_B T \ll E_g$; comme $(k_B T)_{\text{ambiante}} \simeq 25 \text{ meV}$, le changement de comportement conductif induit par la température finie dépend de la valeur numérique de E_g :

- si E_g est de l'ordre de quelques eV (par exemple, pour le diamant $E_g = 5.5 \text{ eV}$), les probabilités d'excitation varient en gros comme $e^{-40 \times 5.5}$: autant dire qu'elles sont nulles. Le solide est isolant à toute température où il existe en tant que tel
- si E_g est nettement plus petit, une fraction d'eV pour fixer les idées (exemple : le germanium pour lequel $E_g \simeq 0.67 \text{ eV}$ à l'ambiante⁶⁰), les probabilités sont de l'ordre de e^{-30} , ce qui est encore très petit. Cependant, compte tenu de la relativement haute densité de porteurs potentiels, la conductivité devient notable, quoique très petite par rapport à un bon conducteur comme le cuivre. Le solide est alors dit semi-conducteur (intrinsèque⁶¹). La résistivité d'un semi-conducteur *décroît* quand la température augmente, ce qui est le comportement inverse de celui observé dans les métaux.

2. la plus haute bande est incomplètement remplie, ce qui situe le niveau de Fermi ε_F à l'intérieur de cette bande⁶² à la hauteur du dernier niveau occupé (niveau d'eau de la mer de Fermi). Les électrons qui se trouvent dans le voisinage inférieur de ε_F se voient offerts des niveaux d'énergie infiniment proches : dès lors, un champ électrique, aussi petit soit-il, peut de fait les mettre en mouvement et leur donner une énergie "orientée" – qui reste toutefois minuscule par rapport à l'énergie thermique – donnant lieu à un vecteur courant, donc au passage d'un courant électrique.

À l'intérieur d'une maille primitive de $\tilde{\mathcal{B}}$, il y a autant d'états \vec{k} distincts qu'il y a de mailles primitives dans \mathcal{B} , et chacun peut recevoir deux électrons compte tenu du spin. Si \mathcal{B} n'a pas de base⁶³ (cristal monoatomique) et si chaque atome fournit un seul électron ($n_e = 1$), alors la dernière bande est à moitié remplie⁶⁴ – toujours dans l'hypothèse de bandes qui ne se recouvrent pas – et le cristal est conducteur ; c'est encore vrai si n_e est impair. La situation est inversée si n_e est pair, et le solide est isolant⁶⁵.

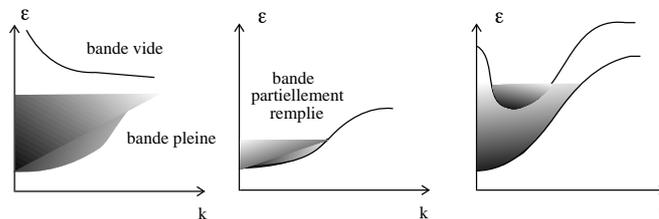


Figure 9.5: Illustration schématique selon la théorie des bandes (théorie à *une* particule) de la différence entre isolant (à gauche), métal (au milieu) et semi-métal (à droite).

L'existence de recouvrement⁶⁶ entre les bandes modifie sérieusement ces dernières conclusions. En pareil cas, si n_e est impair, les dernières bandes se partagent les électrons et on n'obtient évidemment pas un demi-remplissage ; toutefois, rien n'est changé fondamentalement en ce qui concerne les propriétés de conduction : le

⁶⁰La dilatation du réseau avec la température donne une dépendance en température aux énergies des *gaps*.

⁶¹Cette précision est là pour le distinguer d'un matériau où l'on a délibérément introduit des impuretés (dopage). Cette opération produit des états à l'intérieur du *gap*, qui sont autant de barreaux d'échelle permettant le franchissement par étapes de ce dernier.

⁶² ε_F est la limite du potentiel chimique $\mu(T)$ lorsque $T \rightarrow 0$. Pour un isolant ou un semi-conducteur, cette limite se situe dans le *gap* ; sa position précise dépend des courbures des bandes de conduction et de valence. Si celles-ci ont la même courbure, $\varepsilon_F \equiv \mu(0+)$ est au milieu du *gap*.

⁶³On rappelle que la structure cristalline d'un solide est complètement définie en donnant le réseau de Bravais \mathcal{B} et l'objet physique (atome, molécule, ion, ...) situé aux nœuds de ce réseau. Si l'on veut éviter une référence physique, on appelle simplement *motif* (ou *base*) l'objet situé à chaque nœud.

⁶⁴On parle alors de *demi-remplissage*.

⁶⁵Usuellement, pour un solide, on appelle *bande de valence* une bande pleine, et *bande de conduction* une bande partiellement remplie. Dans tous les cas, ce sont les électrons de valence *atomiques* qui les occupent.

⁶⁶Un cas limite est celui où deux bandes, l'une pleine, l'autre vide, entrent en contact sans toutefois se dépasser l'une l'autre en un nombre fini de points de BZ1. En pareil cas, le solide n'est pas conducteur (toujours à température nulle) puisque le canal dans l'espace de \vec{k} permettant de passer de la bande pleine à la bande vide est "infiniment" étroit (dit autrement, la frontière à l'énergie de Fermi ε_F est de mesure nulle).

solide reste conducteur. Au contraire, quand n_e est pair, un solide qui est isolant en l'absence de recouvrement devient conducteur puisque les deux bandes qui se recouvrent sont inévitablement maintenant incomplètes.

Il importe de réaliser le rôle crucial joué par la *dispersion* des bandes électroniques, c'est-à-dire le fait que leur énergie ε dépend de \vec{k} : un électron dans une bande plate ne peut pas gagner d'énergie sous l'effet d'un champ, même si la bande est quasi-vide. Quand la bande est partiellement remplie, les électrons peuvent certes passer d'un état \vec{k} à un état \vec{k}' sous l'effet de collisions (entre eux, sur les ions du réseau, sur des impuretés, ...), mais leur énergie ne change pas puisqu'elle est précisément indépendante de la pseudo-impulsion. La "platitude" de l'énergie d'une bande se reporte immédiatement sur la vitesse (moyenne) d'un électron de Bloch, en vertu de la relation $\vec{v} = \hbar^{-1} \vec{\nabla}_{\vec{k}} \varepsilon$: quand la bande est plate, la vitesse est nulle.

Dans les cas moins extrêmes, les grandeurs relatives des composantes de $\vec{\nabla}_{\vec{k}} \varepsilon$ sont l'une des causes premières des anisotropies de conductivité (d'autres facteurs doivent être pris en compte, notamment la densité d'états au niveau de Fermi). Pour un solide qui se clive facilement (le carbone graphite, par exemple), cette propriété de clivage provient principalement du fait que la distance entre deux plans de réseau en nid d'abeille est plusieurs fois supérieure⁶⁷ à la plus petite distance entre deux atomes de carbone d'un même plan. Comme la dispersion des bandes est pilotée par les amplitudes de saut – lesquelles ont en gros une dépendance exponentielle en distance –, il en résulte que la dispersion des bandes dans la direction perpendiculaire aux plans est beaucoup plus faible que parallèlement aux plans, d'où une forte anisotropie de conductivité. Notons aussi qu'une bande plate donne *ipso facto* une masse effective infinie, une autre façon de garder à l'esprit l'importance fondamentale de la dispersion.

On retiendra également le lien semi-quantitatif entre dispersion et délocalisation des états. Pour les électrons presque libres, la dispersion est très grande, cependant que les états sont essentiellement des ondes planes, faiblement modulées en amplitude par la fonction $u_{\vec{k}}$; il en résulte que la densité électronique est très diffuse⁶⁸ et que, si on prépare un paquet d'ondes localisé (sur un seul site, par exemple), il s'étale très vite. Au contraire, pour des électrons en liaisons fortes, la dispersion est faible ; les états propres sont toujours étendus, puisqu'il s'agit d'états de Bloch, mais ils présentent une forte structure spatiale, traduisant le fait que l'électron, en quelque sorte, passe par effet-tunnel d'un site à l'autre. Ici encore, un paquet d'ondes localisé finit par se diluer sur le réseau, mais l'évolution se fait sur une échelle de temps beaucoup plus longue ($\sim \frac{\hbar}{t}$, t désignant toujours l'amplitude de saut élémentaire). À la limite d'une amplitude de saut infiniment petite, l'énergie ne dépend pas de \vec{k} , la dégénérescence est $\sim N$ et les fonctions atomiques localisées sont, tout autant que les états de Bloch ou n'importe quelle autre combinaison linéaire, des fonctions propres du Hamiltonien d'un électron. En définitive, la dispersion ou la non-dispersion des bandes est un indice commode de la nature, localisée ou étendue, des états électroniques.

Remarques

1. Comme rappelé au début de la discussion, cette explication élémentaire de la distinction entre conducteurs et isolants repose crucialement sur l'hypothèse de la validité d'une théorie à une particule, où les interactions entre électrons sont oubliées – ou en tout cas lorsqu'elles peuvent être traitées par une méthode de type champ moyen qui, par nature, restitue formellement *in fine* le cadre d'un modèle à particules indépendantes. Lorsque tel n'est pas le cas, la question de la distinction isolant *vs* conducteur doit être entièrement reconsidérée. Il est facile de comprendre pourquoi en raisonnant physiquement par référence à un cas-limite.

Pour simplifier la discussion, supposons que chaque atome apporte un seul électron (demi-remplissage), auquel cas, selon la théorie des bandes, on est en présence d'un conducteur. Introduisons maintenant l'interaction entre les électrons (on sort du schéma des bandes) et admettons que leur répulsion mutuelle (traditionnellement notée U) est si forte que leur énergie cinétique puisse être négligée devant l'énergie potentielle d'interaction. Il est bien clair, alors, que chaque électron a envie de rester sur un atome : dès qu'il veut en bouger, les autres le repoussent fortement. En définitive, dans cette image extrême, les

⁶⁷environ 2.4 fois plus grande.

⁶⁸On parle alors usuellement d'états *étendus*.

électrons eux-mêmes ont tendance à former un réseau statique⁶⁹, coïncidant pour cette situation de demi-remplissage, au réseau de Bravais sous-jacent du solide⁷⁰. Il est physiquement évident qu'un tel solide ne peut être conducteur – tous les porteurs de charge sont bloqués. Il est non moins évident qu'il n'y a pas de contradiction : le schéma explicatif en terme de bande est invalide, puisqu'il résulte d'une théorie à une particule. Pour cette raison, il s'agit d'une classe particulière d'isolant : on l'appelle *isolant de Mott*.

Dans un cas moins extrême, la discussion repose sur les valeurs comparées de la répulsion électronique – d'ordre U quand deux électrons sont sur le même site – et de l'énergie cinétique, dont l'ordre de grandeur est noté (traditionnellement) t . Tout (ou presque) dépend du rapport $\alpha = \frac{t}{U}$. Conceptuellement, on peut imaginer une variation graduelle de ce rapport. Quand il est très grand, le solide considéré ci-dessus est sûrement conducteur ; quand il est très petit, on a un isolant. Le passage d'une phase à l'autre porte le nom de transition de Mott, où l'on voit la conductivité chuter brutalement à zéro quand α décroît au voisinage d'une certaine valeur critique α_C , comportement reminiscent d'une transition de phase survenant à une certaine température critique T_C .

Physiquement, on peut imaginer que cette variation graduelle se fait en dilatant progressivement le réseau d'un alcalin. Pour les valeurs usuelles de la maille, le solide est conducteur ; distendre le réseau fait décroître α . Quand le paramètre a de la maille est très grand, les "ions" se trouvent très loin les uns des autres et la structure est celle d'un antiferromagnétique⁷¹, les atomes portant chacun un électron dont le spin alterne d'un site à l'autre (ceci pour l'état singulet de spin total nul). Quand a décroît, t augmente puisque le passage d'un site à l'autre est facilité, le rapport $\frac{t}{U}$ croît et la transition de Mott peut se produire, le cas échéant. En pratique, la variation de a peut s'obtenir en comprimant le cristal : sous très haute pression, un isolant de Mott peut basculer dans l'état conducteur.

2. Dans un tout autre ordre d'idées, l'explication de la distinction isolant - conducteur donnée plus haut repose également sur une autre hypothèse importante, celle d'un réseau *rigide*. Dans certains cas, autoriser une déformation de réseau peut donner lieu à une modification spectaculaire de l'image fournie par la théorie des bandes avec un réseau fixé. L'exemple le plus connu est l'instabilité dite de Peierls.

L'idée est facile à comprendre dans le cas très simple d'un réseau unidimensionnel à un électron par maille. Si le réseau est rigide, la première zone de Brillouin s'étend entre $\pm\frac{\pi}{a}$ et est à moitié remplie : le solide est conducteur. Imaginons maintenant que l'atome initialement en na est déplacé d'une petite quantité $(-1)^n\delta a$. Autrement dit, on "dimérise" le réseau, les atomes $2n$ et $2n+1$ étant maintenant distants de $a-2\delta a$. Dès lors, le réseau de Bravais a une maille élémentaire de dimension $2a$ et la nouvelle première zone de Brillouin s'étend entre $\pm\frac{\pi}{2a}$. En repliant la bande d'énergie considérée⁷², on voit qu'il y a une stricte dégénérescence en bord de zone ($k = \pm\frac{\pi}{2a}$) ; c'est donc à ces endroits que les *gaps* vont s'ouvrir quand le réseau se déforme. Maintenant, la bande inférieure est pleine, la bande supérieure est vide, un *gap* existe et le solide est isolant !

Il s'agit bien d'une instabilité au sens où l'énergie du gaz d'électrons a spontanément tendance à décroître dès que la dimérisation est rendue possible⁷³. Cet abaissement d'énergie entre en compétition avec l'énergie élastique de déformation de réseau, puisque déformer le réseau coûte de l'énergie. En faisant le bilan, on trouve que l'énergie totale présente un minimum pour une certaine valeur δa_0 , qui fixe le taux de dimérisation du réseau, la conclusion restant inchangée : dans ce cas précis, autoriser le réseau à se déformer transforme un conducteur en isolant

⁶⁹Dans le régime des faibles densités ($r_s \gg a_0$), c'est le réseau de Wigner.

⁷⁰Quand on n'est pas à demi-remplissage, la densité d'électrons n'est pas commensurable avec la densité d'atomes, et la physique devient beaucoup plus riche (possibilité d'ondes de densité de charge, appelées CDW, acronyme de *charge density waves*).

⁷¹Cette affirmation, sommaire et désinvolte, laisse de côté de grandes difficultés. En effet, la structure antiferromagnétique ne s'impose pas en l'état actuel des choses puisque tout basculement de spins conservant $M_S = 0$ ne coûte rien ; on peut même imaginer une structure en domaines (des groupes de spins tous alignés, soit vers le haut, soit vers le bas). En définitive, dans la limite $\alpha = 0$, la dégénérescence est infinie. Savoir ce que fait réellement le système à α petit est une question d'une grande difficulté. Le fondamental du Hamiltonien de Hubbard a été obtenu en 1968 par Lieb et Wu ; dans le strict cadre du modèle de Hubbard, il n'y a pas de transition de Mott.

⁷²Après tout, rien n'interdit de représenter le réseau *non déformé* comme un réseau de maille double avec deux atomes par maille. On retrouvera cette façon de faire lors de l'étude des vibrations de réseau.

⁷³Visiblement, l'énergie de la bande pleine avec les *gaps* est plus basse que celle de la bande non déformée.

Chapitre 10

Vibrations d'un solide ordonné

Jusqu'à présent, les ions (atomes, cœurs, ions, *etc.*) ont été considérés comme fixes, situés aux nœuds d'un réseau de Bravais \mathcal{B} . Ce cadre restrictif doit être dépassé, pour plusieurs raisons, dont les plus importantes sont les suivantes :

- les vibrations du réseau mettent en jeu des degrés de liberté intervenant dans la chaleur spécifique C_V . Leur quantification permet de rendre compte du comportement basse température de celle-ci, l'un des premiers succès de la quantification (Einstein, 1905), dans le droit fil de l'idée-maîtresse de Planck (1900) : contrairement à ce qu'affirme la loi (classique) de Dulong et Petit, C_V n'est pas constante et contient un terme $\propto T^3$. Comme toujours, ce terme est essentiel à basse température, $k_B T \lesssim \hbar \omega_{\text{vibr}}$, où ω_{vibr} est une pulsation typique des vibrations du réseau
- la conduction dans un métal relève usuellement de la loi d'Ohm, laquelle ne peut être comprise sans l'intervention d'une dissipation d'énergie. Les vibrations y jouent un rôle majeur : le mouvement des électrons est un peu un jeu de billard où ceux-ci effectuent une sorte de marche au hasard en raison des collisions inélastiques¹ sur les ions (cœurs) du réseau. La dissipation est assurée par le caractère inélastique des collisions, rendu possible par la mise en mouvement des cœurs
- les spectres de diffraction – dont la théorie a été faite comme si les centres diffuseurs étaient fixes – sont évidemment altérés par les vibrations, donnant des élargissements spectraux dont l'analyse apporte de précieuses informations. En outre, il s'agit d'un problème où la dimensionnalité du réseau joue un rôle important.

La question des vibrations des centres massifs se pose en soi, indépendamment de l'arrangement spatial de ces derniers (cristallisé, vitreux, amorphe, *etc.*). Pour simplifier, on ne considère dans la suite que les vibrations d'un cristal parfait, complètement défini par la donnée de sa structure (réseau de Bravais et motif).

10.1 L'approximation harmonique

La considération des vibrations repose cependant sur l'idée que le cristal existe, et notamment que l'énergie (cinétique et potentielle) due à la vibration reste petite par rapport à l'énergie de cohésion. En pratique, ceci contient une hypothèse de petites oscillations : l'amplitude des vibrations doit rester petite devant les distances moyennes, assurant que l'énergie de vibration (potentielle et cinétique) est faible comparée à l'énergie statique d'interaction conduisant à l'énergie de cohésion² (dans le cas contraire, cela signifierait que le cristal est proche

¹Ceci ne préjuge en rien de l'existence d'autres causes de dissipation : impuretés chimiques ou structurales, dislocations,...

²Dans l'hypothèse harmonique qui va être faite, les énergies cinétique et potentielle de vibration sont égales entre elles, chacune représentant très exactement la moitié de l'énergie totale de vibration. On utilise parfois le mot *élastique* au lieu de *harmonique*, par opposition à *plastique* ; dans ce dernier cas, une fois la contrainte relâchée, chaque atome ne revient pas à son point de départ.

de son point de fusion). Dans toute la suite, on maintient donc l'existence d'un réseau de Bravais de référence – les nœuds de ce réseau sont alors les positions *moyennes* (d'équilibre) des objets qui vibrent – et on fait l'hypothèse de petites oscillations. Notant $\vec{r}(\vec{R})$ la position instantanée de l'ion³ vibrant près du nœud \vec{R} , on peut toujours écrire :

$$\vec{r}(\vec{R}) = \vec{R} + \vec{u}(\vec{R}) , \quad (10.1)$$

\vec{u} désignant l'écart à la position d'équilibre (qui est aussi la position moyenne) \vec{R} . Le Hamiltonien du réseau est :

$$H = \sum_{\vec{R} \in \mathcal{B}} \frac{\vec{P}_{\vec{r}(\vec{R})}^2}{2M} + \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} V(\vec{r}(\vec{R}) - \vec{r}(\vec{R}')) \equiv T + U . \quad (10.2)$$

Compte tenu de (10.1), l'énergie potentielle s'écrit :

$$U = \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} V(\vec{R} - \vec{R}' + \vec{u}(\vec{R}) - \vec{u}(\vec{R}')) , \quad (10.3)$$

et, dans l'hypothèse où les écarts sont faibles, le développement de Taylor au voisinage des positions d'équilibre \vec{R} et \vec{R}' donne :

$$U = U_{\text{eq}} + U_{\text{h}} , \quad (10.4)$$

avec :

$$U_{\text{eq}} = \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} V(\vec{R} - \vec{R}') , \quad (10.5)$$

qui correspond au réseau immobile, et :

$$U_{\text{h}} = \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} \sum_{\alpha, \beta=x, y, z} \frac{1}{2!} [\vec{u}(\vec{R}) - \vec{u}(\vec{R}')]_{\alpha} K_{\alpha\beta}(\vec{R} - \vec{R}') [\vec{u}(\vec{R}) - \vec{u}(\vec{R}')]_{\beta} . \quad (10.6)$$

Dans cette expression, les $K_{\alpha\beta}$ sont les dérivées secondes :

$$K_{\alpha\beta} = \left(\frac{\partial^2 V}{\partial r_{\alpha} \partial r_{\beta}} \right)_{\vec{R}, \vec{R}'} , \quad (10.7)$$

calculées aux positions d'équilibre du réseau d'ions fixes ; la notation $K_{\alpha\beta}$ évoque une constante de raideur de ressort parfait ($[K_{\alpha\beta}] = \text{MT}^{-2}$). U_{eq} est le terme donnant lieu à l'énergie de cohésion dans l'approximation des cœurs fixes. Arrêter le développement au second ordre inclus n'est manifestement sensé que dans l'hypothèse acceptée de petites oscillations. En réarrangeant les sommations, le terme harmonique⁴ U_{h} donné par (10.6) peut toujours être mis sous la forme :

$$U_{\text{h}} = \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} \sum_{\alpha, \beta=x, y, z} u_{\alpha}(\vec{R}) Q_{\alpha\beta}(\vec{R} - \vec{R}') u_{\beta}(\vec{R}') , \quad (10.8)$$

où les $Q_{\alpha\beta}$ sont des combinaisons linéaires des constantes de raideur $K_{\alpha\beta}$. En outre, si on donne le même déplacement \vec{u}_0 à tous les atomes (translation en bloc), l'énergie potentielle ne change pas. Il en résulte que :

$$0 = \frac{1}{2} \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} \sum_{\alpha, \beta=x, y, z} u_{0, \alpha}(\vec{R}) Q_{\alpha\beta}(\vec{R} - \vec{R}') u_{0, \beta}(\vec{R}') , \quad (10.9)$$

soit :

$$\sum_{\alpha, \beta=x, y, z} u_{0, \beta}(\vec{R}') u_{0, \alpha}(\vec{R}) \sum_{\vec{R}, \vec{R}' \in \mathcal{B}} Q_{\alpha\beta}(\vec{R} - \vec{R}') = 0 \quad \forall \vec{u}_0(\vec{R}) . \quad (10.10)$$

³Pour simplifier les notations, on suppose ici le réseau sans base. La généralisation, à ce stade, est un pur jeu d'écriture. Le cas d'un réseau avec base est examiné dans la suite.

⁴Le terme U_{eq} ne joue aucun rôle en ce qui concerne les vibrations, et sera omis dans la suite. Il reste que c'est ce terme qui, précisément, fixe la configuration d'équilibre, et donne l'énergie de cohésion dans l'hypothèse d'ions fixes.

La dernière somme est $N \sum_{\vec{R}} Q_{\alpha\beta}(\vec{R})$; d'où :

$$\sum_{\vec{R} \in \mathcal{B}} Q_{\alpha\beta}(\vec{R}) = 0 \quad \forall \alpha, \beta . \quad (10.11)$$

Une énergie potentielle étant définie, il est possible d'écrire les équations du mouvement (classiques⁵) pour chaque coordonnée $\vec{r}(\vec{R})$, étant entendu que $\dot{\vec{r}}(\vec{R}) = \dot{\vec{u}}(\vec{R})$. On a :

$$\begin{aligned} M \ddot{u}_\alpha(\vec{R}) &= -\frac{\partial}{\partial u_\alpha(\vec{R})} \frac{1}{2} \sum_{\vec{R}', \vec{R}''} \sum_{\alpha', \beta'} u_{\alpha'}(\vec{R}') Q_{\alpha'\beta'}(\vec{R}' - \vec{R}'') u_{\beta'}(\vec{R}'') \\ &= -\frac{1}{2} \sum_{\vec{R}', \vec{R}''} \sum_{\alpha', \beta'} [\delta_{\alpha'\alpha} \delta_{\vec{R}' \vec{R}} u_{\beta'}(\vec{R}'') + \delta_{\beta'\alpha} \delta_{\vec{R}'' \vec{R}} u_{\alpha'}(\vec{R}')] Q_{\alpha'\beta'}(\vec{R}' - \vec{R}'') . \end{aligned} \quad (10.12)$$

En jouant avec le nom des indices muets de sommation, ceci se récrit :

$$M \ddot{u}_\alpha(\vec{R}) = -\frac{1}{2} \sum_{\vec{R}'} \sum_{\beta} [Q_{\alpha\beta}(\vec{R} - \vec{R}') + Q_{\beta\alpha}(\vec{R}' - \vec{R})] u_\beta(\vec{R}') . \quad (10.13)$$

$Q_{\alpha\beta}$ est une matrice symétrique, puisque l'ordre des dérivations incluses dans (10.8) – voir aussi (10.7) – est sans importance. Par ailleurs, tout réseau de Bravais a un centre d'inversion⁶ : $Q_{\alpha\beta}(\vec{R})$ est donc invariant dans le changement $\vec{R} \rightarrow -\vec{R}$; au total :

$$Q_{\alpha\beta}(\vec{R}) = Q_{\beta\alpha}(\vec{R}) = Q_{\alpha\beta}(-\vec{R}) . \quad (10.14)$$

Compte tenu de ces symétries, les équations du mouvement (10.13) deviennent :

$$M \ddot{u}_\alpha(\vec{R}, t) = -\sum_{\vec{R}'} \sum_{\beta=x,y,z} Q_{\alpha\beta}(\vec{R} - \vec{R}') u_\beta(\vec{R}', t) . \quad (10.15)$$

10.2 Modes normaux d'un réseau unidimensionnel

Afin d'illustrer les idées précédentes, considérons un réseau unidimensionnel de paramètre a , dont les atomes peuvent vibrer le long du réseau⁷. En vertu du caractère unidimensionnel, chaque déplacement \vec{u} n'a qu'une composante et la matrice $K_{\alpha\beta}$ se réduit à un scalaire. La position de chaque ion est⁸ r_j :

$$r_j = R_j + u_j , \quad R_j = ja \quad (j \in \mathbb{Z}) . \quad (10.16)$$

et (10.6) devient :

$$U_h = \frac{1}{2} \sum_{j, j' \in \mathbb{Z}} \frac{1}{2!} K(R_j - R_{j'}) (u_j - u_{j'})^2 . \quad (10.17)$$

Même dans ce cas très simple, le calcul effectif de la fonction $K(R)$ à partir des premiers principes est en général très difficile, pour ne pas dire impossible. L'usage est de s'en remettre à une nouvelle approximation *ad hoc*, consistant à ne considérer que des termes entre premiers voisins, ce qui n'introduit ici qu'un seul paramètre $K(\pm a)$, noté⁹ simplement K :

$$K(R_j - R_{j'}) = \begin{cases} K & \text{si } |j - j'| = 1 \\ 0 & \text{autrement} \end{cases} . \quad (10.18)$$

⁵On verra en détail dans la section 10.2 que les équations classiques et quantiques sont formellement identiques.

⁶Il en va donc de même pour tout cristal monoatomique formé d'objets ponctuels ou à symétrie sphérique (scalaires, par opposition à vecteurs, ...)

⁷Les résultats sont inchangés si tous les atomes vibrent suivant une direction commune quelconque.

⁸ r_j n'est pas une longueur, c'est une abscisse $\in \mathbb{R}$.

⁹À ne pas confondre avec un vecteur de $\vec{\mathcal{B}}$!

Il n'existe pas vraiment de justification théorique d'une telle approximation, ce d'autant plus que les interactions élastiques sont à longue portée, surtout en basse dimensionnalité. En fait, la validation de ce schéma minimal – à $d = 1$ comme en dimension supérieure – repose sur ses succès quand on le confronte aux résultats expérimentaux. Ceci étant admis, et comme les deux termes $j' = j \pm 1$ sont en fait égaux entre eux (ce qui fait sauter un facteur $\frac{1}{2}$), le potentiel harmonique se réduit à :

$$U_h = \frac{1}{2} K \sum_{j \in \mathbb{Z}} (u_j - u_{j+1})^2 . \quad (10.19)$$

L'interprétation de cette expression est élémentaire ; les équations de Lagrange classiques avec une telle énergie potentielle conduisent à l'équation fondamentale de la Dynamique :

$$M \ddot{u}_j = - \frac{\partial U_h}{\partial u_j} = -K (2u_j - u_{j-1} - u_{j+1}) \iff \ddot{u}_j = -\omega_0^2 (2u_j - u_{j-1} - u_{j+1}) , \quad (10.20)$$

où on a posé :

$$K = M \omega_0^2 \quad \omega_0 = \sqrt{\frac{K}{M}} > 0 . \quad (10.21)$$

D'un autre point de vue, la force harmonique de rappel pour l'ion situé en ja résulte des deux forces exercées par ses deux voisins situés en $(j \pm 1)a$; elle vaut $-K(u_j - u_{j-1}) - K(u_j - u_{j+1})$: c'est très exactement ce qui figure au second membre de (10.20). Autrement dit, dans l'approximation en cours, tout se passe comme si on avait une chaîne de boules identiques reliées par des ressorts parfaits ayant tous la même constante de raideur K : on est finalement réduit à trouver les modes propres (classiques) d'un ensemble d'oscillateurs couplés, au nombre de N s'il y a N ions.

Si l'on regarde les choses d'un point de vue quantique, l'interprétation est la même. Le Hamiltonien de la chaîne d'ions est :

$$H = \sum_{j \in \mathbb{Z}} \frac{P_j^2}{2M} + \frac{1}{2} K \sum_{j \in \mathbb{Z}} (u_j - u_{j+1})^2 , \quad (10.22)$$

où P_j est le moment conjugué (impulsion) de l'ion situé au voisinage du point ja . Compte tenu de $[u_i, P_j] = i\hbar \delta_{ij}$, les équations de Heisenberg sont :

$$i\hbar \frac{d}{dt} u_{j,H} = [u_j, H]_H = \sum_{j' \in \mathbb{Z}} \frac{1}{2M} [u_j, P_{j'}^2]_H = \frac{i\hbar}{M} P_{j,H} \iff \dot{u}_{j,H} = \frac{P_{j,H}}{M} , \quad (10.23)$$

et :

$$\begin{aligned} i\hbar \frac{d}{dt} P_{j,H} &= [P_j, H]_H = \frac{1}{2} K \sum_{j' \in \mathbb{Z}} [P_j, (u_{j'} - u_{j'+1})^2]_H = \frac{1}{2} K [P_j, (u_j - u_{j+1})^2 + (u_{j-1} - u_j)^2]_H \\ &= -i\hbar K (u_j - u_{j+1} + u_j - u_{j-1}) \iff \dot{P}_{j,H} = -K (2u_j - u_{j+1} - u_{j-1})_H . \end{aligned} \quad (10.24)$$

En rapprochant (10.23) et (10.24), on trouve :

$$\ddot{u}_{j,H} = -\omega_0^2 (2u_j - u_{j+1} - u_{j-1})_H , \quad (10.25)$$

qui n'est autre que la version quantique de (10.20). La seule distinction entre ces deux équations est la nature des objets qui y figurent : des coordonnées u_j au sens classique, ou des observables $u_{j,H}$ en représentation de Heisenberg. Dans un cas comme dans l'autre, on est en face d'un système linéaire de N inconnues, qui se traite toujours de la même façon, puisque, en définitive, tous ces objets commutent toujours entre eux : l'algèbre reste donc "classique".

Finalement, notant x_j soit u_j , soit $u_{j,H}$, il convient de résoudre le système linéaire homogène :

$$\ddot{x}_j = -\omega_0^2 (2x_j - x_{j+1} - x_{j-1}) , \quad (10.26)$$

où, à ce stade, la nature précise des inconnues x_j est techniquement sans importance. Pour éviter des difficultés inessentiels, on procède comme d'habitude en considérant dans un premier temps un réseau de taille finie (N

atomes) muni des conditions aux limites commodes précisées en temps utile. Dès lors, (10.26) est un système linéaire $N \times N$ du second ordre en temps, dont la résolution exige la connaissance de $2N$ conditions initiales. Dans le cas classique, il s'agit des positions et des vitesses initiales des atomes de la chaîne ; dans le cas quantique, ce sont les $u_{j\text{H}}(t=0) \equiv u_j$ et des $P_{j\text{H}}(t=0) \equiv P_j$, c'est-à-dire les opérateurs au sens de Schrödinger. Quoiqu'il en soit, la méthode de résolution consiste à effectuer une transformée de Fourier en temps, en posant :

$$x_j(t) = \int_{-\infty}^{+\infty} d\omega X_j(\omega) e^{-i\omega t} . \quad (10.27)$$

En vertu de l'hermiticité – ou de la réalité – de x_j , on doit avoir :

$$X_j(-\omega) = X_j(\omega)^\dagger \quad \text{ou} \quad X_j(-\omega) = X_j(\omega)^* . \quad (10.28)$$

Le report dans (10.26) fournit :

$$\int_{-\infty}^{+\infty} d\omega [\omega^2 X_j - \omega_0^2 (2X_j - X_{j+1} - X_{j-1})] e^{-i\omega t} = 0 \quad \forall \omega , \quad (10.29)$$

ce qui impose aux X_j de satisfaire :

$$\omega^2 X_j - \omega_0^2 (2X_j - X_{j+1} - X_{j-1}) = 0 \quad (j \in \mathbb{Z}) . \quad (10.30)$$

Il s'agit d'un système linéaire homogène, qui n'a de solution non triviale que si son déterminant est nul. Cette condition fixe une fois pour toutes les seules valeurs possibles de ω^2 , en tant que zéros d'un polynôme de degré N , $\mathcal{P}_N(\omega^2)$. À ce stade – même si le calcul explicite n'est pas encore fait –, les valeurs possibles de ω^2 sont formellement déterminées¹⁰ par l'équation $\mathcal{P}_N(\omega^2) = 0$.

Par ailleurs, le système homogène (10.30) est une équation aux différences finies, dont les solutions sont de la forme $X_j = (\text{quelque chose})^j$ – et on peut toujours dire que (quelque chose) est une exponentielle. Posons donc :

$$X_j = C e^{ij\phi} ; \quad (10.31)$$

ϕ est une quantité réelle, sinon on aurait d'un côté ($j > 0$) ou de l'autre ($j < 0$) des divergences exponentielles inacceptables physiquement, tandis que C est un facteur indéterminé¹¹. Alors (10.30) devient :

$$[\omega^2 - \omega_0^2 (2 - e^{i\phi} - e^{-i\phi})] e^{ij\phi} = 0 \quad \iff \quad \omega^2 - 4\omega_0^2 \sin^2 \frac{\phi}{2} = 0 . \quad (10.32)$$

ω étant formellement déjà déterminé, ceci fixe ϕ complètement, ou *vice versa*. En pratique¹², il est plus commode d'obtenir d'abord ϕ , en choisissant des conditions aux limites agréables dont la nature précise est sans importance pour les résultats physiques tant que $N \gg 1$; comme d'habitude, on prend des conditions cycliques identifiant X_{N+1} et X_1 :

$$X_{N+1} = X_1 \quad \iff \quad e^{i(N+1)\phi} = e^{i\phi} \quad \iff \quad e^{iN\phi} = 1 . \quad (10.33)$$

Comme déjà vu à plusieurs reprises, ceci quantifie les valeurs de ϕ , un angle défini à 2π près ; ϕ est sans dimension, X_j est relatif au site situé en ja : il est donc naturel d'introduire un vecteur d'onde k , ici à une dimension¹³, k , tel que $\phi = ka$. Alors :

$$e^{iNka} = 1 \quad \iff \quad k = \text{entier} \times \frac{2\pi}{Na} \quad \iff \quad k = n \frac{2\pi}{Na} \equiv k_n . \quad (10.34)$$

¹⁰Il resterait à prouver que ce polynôme a non seulement tous ses zéros réels mais encore qu'ils sont tous positifs ou nuls. La réalité est assurée par le fait que \mathcal{P}_N est associé à une matrice symétrique (hermitique réelle). La positivité peut se démontrer soit par récurrence en N , soit en utilisant des propriétés générales (pour une chaîne ouverte – non refermée à la Born - von Karman –, \mathcal{P}_N est un polynôme bien connu, appelé polynôme de Gegenbauer).

¹¹Le système (10.30) étant *homogène*, les X_j s'en trouvent fixés à un facteur multiplicatif près

¹²Les deux procédures sont strictement équivalentes pour la raison suivante. Appliquer les conditions cycliques, c'est finalement jouer implicitement avec les opérateurs de rotation d'angle entier $\frac{2\pi}{N}$, qui représentent des symétries de la chaîne refermée par l'identification $N + 1 = 1$. Ces opérateurs forment un groupe abélien et, ce faisant, on diagonalise complètement *ipso facto* la matrice associée au système linéaire (10.30).

¹³ k est donc défini à $\frac{2\pi}{a}$ près.

Puisqu'il y a N modes propres au total, n prend N valeurs, par exemple¹⁴ $1, 2, \dots, N$; les valeurs de ω sont alors explicites :

$$\omega = \pm\omega(k) , \quad \omega(k) = 2\omega_0 \left| \sin \frac{ka}{2} \right| , \quad k \in \{k_n\}_n . \quad (10.35)$$

Au total, il n'y a bien que N valeurs distinctes pour k , ce qui donne également N valeurs distinctes pour les deux branches $\pm\omega(k)$, eu égard à la parité de $\omega(k)$ (pour chaque valeur de k , la dégénérescence est égale à 2). L'usage ordinaire¹⁵ consiste à faire varier k symétriquement autour de l'origine, essentiellement entre $\pm\frac{N}{2}\frac{2\pi}{Na} = \pm\frac{\pi}{a}$. La variation de la fonction $\omega(k)$ avec k constitue la loi de dispersion pour les modes de vibration ; pour un k donné, les X_j sont donnés par¹⁶ :

$$X_j(\omega(k)) = C e^{ikja} . \quad (10.36)$$

Il est instructif d'examiner en détail la collection des déplacements u_j résultant d'un état initial précis. Il suffit pour cela d'inverser la relation de Fourier (10.27) ; les valeurs possibles de ω sont en fait discrètes¹⁷, de sorte que l'intégrale dans (10.27) est à ce stade une somme \sum_k . Compte tenu des deux branches $\pm\omega(k)$ de la loi de dispersion, on peut toujours écrire $x_j(t)$ sous la forme :

$$x_j(t) = \sum_k X_j(\omega(k)) (\alpha_{k+} e^{i\omega(k)t} + \alpha_{k-} e^{-i\omega(k)t}) , \quad (10.37)$$

soit :

$$x_j(t) = \sum_k C e^{ikja} (\alpha_{k+} e^{i\omega(k)t} + \alpha_{k-} e^{-i\omega(k)t}) , \quad (10.38)$$

où les $\alpha_{k\pm}$ sont deux constantes d'intégration, à k fixé. Au total, il faut trouver $2N$ constantes $C\alpha_{k\pm}$, ce qui est possible en utilisant les $2N$ conditions initiales¹⁸. Raisonnons d'abord dans le cadre classique ; les x_j sont les u_j et selon (10.38) :

$$u_j(t) = \sum_k C e^{ikja} (\alpha_{k+} e^{i\omega(k)t} + \alpha_{k-} e^{-i\omega(k)t}) , \quad (10.39)$$

et les moments sont ($P_j \equiv M\dot{r}_j = M\dot{u}_j$) :

$$P_j(t) = M \sum_k i\omega(k) C e^{ikja} (\alpha_{k+} e^{i\omega(k)t} - \alpha_{k-} e^{-i\omega(k)t}) . \quad (10.40)$$

Les $2N$ conditions initiales s'écrivent :

$$u_j(0) = \sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} C e^{ikja} (\alpha_{k+} + \alpha_{k-}) , \quad (10.41)$$

$$P_j(0) = M \sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} i\omega(k) C e^{ikja} (\alpha_{k+} - \alpha_{k-}) . \quad (10.42)$$

Pour simplifier, on suppose toutes les vitesses initiales nulles ; (10.42) dit alors que $\alpha_{k+} = \alpha_{k-} = \alpha_k$ ce qui donne pour (10.39) :

$$u_j(t) = \sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} 2\alpha_k C e^{ikja} \cos \omega(k)t , \quad (10.43)$$

avec maintenant :

$$u_j(0) = \sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} 2\alpha_k C e^{ikja} . \quad (10.44)$$

¹⁴On pourrait tout autant prendre $n = -\frac{N}{2}, -\frac{N}{2} + 1, \dots, +\frac{N}{2}$, en supposant N pair – hypothèse non-restrictive dans la limite $N \gg 1$.

¹⁵ $\omega(k)$ est de toute façon une fonction périodique de k , avec la (plus petite) période $\frac{2\pi}{a}$.

¹⁶La constante multiplicative C vient du fait que le système (10.30) est homogène. C peut dépendre de k , mais cette dépendance est sous-entendue par simplicité.

¹⁷Ce sont les N solutions de $\mathcal{P}_N(\omega^2) = 0$.

¹⁸qui, dans le cas quantique, contiennent les opérateurs au sens de Schrödinger.

Cette dernière relation s'inverse facilement en utilisant¹⁹ :

$$\sum_{j=1}^N e^{i(k-k')ja} = N \delta_{kk'} \quad (N \gg 1) , \quad (10.45)$$

où, tant que N est fini, $\delta_{kk'}$ est le *symbole* de Kronecker. On trouve ainsi :

$$2N\alpha_k C = \sum_{j=1}^N u_j(0) e^{-ikja} . \quad (10.46)$$

Finalement, les déplacements pour une chaîne classique abandonnée hors de son point d'équilibre sans vitesse initiale pour les atomes :

$$u_j(t) = \frac{1}{N} \sum_{j'=1}^N u_{j'}(0) \sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} e^{ik(j-j')a} \cos \omega(k)t . \quad (10.47)$$

Dans la limite $N \gg 1$, la somme peut être remplacée par une intégrale :

$$\sum_{k \in [-\frac{\pi}{a}, +\frac{\pi}{a}]} \rightarrow \frac{L}{2\pi} \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} dk , \quad (10.48)$$

et comme $L = Na$, on obtient :

$$u_j(t) = \sum_{j'=1}^N u_{j'}(0) \frac{a}{2\pi} \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} dk e^{ik(j-j')a} \cos \omega(k)t . \quad (10.49)$$

Compte tenu de la dépendance de ω en k (voir (10.35)), l'intégrale porte sur des exponentielles du genre $e^{ik(j-j')a} e^{\pm 2i\omega_0 t \sin(ka/2)}$; le changement de variable $ka/2 = \phi$ fait apparaître des intégrales $e^{2i(j-j')\phi} e^{\pm 2i\omega_0 t \sin \phi}$, caractéristiques de la définition des fonctions de Bessel²⁰. Au total, $u_j(t)$ est une combinaison linéaire de ces fonctions spéciales :

$$u_j(t) = \sum_{j'=1}^N u_{j'}(0) J_{2(j-j')}(2\omega_0 t) . \quad (10.51)$$

En particulier, si un seul atome est déplacé au départ de u_0 , celui en j_0 , on a $u_j(0) = u_0 \delta_{jj_0}$ et :

$$u_j(t) = u_0 J_{2(j-j_0)}(2\omega_0 t) ; \quad (10.52)$$

aux grands temps $t \gg \omega_0^{-1}$, on a à peu près :

$$u_j(t) \sim \frac{u_0}{\sqrt{\pi\omega_0 t}} \cos \left[2\omega_0 t - (j - j_0) \frac{\pi}{2} - \frac{\pi}{4} \right] ; \quad (10.53)$$

à t fixé et grand, le champ des déplacements a en gros l'allure d'une sinusoïde, avec une amplitude $\sim t^{-\frac{1}{2}}$. Clairement, le "paquet d'ondes" de la déformation initiale s'élargit et se dilue le long de la chaîne ; notamment, le déplacement initial en j_0 s'évanouit et tend vers zéro comme $t^{-\frac{1}{2}}$.

En $t = 0$, le second membre de (10.47) est²¹ :

$$\frac{1}{N} \sum_{j'} u_{j'}(0) \sum_k e^{ik(j-j')a} = \frac{1}{N} \sum_{j'} u_{j'}(0) N \delta_{jj'} = u_j(0) , \quad (10.54)$$

¹⁹Cette relation a déjà été utilisée à propos de l'approximation des liaisons fortes et aussi pour établir la condition de Bragg.

²⁰La fonction de Bessel ordinaire J_n , n entier, est l'intégrale :

$$J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} d\phi e^{-in\phi} e^{ix \sin \phi} \quad (n \in \mathbb{Z}) . \quad (10.50)$$

Pour $x \gg 1$ (et $x \lesssim |n|$), on a $J_n(x) \sim \sqrt{\frac{2}{\pi x}} \cos(x - n\frac{\pi}{2} - \frac{\pi}{4})$

²¹Il en va de même si on fait $t = 0$ au second membre de (10.49).

comme il se doit. Par ailleurs, si on déplace initialement chaque atome de la même quantité $u(0)$, (10.47) donne :

$$u_j(t) = \frac{u(0)}{N} \sum_{j'=1}^N \sum_k e^{ik(j-j')a} \cos \omega(k)t = \frac{u(0)}{N} \sum_k e^{ikja} N \delta_{ka0} \cos \omega(k)t = u(0) . \quad (10.55)$$

Dans ces circonstances, chaque atome garde sa position initiale à tout temps, ce qui est bien normal : un tel état de départ est en fait une translation en bloc du réseau, sans déformation, n'induisant aucune force de rappel et ne provoquant donc aucun mouvement ultérieur. Ceci permet de retrouver le fait que le mode $k = 0$ est en réalité associé à un déplacement (translation) en bloc.

D'une façon générale, les modes de petit k ($ka \ll 1$) correspondent à des déformations dynamiques $u_j(t)$ variant lentement en fonction de l'entier j . Autrement dit, les déformations correspondantes ont une variation spatiale lente à l'échelle du pas du réseau (si les déplacements initiaux $u_{j'}(0)$ possèdent cette propriété, il en ira évidemment de même à tout temps). Dès lors, ce sont les modes de grande longueur d'onde qui sont prépondérants : on ne peut plus bien sûr se contenter strictement de $k = 0$, mais, à l'inverse, le *voisinage* de $k = 0$ suffit. Alors, dans (10.49), on peut remplacer $\omega(k)$ par son développement limité (voir(10.35)) :

$$\omega(k) = 2\omega_0 \frac{|k|a}{2} + \dots \equiv |k|c + \dots \quad (10.56)$$

où a été introduite une certaine vitesse c :

$$c = a\omega_0 , \quad (10.57)$$

qui n'est autre que la vitesse du son (vitesse de groupe et vitesse de phase coïncident dans la limite du solide continu²²). De la sorte, en posant $x = ja$, $x' = j'a$, $u(x = ja, t) = u_j(t)$, (10.49) devient :

$$u(x, t) \simeq \frac{a}{2\pi} \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} dk \sum_{j'=1}^N u(x', 0) e^{ik(x-x')} \cos kct . \quad (10.58)$$

Dans cette optique à grande échelle, tout se passe comme si le pas a du réseau était un "infiniment petit". Il est donc licite de remplacer $a \sum_{j'}$ par $\int dx'$; (10.58) a donc le même sens que²³ :

$$u(x, t) \simeq \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} \frac{dk}{2\pi} \int dx' u(x', 0) e^{ik(x-x')} \cos kct . \quad (10.60)$$

Il s'agit de détails ; l'une ou l'autre de ces dernières expressions montre que la déformation initiale se propage le long du réseau sous la forme d'un paquet d'ondes dispersé suivant la loi (10.35), dont vitesse de phase et vitesse de groupe coïncident dans la limite des grandes longueurs d'onde. Une branche de vibration qui tend vers zéro linéairement en k est appelée branche acoustique²⁴ :

$$\lim_{k \rightarrow 0} \omega(k) = 0 \quad \text{et} \quad \lim_{k \rightarrow 0} \frac{\omega(k)}{k} = c \equiv C^{\text{ste}} \iff \text{branche acoustique} . \quad (10.61)$$

D'un autre côté, si les déplacements initiaux varient vite avec j , il faut conserver la loi de dispersion (10.35) dans toute son intégrité. La vitesse de groupe associée à la loi de dispersion (10.35) est :

$$v_g(k) = \frac{d\omega(k)}{dk} = a\omega_0 \cos \frac{ka}{2} . \quad (10.62)$$

²²Dans la limite $ka \rightarrow 0$, la nature discrète du réseau s'efface, toutes les longueurs d'onde du champ des déplacements étant très grandes devant le pas a du réseau.

²³La linéarisation en k de la loi de dispersion – et la limite $a \rightarrow 0$ – rend élémentaire l'intégration en k de (10.60), qui fait apparaître $\delta(x - x' \pm ct)$, d'où :

$$u(x, t) = \frac{1}{2} [u(x - ct, t = 0) + u(x + ct, t = 0)] ; \quad (10.59)$$

cette expression est bien la solution de l'équation de propagation ($c^{-2}\partial_{t^2} - \partial_{x^2}$) $u = 0$ et représente la dynamique d'une onde dans un milieu non dispersif (toutes les fréquences avancent à la même vitesse) ; dans ces conditions, tout paquet d'ondes se propage en bloc, sans déformation. Retomber sur l'équation de propagation ne doit pas surprendre : le deuxième membre de (10.26) n'est rien d'autre qu'une dérivée seconde *discrète* par rapport à l'espace.

²⁴L'origine de la terminologie sera claire par la suite.

Elle s'annule lorsque $ka = \pi$, c'est-à-dire en bord de zone. Un paquet d'ondes ayant une amplitude maximum en $k = k_0 = \frac{\pi}{a}$ ne se propage donc pas et fait du sur-place. Pour comprendre ce que ceci signifie en termes de déplacements, choisissons un état initial où :

$$u_j(t=0) = (-1)^j u_0 . \quad (10.63)$$

Ceci représente un état où les sites pairs sont tous écartés d'une même quantité u_0 , tous les sites impairs étant écartés de $-u_0$; visiblement, c'est bien un cas où le champ des écarts initiaux varie vite à l'échelle du réseau (il est nul en moyenne sur deux mailles adjacentes). Le report de (10.63) dans (10.49) fait apparaître la somme :

$$\sum_{j'=1}^N e^{-ij'ka} u_{j'}(0) = u_0 \sum_{j'=1}^N e^{-ij'ka} (-1)^{j'} = u_0 \sum_{j'=1}^N e^{ij'(\pi-ka)} = u_0 N \delta_{ka\pi} . \quad (10.64)$$

Insérée dans une intégrale sur k , $N \delta_{ka\pi}$ se comporte comme une fonction de Dirac, $C \delta(ka-\pi)$ – soit $\frac{C}{a} \delta(k-\frac{\pi}{a})$; en outre, il est facile de voir²⁵ que $C = 2\pi$. L'expression de $u_j(t)$ selon (10.49) est alors :

$$u_j(t) = u_0 \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} dk e^{ijk a} \delta(k - \frac{\pi}{a}) \cos \omega(k)t = u_0 e^{ij\pi} \cos 2\omega_0 t , \quad (10.65)$$

où l'équation (10.35) a été utilisée ($\omega(k = \frac{\pi}{a}) = 2\omega_0$). Finalement :

$$u_j(t) = u_j(0) \cos 2\omega_0 t . \quad (10.66)$$

Tous les sites vibrent donc à la même fréquence $2\omega_0$ et deux sites voisins restent en opposition de phase, tout comme ils l'étaient²⁶ à $t = 0$. Si l'on regarde le “paquet d'ondes” réduit à deux tels sites voisins, on voit en effet qu'il reste en place : il ne bouge pas et ne se déforme pas.

Le cas quantique se traite de la même façon ; en prenant les valeurs moyennes des opérateurs à la Heisenberg sur un état initial des atomes où toutes les vitesses sont nulles, on obtient exactement les mêmes expressions, $u_j(t)$ donné par (10.49) étant alors la valeur moyenne (quantique) du déplacement à l'instant t quand les noyaux sont immobiles (en moyenne) dans l'état initial.

10.3 Modes normaux d'un réseau unidimensionnel à deux atomes par maille

On a vu dans la section 10.2 que les modes $k \approx 0$ d'une chaîne monoatomique correspondent à un champ de déplacement de grande longueur d'onde. À la limite où $k = 0$, ceci donne un mode de translation en bloc d'énergie nulle : il n'est donc pas surprenant que la pulsation $\omega(k)$ s'annule en $k = 0$.

Par ailleurs, rien n'interdit de considérer le réseau précédent comme un réseau de maille $a' = 2a$, avec une base à deux atomes séparés de la même distance a , donc constituée de molécules diatomiques (“dimères”) ; dans cette optique, le plus petit vecteur du réseau réciproque est de norme²⁷ $K = \frac{\pi}{2a}$. En suivant le même procédé que pour les électrons dans un potentiel périodique, on peut récrire les résultats de la section 10.2 dans un schéma en zones réduites, où k varie maintenant entre $\pm \frac{\pi}{2a}$. La première branche qui apparaît, notée désormais ω_{ac} (acoustique) est donc identique à (10.35), mais est maintenant bornée entre $\pm \frac{\pi}{2a}$:

$$\omega_{ac}(k) = 2\omega_0 \left| \sin \frac{ka}{2} \right| , \quad k \in \left[-\frac{\pi}{2a}, +\frac{\pi}{2a} \right] . \quad (10.67)$$

²⁵ Le pas en k est $\delta k = \frac{2\pi}{a}$ et $L = Na$.

²⁶ Ainsi, la symétrie “locale” présente dans l'état de départ est préservée à tout instant ultérieur. Ce n'est pas surprenant : aucun terme dans le Hamiltonien n'est susceptible de la briser.

²⁷ Ne pas confondre avec la constante de raideur !

L'arc de (10.35) (*resp.* son symétrique) initialement situé entre $\frac{\pi}{a}$ et $\frac{\pi}{2a}$ (*resp.* $-\frac{\pi}{a}$ et $-\frac{\pi}{2a}$) doit être décalé vers la gauche (*resp.* vers la droite) d'un multiple de $K = \frac{\pi}{2a}$, de façon à se retrouver entre $\pm\frac{\pi}{2a}$: on voit tout de suite qu'il faut le translater de $2\frac{\pi}{2a} \equiv 2K$; notant ω_{opt} cette branche, appelée branche optique²⁸, on a :

$$\omega_{\text{opt}}(k) = 2\omega_0 \left| \sin \frac{(k \mp \frac{\pi}{a})a}{2} \right| \equiv 2\omega_0 \left| \cos \frac{ka}{2} \right|, \quad k \in \left[-\frac{\pi}{2a}, +\frac{\pi}{2a} \right]. \quad (10.68)$$

Dans le schéma en zones réduites, la pulsation est maintenant *bivaluée*, les deux branches se rejoignant en $\pm\frac{\pi}{2a}$. Notamment, il y a deux modes $k = 0$. Celui de la branche acoustique représente toujours la translation en bloc et a toujours une énergie nulle ; en revanche, celui de la branche optique correspond à une vibration en opposition de phase des deux membres du dimère, tous les dimères vibrant en phase (voir (10.66)) ; son énergie n'est pas nulle, évidemment. Un mode optique se différencie notamment d'un mode acoustique par le fait que sa pulsation ne tend pas vers zéro aux grandes longueurs d'onde puisqu'il représente, dans le schéma en zones réduites, le mode où la vitesse de groupe s'annule et où deux atomes voisins vibrent en opposition de phase.

La description qui précède peut paraître artificielle lorsque les deux partenaires du dimère sont identiques, mais elle permet une transition vers le cas où ils sont, de fait, différents, une situation qui correspond à un vrai cristal unidimensionnel avec deux atomes par maille. Dans l'immédiat, on va se borner à différencier un atome sur deux du réseau de pas a (par leur masse, par exemple). Le traitement sera finalement étendu à une base diatomique où les partenaires sont à une distance d l'un de l'autre (en pratique, $d < a$).

On peut deviner ce qui va se passer, en raisonnant dans un cas limite. Puisque les atomes sont différents, les fréquences propres vont être différentes à l'intérieur d'un dimère (dont les deux masses sont différentes) ou entre deux dimères distincts (qui ont la même masse). À titre d'exemple, si les deux atomes sont seulement isotopes l'un de l'autre (masses M_1 et M_2 , $M_1 < M_2$), le potentiel harmonique – qui ne dépend que de la distribution électronique périphérique – sera le même sur tous les sites, donc la constante de raideur K sera la même d'un bout à l'autre du réseau. En revanche, on aura maintenant deux pulsations propres :

$$\omega_0 = \sqrt{\frac{K}{M_1}} \quad \Omega_0 = \sqrt{\frac{K}{M_2}} < \omega_0. \quad (10.69)$$

D'une façon générale, deux branches vont donc apparaître si deux masses sont en présence.

Il est facile de prévoir l'allure des lois de dispersion dans le cas extrême où $M_1 \ll M_2$. En effet, dans ce cas, chaque dimère a une masse à peu près égale à M_2 ; le réseau de dimères a un pas égal à $2a$. Donc, en considérant chacun d'entre eux comme une grosse particule, on voit qu'il existe des modes de type acoustique, la loi étant approximativement²⁹ donnée par :

$$\omega_{\text{ac}}(k) \simeq 2\frac{\Omega_0}{\sqrt{2}} \left| \sin \frac{k(2a)}{2} \right|, \quad k \in \left[-\frac{\pi}{2a}, +\frac{\pi}{2a} \right] \quad (M_1 \ll M_2); \quad (10.70)$$

c'est bien la pulsation $\frac{\Omega_0}{\sqrt{2}}$ qui doit être insérée dans (10.68), puisque chaque grosse particule a une masse à peu près égale à M_2 et que chacune des deux grosses masses est reliée à la petite par un ressort de constante K . Le ressort effectif³⁰ entre les deux grosses masses a donc une constante K_{eff} telle que $K_{\text{eff}}^{-1} = K^{-1} + K^{-1} = 2K^{-1}$, soit $K_{\text{eff}} = \frac{K}{2}$. La fréquence d'oscillation est donc $\Omega_{\text{eff}}^2 = \frac{K}{2}M_2$, soit $\Omega_{\text{eff}} = \frac{\Omega_0}{\sqrt{2}}$. La branche part de 0 en $k = 0$ et atteint sa valeur maximale $\sqrt{2}\Omega_0$ en $k = \frac{\pi}{2a}$.

Ceci définit une première classe de modes de vibration. Il existe aussi un mode de vibration intradimère, représentant cette fois une oscillation rapide de la petite masse M_1 rappelée par la grande masse M_2 quasiment immobile. En considérant un seul dimère, cette oscillation se fait strictement à la pulsation³¹ $\sqrt{2}\omega_0 \gg \Omega_0$, et a

²⁸ À nouveau, l'origine de la terminologie sera claire par la suite.

²⁹ On va évidemment trouver des corrections, puisque le rapport $\frac{M_2}{M_1}$ n'est pas infini. Noter que $M_1 \ll M_2$ est équivalent à $\Omega_0 \ll \omega_0$.

³⁰ Quand on met des ressorts en série, ce sont les inverses des constantes de raideur qui s'additionnent (au total, c'est plus "mou") ; au contraire, si les ressorts sont mis en parallèle, c'est plus "dur" et les constantes de raideur s'additionnent. Une chaîne 1-d infinie a une compressibilité infinie (elle est infiniment molle).

³¹ Pour une masse M_1 liée de part et d'autre à deux masses inertes ($u_{j\pm 1} = 0$), l'équation est $M_1\ddot{u}_j = -2Ku_j$, d'où le facteur $\sqrt{2}$.

donc une grande énergie³². En raison du couplage mécanique entre les petites masses³³, l'énergie correspondante dépendra de k , mais on s'attend à ce qu'elle soit peu dispersée autour de ω_0 – elle est donc relativement plate – et, en tout cas ressemble à un mode optique, tendant vers une valeur finie en $k = 0$.

En définitive, dans le cas de deux masses très différentes, $M_1 \ll M_2$, on attend deux branches, l'une optique (étroite), l'autre acoustique, qui ne se rejoignent plus en $k = \frac{\pi}{2a}$: un *gap* (relativement grand) s'est ouvert en bord de zone par comparaison au cas $M_1 = M_2$ traité comme ci-dessus (deux atomes identiques par maille) – tout comme dans le cas des électrons presque libres. Pour deux masses très voisines, $M_1 \lesssim M_2$, des arguments analogues montrent que les dispersions des deux branches sont comparables, et que le *gap* est cette fois tout petit.

Le calcul exact détaillé est facile, avec M_1 et M_2 quelconques (mais toujours $M_1 < M_2$ pour fixer les idées). On repart des équations du mouvement (10.20) adaptées au cas analysé maintenant ; supposons les petites masses M_1 situées sur les sites impairs, les grandes masses M_2 sur les sites pairs. Alors :

$$M_1 \ddot{u}_{2j+1} = -\frac{\partial U_h}{\partial u_{2j+1}} = -K (2u_{2j+1} - u_{2j} - u_{2j+2}) , \quad (10.71)$$

et :

$$M_2 \ddot{u}_{2j} = -\frac{\partial U_h}{\partial u_{2j}} = -K (2u_{2j} - u_{2j-1} - u_{2j+1}) . \quad (10.72)$$

Avec la même convention que précédemment (x_j est soit un déplacement classique $u(t)$, soit la représentation de Heisenberg $u_H(t)$), on obtient le double système (c'est la généralisation du système unique (10.26)) :

$$\ddot{x}_{2j+1} = -\omega_0^2 (2x_{2j+1} - x_{2j} - x_{2j+2}) , \quad (10.73)$$

et :

$$\ddot{x}_{2j} = -\Omega_0^2 (2x_{2j} - x_{2j-1} - x_{2j+1}) . \quad (10.74)$$

L'analyse de Fourier donne maintenant :

$$\omega^2 X_{2j+1} - \omega_0^2 (2X_{2j+1} - X_{2j} - X_{2j+2}) = 0 , \quad (10.75)$$

et :

$$\omega^2 X_{2j} - \Omega_0^2 (2X_{2j} - X_{2j-1} - X_{2j+1}) = 0 . \quad (10.76)$$

Tout naturellement, les petites masses sont couplées aux grandes masses et, à nouveau, il s'agit d'équations aux différences finies dont les solutions sont les puissances d'exponentielles complexes. Les X sont des amplitudes de Fourier des déplacements ; comme les déplacements sont d'autant plus faibles que l'inertie est grande, et qu'il y a deux masses, il convient d'introduire deux constantes de proportionnalité C_1 et C_2 pour les sites impairs et pairs respectivement. On pose donc :

$$X_{2j+1} = C_1 e^{i(2j+1)ka} \quad X_{2j} = C_2 e^{i2jka} . \quad (10.77)$$

Le report dans (10.75) et (10.76) donne le système :

$$(\omega^2 - 2\omega_0^2) C_1 + 2\omega_0^2 \cos ka C_2 = 0 , \quad (10.78)$$

$$2\Omega_0^2 \cos ka C_1 + (\omega^2 - 2\Omega_0^2) C_2 = 0 . \quad (10.79)$$

La solution est non-triviale si et seulement si le déterminant est nul, soit :

$$\omega^4 - 2(\omega_0^2 + \Omega_0^2) \omega^2 + 4\omega_0^2 \Omega_0^2 \sin^2 ka = 0 . \quad (10.80)$$

Les solutions en ω^2 sont :

$$\omega^2 = \omega_0^2 + \Omega_0^2 \pm \sqrt{(\omega_0^2 + \Omega_0^2)^2 - 4\omega_0^2 \Omega_0^2 \sin^2 ka} . \quad (10.81)$$

³²La relation entre haute fréquence et haute énergie tient tant dans le cas classique que dans le cas quantique. Dans le cas quantique, elle vient de $E = \hbar\omega$, évidemment. Dans le cas classique, elle vient du fait que l'énergie mécanique d'un oscillateur écarté de x_0 de sa position initiale est égale à $\frac{1}{2}m\omega_0^2 x_0^2$, d'autant plus grande que ω_0 est grand.

³³Comme les grosses masses M_2 ne sont pas totalement inertes, il existe évidemment un couplage résiduel entre les petites masses M_1 .

La quantité sous la racine carrée est aussi $(\omega_0^2 - \Omega_0^2)^2 + 4\omega_0^2\Omega_0^2 \cos^2 ka$: elle est bien positive). Par ailleurs, ω^2 est visiblement positive quel que soit le signe au second membre. L'équation (10.81) donne notamment les deux branches positives :

$$\omega_{\pm}(k) = \left[\omega_0^2 + \Omega_0^2 \pm \sqrt{(\omega_0^2 + \Omega_0^2)^2 - 4\omega_0^2\Omega_0^2 \sin^2 ka} \right]^{1/2}, \quad (10.82)$$

ou encore :

$$\omega_{\pm}(k) = \left[\omega_0^2 + \Omega_0^2 \pm \sqrt{(\omega_0^2 - \Omega_0^2)^2 + 4\omega_0^2\Omega_0^2 \cos^2 ka} \right]^{1/2}. \quad (10.83)$$

En $k = 0$, elles valent $[\omega_0^2 + \Omega_0^2 \pm (\omega_0^2 + \Omega_0^2)]^{1/2} = 0, \sqrt{2(\omega_0^2 + \Omega_0^2)}$; plus précisément, un développement limité donne ($ka \simeq 0$) :

$$\omega_-(k) \simeq \sqrt{2 \frac{\omega_0^2 \Omega_0^2}{\omega_0^2 + \Omega_0^2}} ka, \quad \omega_+(k) \simeq \sqrt{2(\omega_0^2 + \Omega_0^2)} \left[1 - \frac{\omega_0^2 \Omega_0^2}{2(\omega_0^2 + \Omega_0^2)^2} k^2 a^2 \right] \quad (10.84)$$

ω_- est la branche acoustique (elle s'annule linéairement en k), ω_+ est la branche optique, parabolique près de $k = 0$.

Pour examiner ce qui se passe près du bord de zone, $k \simeq \frac{\pi}{2a}$, on part de :

$$\omega_{\pm} = \omega_0^2 + \Omega_0^2 \pm \sqrt{(\omega_0^2 - \Omega_0^2)^2 + 4\omega_0^2\Omega_0^2 \sin^2(\frac{\pi}{2} - ka)}; \quad (10.85)$$

avec $\omega_0 > \Omega_0$; le développement limité donne :

$$\omega_+(k) \simeq \sqrt{2}\omega_0 \left[1 + \frac{\omega_0^2 \Omega_0^2}{2(\omega_0^2 - \Omega_0^2)^2} (\frac{\pi}{2} - ka)^2 \right], \quad (10.86)$$

$$\omega_-(k) \simeq \sqrt{2}\Omega_0 \left[1 - \frac{\omega_0^2 \Omega_0^2}{2(\omega_0^2 - \Omega_0^2)^2} (\frac{\pi}{2} - ka)^2 \right]. \quad (10.87)$$

Les deux branches sont donc à peu près paraboliques en bord de zone et séparées par le *gap* :

$$(\omega_+ - \omega_-)(k = \frac{\pi}{2a}) = \sqrt{2}(\omega_0 - \Omega_0), \quad (10.88)$$

qui s'annule bien si on en revient à deux masses identiques. À l'inverse, si $M_1 \ll M_2$ (soit $\omega_0 \gg \Omega_0$), (10.82) montre que la branche optique est peu dispersée ; en effet, on trouve alors :

$$\omega_+(k) \simeq \sqrt{2}\omega_0^2 \left[1 + \frac{\Omega_0^2}{\omega_0^2} \cos^2 ka \right]. \quad (10.89)$$

Comme $\frac{\Omega_0}{\omega_0} \ll 1$, la variation induite par $\cos^2 ka$ est faible : la bande optique est presque plate.

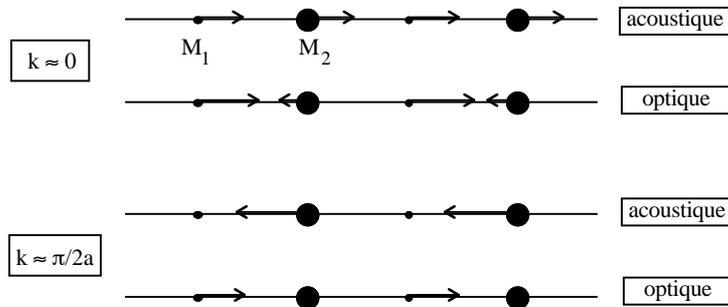


Figure 10.1: Représentation schématique des amplitudes de vibration pour un réseau à deux atomes par maille.

Une fois obtenues les valeurs propres, il est possible d'analyser les mouvements respectifs des différents atomes dans chacun des modes obtenus (voir fig. 10.1). Il suffit pour cela d'examiner le rapport $\frac{C_1}{C_2}$, donné (par exemple) par (10.78) :

$$\left(\frac{C_1}{C_2}\right)_{\pm} = -\frac{2\omega_0^2 \cos ka}{\omega_{\pm}^2 - 2\omega_0^2} = \frac{2 \cos ka}{1 - \mu \mp \sqrt{(1 + \mu)^2 - 4\mu \sin^2 ka}} , \quad \mu = \frac{\Omega_0^2}{\omega_0^2} = \frac{M_1}{M_2} \leq 1 . \quad (10.90)$$

En utilisant les développements limités de $\omega_{\pm}(k)$, il est facile de voir que :

- $k \rightarrow 0$

$$\text{mode optique : } \left(\frac{C_1}{C_2}\right)_{+} \simeq -\frac{1}{\mu} \left[1 - \frac{1 - \mu}{2(1 + \mu)} (ka)^2\right] = -\frac{1}{\mu} [1 + \mathcal{O}(k^2 a^2)] . \quad (10.91)$$

Les deux atomes vibrent en opposition de phase – ce que l'on attend d'un mode optique, avec une amplitude inversement proportionnelle à leur masse ; on a $C_1 > C_2$ si $M_1 < M_2$ (inertie). Pour le mode acoustique ω_{-} , on trouve tout naturellement que les deux amplitudes sont du même ordre de grandeur par rapport au rapport des masses :

$$\text{mode acoustique : } \left(\frac{C_1}{C_2}\right)_{-} \simeq 1 - \frac{1 - \mu}{2(1 + \mu)} (ka)^2 . \quad (10.92)$$

- $k \rightarrow \frac{\pi}{2a}$

$\omega_{+}^2 - 2\omega_0^2$ s'annule quadratiquement par rapport à $\epsilon = k - \frac{\pi}{2a} < 0$. Au contraire, le numérateur de (10.90) s'annule linéairement. Il en résulte que le rapport $\left(\frac{C_1}{C_2}\right)_{+}$ diverge ; très précisément, un développement limité donne :

$$\text{mode optique : } \left(\frac{C_1}{C_2}\right)_{+} \simeq \left(\frac{1}{\mu} - 1\right) \left(k - \frac{\pi}{2a}\right)^{-1} \rightarrow -\infty . \quad (10.93)$$

Dans le mode optique, près du bord de zone, l'atome lourd est presque immobile et seul l'atome léger vibre avec une amplitude finie (comme le montre le signe $-$, le mouvement est toujours en opposition de phase). C'est un mode de haute énergie, correspondant pratiquement à un mode de vibration intramoléculaire.

Pour ω_{-} (mode acoustique), il en va tout autrement. Maintenant, le dénominateur a une limite finie. Il en résulte :

$$\text{mode acoustique : } \left(\frac{C_1}{C_2}\right)_{-} \simeq \frac{1}{1 - \mu} \left(\frac{\pi}{2} - ka\right) > 0 . \quad (10.94)$$

Cette fois, c'est l'atome massif qui a une grande amplitude, cependant que, relativement, l'atome léger est à peu près fixe ; l'énergie de vibration est nettement plus petite que pour le mode optique.

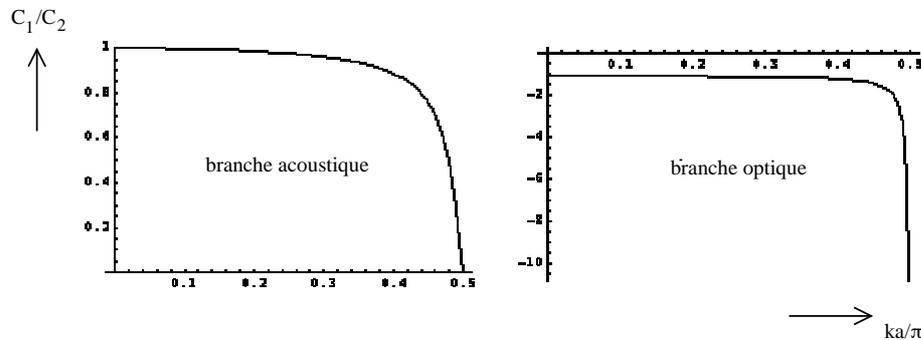
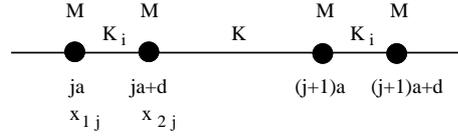


Figure 10.2: Rapport des amplitudes de vibration $\frac{C_1}{C_2}$ des deux masses M_1 et M_2 (voir (10.90)) ; les courbes sont tracées pour $\mu \stackrel{\text{d'ef}}{=} \frac{M_1}{M_2} = 0.9$.

Le cas précédent a été analysé dans la lignée du réseau monoatomique, en alternant les masses et en définissant donc un réseau de maille deux fois plus grande avec une base. Comme dernier exemple, on considère

Figure 10.3: Chaîne linéaire à deux atomes par maille (même masse M).

un réseau de paramètre³⁴ a , avec une base diatomique constituée de deux atomes de même masse M séparés (à l'équilibre) de la distance $d < a$ (molécule homonucléaire) (voir fig. 10.3). Il est clair physiquement que les constantes de raideur sont maintenant différentes : K_i pour le ressort entre les deux atomes séparés de d (vibration intramoléculaire), K pour deux atomes appartenant à deux molécules adjacentes ; afin de fixer les idées, on suppose $K_i > K$, ce qui est physiquement bien naturel. On désigne par $x_{1,j}$ l'écart de l'atome qui vibre autour de ja , $x_{2,j}$ l'écart de celui qui est près de $ja + d$. Les équations du mouvement sont :

$$M\ddot{x}_{1,j} = -K(x_{1,j} - x_{2,j-1}) - K_i(x_{1,j} - x_{2,j}) = -(K + K_i)x_{1,j} + Kx_{2,j-1} + K_ix_{2,j} , \quad (10.95)$$

$$M\ddot{x}_{2,j} = -K_i(x_{2,j} - x_{1,j}) - K(x_{2,j} - x_{1,j+1}) = -(K_i + K)x_{2,j} + K_ix_{1,j} + Kx_{1,j+1} . \quad (10.96)$$

Pour les composantes de Fourier, on a :

$$[M\omega^2 - (K + K_i)] X_{1,j} + K X_{2,j-1} + K_i X_{2,j} = 0 , \quad (10.97)$$

et :

$$[M\omega^2 - (K_i + K)] X_{2,j} + K_i X_{1,j} + K X_{1,j+1} = 0 . \quad (10.98)$$

Posant :

$$\omega_0^2 = \frac{K_i + K}{M} , \quad \Omega_0^2 = \frac{K}{M} , \quad \Omega_i^2 = \frac{K_i}{M} , \quad \Omega_0^2 + \Omega_i^2 = 2\omega_0^2 , \quad (10.99)$$

il vient :

$$(\omega^2 - 2\omega_0^2) X_{1,j} + \Omega_0^2 X_{2,j-1} + \Omega_i^2 X_{2,j} = 0 , \quad (10.100)$$

et :

$$(\omega^2 - 2\omega_0^2) X_{2,j} + \Omega_i^2 X_{1,j} + \Omega_0^2 X_{1,j+1} = 0 . \quad (10.101)$$

Avec :

$$X_{1,j} = C_1 e^{ijk a} , \quad X_{2,j} = C_2 e^{ijk a} , \quad (10.102)$$

les deux équations (10.100) et (10.101) donnent le système homogène :

$$(\omega^2 - 2\omega_0^2) C_1 + \Omega_i^2 C_2 e^{-ika} + \Omega_0^2 C_2 = 0 , \quad (\omega^2 - 2\omega_0^2) C_2 + \Omega_i^2 C_1 + \Omega_0^2 C_1 e^{ika} = 0 . \quad (10.103)$$

La condition d'annulation du déterminant est :

$$\omega^2 = \Omega_0^2 + \Omega_i^2 \pm (\Omega_0^4 + 2\Omega_0^2\Omega_i^2 \cos ka + \Omega_i^4)^{1/2} = \frac{1}{M} \left\{ K + K_i \pm [(K + K_i)^2 - 4KK_i \sin^2 \frac{ka}{2}]^{1/2} \right\} , \quad (10.104)$$

et donne à nouveau une branche acoustique et une branche optique. Au voisinage des points remarquables, on a :

- $k \rightarrow 0$

$$\omega_+ \simeq \sqrt{2\frac{K + K_i}{M}} \left[1 - \frac{KK_i}{8(K + K_i)^2} k^2 a^2 \right] , \quad \omega_- \simeq \sqrt{\frac{KK_i}{2M(K + K_i)}} ka . \quad (10.105)$$

³⁴Si on veut comparer les expressions obtenues ci-dessous avec celles trouvées pour le réseau à deux masses, il faut se souvenir qu'alors la maille avait pour côté $2a$.

- $k \rightarrow \frac{\pi}{2a}$ et si $K_i > K$

$$\omega_+ \simeq \sqrt{2\frac{K_i}{M}} \left[1 + \frac{K}{2(K_i - K)} \left(\frac{\pi}{2} - ka \right)^2 \right], \quad \omega_- \simeq \sqrt{2\frac{K}{M}} \left[1 - \frac{K_i}{2(K_i - K)} \left(\frac{\pi}{2} - ka \right)^2 \right]. \quad (10.106)$$

Le gap en bord de zone est donc :

$$(\omega_+ - \omega_-)_{k=\frac{\pi}{2a}} = \sqrt{2(\Omega_i^2 - \Omega_0^2)}. \quad (10.107)$$

Notons que dans tous les cas, les déplacements s'écrivent finalement comme des combinaisons linéaires d'ondes planes $e^{i[\vec{k} \cdot n\vec{a} - \omega(\vec{k})t]}$ où $n\vec{a}$ est un vecteur de \mathcal{B} et où $\omega(\vec{k})$ est la loi de dispersion.

10.4 Modes normaux d'un réseau tridimensionnel

Les vibrations d'un réseau tridimensionnel s'obtiennent exactement comme précédemment, en tenant compte simplement du fait que le déplacement de chaque atome est déterminé par un vecteur $\vec{u}(\vec{R})$.

10.4.1 Réseau sans base

Lorsque le réseau est sans base, les composantes de $\vec{u}(\vec{R})$ obéissent aux équations du mouvement (10.15). Tout comme avant, des solutions particulières (complètes) peuvent s'écrire sous la forme d'ondes planes, dont il faut maintenant de surcroît préciser la polarisation $\vec{\varepsilon}'$:

$$\vec{f}(\vec{R}, t) = \vec{\varepsilon}' e^{i(\vec{k} \cdot \vec{R} - \omega t)}. \quad (10.108)$$

Pour chaque composante ε_α de $\vec{\varepsilon}'$, le report de cette expression dans (10.15) donne :

$$-M\omega^2 \varepsilon_\alpha e^{i\vec{k} \cdot \vec{R}} = -\sum_{\beta} \sum_{\vec{R}'} Q_{\alpha\beta}(\vec{R}') \varepsilon_\beta e^{i\vec{k} \cdot \vec{R}'}. \quad (10.109)$$

Pour un réseau macroscopique, ceci se réécrit comme suit :

$$M\omega^2 \varepsilon_\alpha = \sum_{\beta} \sum_{\vec{R}''} Q_{\alpha\beta}(\vec{R}'') \varepsilon_\beta e^{-i\vec{k} \cdot \vec{R}''} \equiv \sum_{\beta} Q_{\alpha\beta}(\vec{k}) \varepsilon_\beta, \quad (10.110)$$

ce que l'on peut aussi écrire sous forme matricielle :

$$M\omega^2 \vec{\varepsilon}' = \mathcal{Q}(\vec{k}) \vec{\varepsilon}'. \quad (10.111)$$

$\mathcal{Q}(\vec{k})$ est appelé matrice dynamique. Si N désigne toujours le nombre de nœuds (atomes), il existe maintenant $3N$ modes normaux distincts³⁵. En vertu des symétries de la matrice Q (voir (10.14)), on a :

$$\mathcal{Q}(\vec{k}) \equiv \frac{1}{2} \left[\sum_{\vec{R}} Q(\vec{R}) e^{-i\vec{k} \cdot \vec{R}} + \sum_{\vec{R}} Q(-\vec{R}) e^{+i\vec{k} \cdot \vec{R}} \right] = \sum_{\vec{R}} Q(\vec{R}) \cos \vec{k} \cdot \vec{R}. \quad (10.112)$$

Par ailleurs, compte tenu de (10.11), ceci vaut aussi :

$$\mathcal{Q}(\vec{k}) = \sum_{\vec{R}} Q(\vec{R}) \cos \vec{k} \cdot \vec{R} - \sum_{\vec{R}} Q(\vec{R}) = -2 \sum_{\vec{R}} Q(\vec{R}) \sin^2 \frac{1}{2} \vec{k} \cdot \vec{R}. \quad (10.113)$$

³⁵On admet que le cristal ne peut pas tourner – pas de degrés de liberté de rotation. Les trois modes de translation sont inclus dans la suite ($\vec{k} = \vec{0}$). De toute façon, avec $N \gg 1$, $3N$ ou $3N - 6$ font le même effet.

$\mathcal{Q}(\vec{k})$ est donc une combinaison réelle de matrices $Q(\vec{R})$ symétriques réelles de dimension 3, c'est donc une matrice symétrique réelle de dimension 3. En tant que telle, elle a 3 vecteurs propres réels (orthogonaux) satisfaisant :

$$\mathcal{Q}(\vec{k}) \vec{\varepsilon}_\lambda = q_\lambda(\vec{k}) \vec{\varepsilon}_\lambda \quad (\lambda = 1, 2, 3) ; \quad (10.114)$$

dans la suite, on supposera éventuellement que ces vecteurs sont de surcroît normalisés, de sorte que :

$$\vec{\varepsilon}_\lambda \cdot \vec{\varepsilon}_{\lambda'} = \delta_{\lambda\lambda'} . \quad (10.115)$$

Il y a trois polarisations distinctes (et orthogonales) pour un même vecteur³⁶ \vec{k} . Les pulsations propres $\omega_\lambda(\vec{k})$ sont données par :

$$M \omega_\lambda^2(\vec{k}) = q_\lambda(\vec{k}) \iff \omega_\lambda(\vec{k}) = \sqrt{\frac{q_\lambda(\vec{k})}{M}} . \quad (10.116)$$

Tout comme le réseau monoatomique unidimensionnel, les trois branches s'annulent linéairement en $\vec{k} = 0$. Dans cette limite :

$$\mathcal{Q}(\vec{k}) \simeq -\frac{1}{2} \sum_{\vec{R}} Q(\vec{R}) (\vec{k} \cdot \vec{R})^2 \equiv -\frac{1}{2} k^2 \sum_{\vec{R}} Q(\vec{R}) (\vec{\kappa} \cdot \vec{R})^2 , \quad k \rightarrow 0 \quad (10.117)$$

où $\vec{\kappa}$ est le vecteur unitaire porté par \vec{k} . Conformément à (10.116), les pulsations sont dans cette limite approximativement données par les racines carrées des valeurs propres de la matrice $-\frac{1}{2M} \sum_{\vec{R}} Q(\vec{R}) (\vec{\kappa} \cdot \vec{R})^2$; elles dépendent de la direction de \vec{k} par l'intermédiaire de $\vec{\kappa}$. Les notant $c_\lambda(\vec{\kappa})$, il vient :

$$\omega_\lambda(\vec{k}) \simeq k c_\lambda(\vec{\kappa}) , \quad k \rightarrow 0 . \quad (10.118)$$

$c_\lambda(\vec{\kappa})$ est la vitesse (anisotrope) de propagation du son dans le cristal.

10.4.2 Réseau avec base

Lorsque le réseau contient une base ("molécule") à n atomes, il existe au total $3nN$ modes de vibrations, c'est-à-dire $3n$ modes pour chaque valeur de \vec{k} . Parmi ces derniers, il y en a forcément 3 dont la fréquence tend vers zéro à petit k , donnant à la limite la translation d'énergie nulle. Il reste donc $3(n-1)$ modes, qui sont de nature optique, leur fréquence ayant une valeur finie en $\vec{k} = 0$. Ils sont en fait issus des $3(n-1)$ modes de vibration-rotation d'une molécule³⁷, qui présentent une dispersion en \vec{k} compte tenu de l'interaction entre les molécules constituant le cristal.

10.5 Quantification des vibrations de réseau

La quantification des vibrations harmoniques du réseau est immédiate. Il suffit de revenir au Hamiltonien (10.2) où U est remplacé par la forme quadratique (10.8). Les $3N$ composantes $u_\alpha(\vec{R})$ – pour un réseau sans base – peuvent être exprimées à l'aide des déplacements normaux $\xi_\lambda(\vec{k})$ qui diagonalisent la matrice dynamique (10.111). La transformation correspondante est unitaire, de sorte que la partie cinétique T de (10.2) est invariante en forme et devient :

$$T = \sum_{\lambda, \vec{k}} \frac{\Pi_{\lambda \vec{k}}^2}{2M} , \quad \Pi_{\lambda \vec{k}} = -i\hbar \frac{\partial}{\partial \xi_{\lambda \vec{k}}} , \quad (10.119)$$

³⁶La disposition relative des polarisations $\vec{\varepsilon}_\lambda$ et de \vec{k} dépend de la symétrie (ponctuelle) du cristal. Pour un cristal isotrope, pour chaque \vec{k} , on peut toujours choisir les solutions de sorte qu'il existe un mode longitudinal et deux modes transverses. Lorsque le cristal est anisotrope, il n'en va plus ainsi, sauf si \vec{k} est lui-même invariant dans certaines opérations de symétrie.

³⁷Une molécule à n atomes possède $3n$ degrés de liberté nucléaires. 3 d'entre eux représentent une translation en bloc de la molécule, tous les autres sont associés au mouvement du squelette moléculaire et peuvent se subdiviser en deux : la rotation en bloc de la molécule (autour d'une configuration fixée des noyaux, usuellement la configuration stable) qui exige le recours à 3 angles d'Euler pour une molécule quelconque, 2 seulement si la molécule est linéaire. En conséquence, il reste $3n-6$ ou $3n-5$ degrés de liberté de vibration, qui correspondent à des déformations de la molécule.

cependant que la partie potentielle devient par construction :

$$U_h = \sum_{\lambda, \vec{k}} \frac{1}{2} M \omega_\lambda^2(\vec{k}) \xi_\lambda^2(\vec{k}) . \quad (10.120)$$

Les degrés de liberté normaux $\xi_\lambda(\vec{k})$ et les moments conjugués $\Pi_{\lambda \vec{k}}$ satisfont les relations de commutation canoniques :

$$[\xi_{\lambda \vec{k}}, \Pi_{\lambda' \vec{k}'}] = i\hbar \delta_{\lambda \lambda'} \delta_{\vec{k} \vec{k}'} , \quad (10.121)$$

en conséquence du fait que le passage des $\{u\}$ aux $\{\xi\}$ s'effectue par une transformation unitaire.

Il en résulte que H est maintenant une *somme* de $3N$ Hamiltoniens d'oscillateurs harmoniques :

$$H = \sum_{\lambda, \vec{k}} \left[\frac{\Pi_{\lambda \vec{k}}^2}{2M} + \frac{1}{2} M \omega_\lambda^2(\vec{k}) \xi_\lambda^2(\vec{k}) \right] . \quad (10.122)$$

L'additivité de H signifie que chacun des oscillateurs est indépendant de tous les autres. De ce fait, la construction des états de vibration du cristal s'effectue en formant les produits des fonctions propres d'oscillateurs harmoniques $\psi_{\lambda \vec{k}}$, associées aux énergies $E_\lambda(\vec{k}) = \hbar \omega_\lambda(\vec{k})$. Chaque *quantum* d'excitation est appelé phonon ; si le mode (λ, \vec{k}) est au fondamental, il contient zéro phonon, s'il est dans le premier état excité, il y a un phonon, *etc.* Les phonons sont au champ des vibrations de réseau ce que les photons sont au champ électromagnétique³⁸. Notons qu'un phonon acoustique de tout petit \vec{k} a, comme le photon, une relation de dispersion linéaire.

En définitive, l'état vibrationnel du cristal est complètement spécifié en donnant le nombre $n_{\lambda \vec{k}}$ de phonons dans chaque mode. Comme rien n'interdit un état de vibration arbitrairement excité pour un mode donné, $n_{\lambda \vec{k}}$ pouvant prendre n'importe quelle valeur $\in \mathbb{N}$, on peut caser un nombre arbitraire de phonons dans un seul mode : du point de vue de la statistique, les phonons sont des bosons³⁹. Si le mode (λ, \vec{k}) contient $n_{\lambda \vec{k}}$ phonons, il a l'énergie $\hbar \omega_\lambda(\vec{k})(n_{\lambda \vec{k}} + \frac{1}{2})$. Pour une distribution des nombres d'occupation $n_{\lambda \vec{k}}$ sur l'ensemble des modes, l'énergie de vibration totale du réseau est :

$$E = \sum_{\lambda, \vec{k}} \hbar \omega_\lambda(\vec{k}) \left(n_{\lambda \vec{k}} + \frac{1}{2} \right) . \quad (10.125)$$

Le demi-quantum est une simple constante additive⁴⁰, délaissée dans la suite puisque seules comptent les *différences* d'énergie. L'expression retenue pour l'énergie totale de vibration est ainsi :

$$E = \sum_{\lambda, \vec{k}} n_{\lambda \vec{k}} \hbar \omega_\lambda(\vec{k}) . \quad (10.126)$$

10.6 Contribution des phonons à la chaleur spécifique

Pour trouver la chaleur spécifique due aux phonons, il suffit d'écrire l'expression de la valeur moyenne de l'énergie du réseau en équilibre thermique à la température T . La distribution des nombres moyens d'occupation est celle de bosons, donnée par la statistique de Bose - Einstein :

$$n_{\lambda \vec{k}} = \frac{1}{e^{\beta \hbar \omega_\lambda(\vec{k})} - 1} , \quad \beta = \frac{1}{k_B T} . \quad (10.127)$$

³⁸Toutefois, au contraire du champ électromagnétique, le champ de phonons contient un mode propagatif longitudinal.

³⁹Ceci se voit encore plus directement en introduisant les opérateurs d'annihilation et de création pour chaque mode. Ces opérateurs satisfont :

$$[b_{\lambda \vec{k}}, b_{\lambda' \vec{k}'}^\dagger] = \delta_{\lambda \lambda'} \delta_{\vec{k} \vec{k}'} . \quad (10.123)$$

La commutation de tels opérateurs est caractéristique des bosons. Pour des fermions, on a des relations d'anticommutation :

$$\{c_{\lambda \vec{k}}, c_{\lambda' \vec{k}'}^\dagger\} = \delta_{\lambda \lambda'} \delta_{\vec{k} \vec{k}'} , \quad (10.124)$$

où $\{A, B\} \equiv AB + BA$.

⁴⁰éventuellement infinie...

Il en résulte immédiatement que l'énergie moyenne (au sens de la statistique quantique) à la température T est donnée par (10.125) où l'expression (10.127) est utilisée. Si V désigne le volume de l'échantillon, la chaleur spécifique due aux phonons est⁴¹ :

$$c_{V \text{ ph}} = \frac{1}{V} \frac{\partial E}{\partial T} = \frac{1}{V} \frac{\partial}{\partial T} \sum_{\lambda, \vec{k}} n_{\lambda \vec{k}} \hbar \omega_{\lambda}(\vec{k}) = \frac{1}{V} \frac{\partial}{\partial T} \sum_{\lambda, \vec{k}} \frac{\hbar \omega_{\lambda}(\vec{k})}{e^{\beta \hbar \omega_{\lambda}(\vec{k})} - 1} . \quad (10.128)$$

À très haute température ($\beta \hbar \omega_{\lambda}(\vec{k}) \ll 1$) :

$$c_{V \text{ ph}} \simeq \frac{1}{V} \frac{\partial}{\partial T} \sum_{\lambda, \vec{k}} \hbar \omega_{\lambda}(\vec{k}) (\beta \hbar \omega_{\lambda}(\vec{k}))^{-1} = \frac{1}{V} \frac{\partial}{\partial T} \sum_{\lambda, \vec{k}} k_{\text{B}} T = \frac{3N}{V} k_{\text{B}} . \quad (10.129)$$

C'est le résultat classique, appelé loi de Dulong et Petit. Comme toujours, la limite haute température efface le *quantum* porté par \hbar et la physique devient classique⁴². Les corrections quantiques sont faciles à écrire à partir de (10.128), mais sont souvent non-significatives car du même ordre de grandeur que les anharmonicités, ignorées d'un bout à l'autre.

L'autre limite, basse température, est évidemment plus intéressante puisque révélatrice des effets quantiques. Il est plus commode de discuter la forme intégrale de (10.128) ; comme $\delta \vec{k} = \left(\frac{2\pi}{L}\right)^3 = \frac{(2\pi)^3}{V}$, il vient :

$$c_{V \text{ ph}} = \frac{\partial}{\partial T} \sum_{\lambda} \int_{\text{BZ1}} \frac{d^3 k}{(2\pi)^3} \frac{\hbar \omega_{\lambda}(\vec{k})}{e^{\beta \hbar \omega_{\lambda}(\vec{k})} - 1} . \quad (10.130)$$

En se plaçant à température assez basse, on peut négliger tous les modes optiques éventuels⁴³, qui donnent des contributions exponentiellement petites par rapport à celle des modes acoustiques. Par ailleurs, en supposant de plus $k_{\text{B}} T \ll \hbar k c_{\lambda}$ pour toutes les branches acoustiques retenues, on peut remplacer les vraies lois de dispersion $\hbar \omega_{\lambda}(\vec{k})$ par leurs expressions approchées à petit \vec{k} : à nouveau, la substitution n'introduit que des erreurs exponentiellement petites⁴⁴. Enfin, dans le droit-fil, on rejette à l'infini les bornes d'intégration. En conséquence, une expression admissible de $c_{V \text{ ph}}$ est :

$$c_{V \text{ ph}} = \frac{\partial}{\partial T} \sum_{\text{acoustiques}} \int_{\mathbb{R}^3} \frac{d^3 k}{(2\pi)^3} \frac{\hbar k c_{\text{ac}}(\vec{k})}{e^{\beta \hbar k c_{\text{ac}}(\vec{k})} - 1} , \quad (10.131)$$

où la somme ne porte que sur les modes acoustiques, caractérisés par les différentes vitesses $c_{\text{ac}}(\vec{k})$, celles-ci ne dépendant que de la direction de \vec{k} , pas de son module. En séparant les dépendances radiale et angulaire, chaque intégrale de la somme se transforme comme suit ($d^3 k = k^2 dk d\Omega$) :

$$\int_0^{+\infty} \frac{k^2 dk}{2\pi^2} \int_{4\pi} \frac{d\Omega(\vec{k})}{4\pi} \frac{\hbar k c_{\text{ac}}(\vec{k})}{e^{\beta \hbar k c_{\text{ac}}(\vec{k})} - 1} ; \quad (10.132)$$

comme $c_{\text{ac}}(\vec{k})$ ne dépend que de l'orientation de \vec{k} , on peut extraire la fraction contenant la température de l'intégrale sur $d\Omega$. En posant $x = \beta \hbar k c_{\text{ac}}(\vec{k})$, il vient ainsi :

$$\frac{1}{2\pi^2 \beta^4 \hbar^3} \int_0^{+\infty} \frac{x^3 dx}{e^x - 1} \int_{4\pi} \frac{d\Omega(\vec{k})}{4\pi} \frac{1}{[c_{\text{ac}}(\vec{k})]^3} . \quad (10.133)$$

C'est le $\beta^{-4} \propto T^4$ qui va donner par dérivation $c_{V \text{ ph}} \propto T^3$. Dans les détails, on peut s'y prendre comme suit ; en introduisant la moyenne angulaire de $\frac{1}{c_{\text{ac}}^3}$:

$$\frac{1}{c^3} \equiv \frac{1}{3} \sum_{\text{acoustiques}} \int_{4\pi} \frac{d\Omega(\vec{k})}{4\pi} \frac{1}{[c_{\text{ac}}(\vec{k})]^3} , \quad (10.134)$$

⁴¹On remarque que la constante additive d'écart entre (10.125) et (10.126) ne joue aucun rôle pour $c_{V \text{ ph}}$.

⁴²Quoique, comme toujours, l'approche est singulière.

⁴³présents s'il existe un motif.

⁴⁴L'exponentielle $e^{\beta \hbar \omega_{\lambda}(\vec{k})}$ du dénominateur agit comme une fonction de coupure aux grands k et ne retient que les vecteurs d'onde tels que $\hbar \omega_{\lambda}(\vec{k}) \lesssim k_{\text{B}} T$.

et en utilisant :

$$\int_0^{+\infty} \frac{x^3 dx}{e^x - 1} = 3! \sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{15}, \quad (10.135)$$

on trouve finalement :

$$c_{V \text{ ph}} = \frac{\pi^2 k_B^4}{10 \hbar^3 c^3} \frac{\partial}{\partial T} T^4 = \frac{2\pi^2}{5} \left(\frac{k_B T}{\hbar c} \right)^3 k_B. \quad (10.136)$$

À basse température, la contribution des phonons à la chaleur spécifique varie bien essentiellement comme T^3 . Toutefois, ce comportement n'est observé qu'à *très* basse température : comme les lois de dispersion s'éloignent assez vite de la linéarité, il existe en général un grand intervalle de température où, la loi de Dulong et Petit étant déjà invalide, la loi en T^3 n'est pas pour autant vérifiée.

Pour un métal, cette variation typique en T^3 s'ajoute à la contribution du gaz de Fermi constitué par les électrons, qui donne un comportement linéaire en température pour un métal : seuls les électrons situés au niveau de Fermi sont susceptibles d'être excités thermiquement, et leur nombre est d'ordre $\mathcal{D}(\varepsilon_F) k_B T$. La variation d'énergie par fluctuation thermique est donc d'ordre $\mathcal{D}(\varepsilon_F) k_B T \times k_B T$; par dérivation, on en déduit $c_{V \text{ el}} \propto T$. Au total, la chaleur spécifique à basse température est une loi du genre :

$$c_V = (\gamma T + AT^3) k_B, \quad (10.137)$$

où γ est appelé "constante de Sommerfeld" (qui contient la température de Fermi T_F du gaz d'électrons).

10.7 Manifestation des phonons dans les spectres de diffraction

La théorie élémentaire de la diffraction exposée dans un chapitre précédent repose sur l'hypothèse d'un cristal parfait (pas de défauts, ni structurels, ni chimiques), dont les objets physiques sont immobiles. La conclusion majeure alors obtenue, en ne considérant que la diffusion élastique, est l'apparition d'une tache ponctuelle de diffraction à chaque fois que le transfert de moment $\vec{q} = \vec{k}_i - \vec{k}_f$ coïncide avec un vecteur \vec{K} du réseau réciproque.

Bien évidemment, cette conclusion doit être reconsidérée si les objets diffuseurs ont la liberté de se mouvoir, et notamment s'ils effectuent des petites oscillations autour d'une position d'équilibre, définie par le réseau de Bravais. Cette nécessité s'impose quand on se souvient que c'est précisément l'ordre à longue distance, parfaitement déterminé pour un réseau sans défauts et immobile, qui produit l'effet coopératif : l'action collective d'un grand nombre de diffuseurs produit une hypersélectivité donnant lieu à des taches ponctuelles de diffraction. Par ailleurs, lorsqu'il s'agit de diffusion neutronique, il y a maintenant possibilité de diffusion *inélastique* avec création ou annihilation de phonons⁴⁵, puisque les bons neutrons ont une énergie thermique et que, dès lors, on peut satisfaire les deux lois de conservation énergie et impulsion⁴⁶. Afin d'en rester au niveau élémentaire, on continue toutefois à ne considérer que la diffusion élastique.

L'objet central est l'amplitude de diffusion, dont on a vu qu'elle contenait le facteur décisif :

$$S = \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{q} \cdot \vec{R}}, \quad \vec{q} = \vec{k}_i - \vec{k}_f. \quad (10.138)$$

Dans le cas où les atomes effectuent des petites vibrations, compte tenu des notations introduites en (10.1), on doit maintenant écrire :

$$S = \sum_{\vec{R} \in \mathcal{B}} e^{i\vec{q} \cdot [\vec{R} + \vec{u}(\vec{R})]}. \quad (10.139)$$

L'intensité de diffraction est proportionnelle à $|S|^2$

$$|S|^2 = \sum_{\vec{R}, \vec{R}'} e^{i\vec{q} \cdot [\vec{R} + \vec{u}(\vec{R})]} e^{-i\vec{q} \cdot [\vec{R}' + \vec{u}(\vec{R}')] } = \sum_{\vec{R}, \vec{R}'} e^{i\vec{q} \cdot (\vec{R} - \vec{R}')} e^{i\vec{q} \cdot [\vec{u}(\vec{R}) - \vec{u}(\vec{R}')] }; \quad (10.140)$$

⁴⁵ En ce qui concerne l'irradiation X, on continue à laisser de côté les diffusions inélastiques à très haute énergie (~ 10 keV).

⁴⁶ D'ailleurs, la diffusion inélastique des neutrons est une méthode puissante pour obtenir expérimentalement les lois de dispersion des phonons.

ceci est l'amplitude pour une configuration donnée des écarts mais ce que l'on observe, c'est l'amplitude *moyenne* à une température donnée T . Ainsi, pour un réseau en équilibre thermique, la quantité importante à trouver est :

$$\langle |S|^2 \rangle = \sum_{\vec{R}, \vec{R}'} e^{i\vec{q} \cdot (\vec{R} - \vec{R}')} \langle e^{i\vec{q} \cdot [\vec{u}(\vec{R}) - \vec{u}(\vec{R}')] } \rangle , \quad (10.141)$$

où les crochets représentent la moyenne prise au sens de la statistique (classique à très haute température, forcément quantique autrement). La moyenne de droite est évidemment invariante dans toute translation du réseau, de sorte que :

$$\langle |S|^2 \rangle = N \sum_{\vec{R}} e^{-i\vec{q} \cdot \vec{R}} \langle e^{i\vec{q} \cdot \vec{u}(\vec{0})} e^{-i\vec{q} \cdot \vec{u}(\vec{R})} \rangle ; \quad (10.142)$$

en définitive, il faut en premier lieu obtenir la quantité :

$$Q = \langle e^{i\vec{q} \cdot \vec{u}(\vec{0})} e^{-i\vec{q} \cdot \vec{u}(\vec{R})} \rangle , \quad (10.143)$$

puis effectuer la sommation sur \vec{R} .

On sait que la coordonnée et le moment conjugué d'un oscillateur harmonique peuvent s'exprimer en combinaison linéaire des opérateurs de création et d'annihilation b^\dagger et b . Les déplacements $\vec{u}(\vec{R})$ – et leurs moments conjugués – sont donc des combinaisons linéaires des opérateurs $b_{\vec{k}\lambda}^\dagger$ et $b_{\vec{k}\lambda}$ associés aux modes normaux déterminés plus haut. Il en résulte que Q est de la forme :

$$Q = \langle e^A e^B \rangle , \quad (10.144)$$

où A et B sont des formes linéaires des $b_{\vec{k}\lambda}^\dagger$ et $b_{\vec{k}\lambda}$. Ces opérateurs satisfont :

$$[b_{\vec{k}\lambda}, b_{\vec{k}'\lambda'}^\dagger] = \delta_{\vec{k}\vec{k}'} \delta_{\lambda\lambda'} , \quad (10.145)$$

de sorte que l'on peut remplacer sans complication l'exponentielle d'une somme par le produit des exponentielles. Q apparaît alors sous la forme :

$$Q = \left\langle \prod_{\vec{k}, \lambda} e^{A_{\vec{k}\lambda}} e^{B_{\vec{k}\lambda}} \right\rangle . \quad (10.146)$$

Maintenant, $A_{\vec{k}\lambda}$ et $B_{\vec{k}\lambda}$ sont des combinaisons linéaires des deux seuls opérateurs $b_{\vec{k}\lambda}^\dagger$ et $b_{\vec{k}\lambda}$. En tant que tels, ils relèvent de la formule de Glauber (voir [1], p. 376 ou [8], complément B_{II}) :

$$e^{A_{\vec{k}\lambda} + B_{\vec{k}\lambda}} = e^{A_{\vec{k}\lambda}} e^{B_{\vec{k}\lambda}} e^{-\frac{1}{2} [A_{\vec{k}\lambda}, B_{\vec{k}\lambda}]} , \quad (10.147)$$

où le commutateur est un simple nombre (qui, notamment, est insensible aux opérations de moyennes). Lue à l'envers, cette relation donne :

$$e^{A_{\vec{k}\lambda}} e^{B_{\vec{k}\lambda}} = e^{A_{\vec{k}\lambda} + B_{\vec{k}\lambda}} e^{+\frac{1}{2} [A_{\vec{k}\lambda}, B_{\vec{k}\lambda}]} . \quad (10.148)$$

Il en résulte :

$$Q = \prod_{\vec{k}, \lambda} \langle e^{A_{\vec{k}\lambda} + B_{\vec{k}\lambda}} \rangle e^{+\frac{1}{2} [A_{\vec{k}\lambda}, B_{\vec{k}\lambda}]} . \quad (10.149)$$

$A_{\vec{k}\lambda} + B_{\vec{k}\lambda}$ est à nouveau une combinaison linéaire des deux opérateurs $b_{\vec{k}\lambda}^\dagger$ et $b_{\vec{k}\lambda}$:

$$A_{\vec{k}\lambda} + B_{\vec{k}\lambda} = \gamma_{\vec{k}\lambda} b_{\vec{k}\lambda} + \mu_{\vec{k}\lambda} b_{\vec{k}\lambda}^\dagger , \quad (10.150)$$

de sorte que l'on doit finalement calculer des valeurs moyennes du genre $q = \langle e^{\gamma b + \mu b^\dagger} \rangle$ où b et b^\dagger sont relatifs à un certain mode. Explicitement :

$$q = Z^{-1} \sum_{n=0}^{+\infty} e^{-\beta \hbar \omega [n + \frac{1}{2}]} \langle n | e^{\gamma b + \mu b^\dagger} | n \rangle , \quad Z = \sum_{n=0}^{+\infty} e^{-\beta \hbar \omega [n + \frac{1}{2}]} . \quad (10.151)$$

Un calcul facile montre que⁴⁷ :

$$q = e^{\frac{1}{2}\langle(\gamma b + \mu b^\dagger)^2\rangle} . \quad (10.152)$$

Alors ($j \equiv \vec{k}, \lambda$) :

$$Q = \prod_j e^{\frac{1}{2}\langle(\gamma_j b_j + \mu_j b_j^\dagger)^2\rangle} e^{+\frac{1}{2}[A_j, B_j]} = \prod_j e^{\frac{1}{2}\langle(A_j + B_j)^2\rangle} e^{+\frac{1}{2}[A_j, B_j]} . \quad (10.153)$$

Maintenant :

$$\prod_j e^{\frac{1}{2}\langle(A_j + B_j)^2\rangle} = e^{\frac{1}{2}\sum_j \langle(A_j + B_j)^2\rangle} = e^{\frac{1}{2}\langle[\sum_j A_j + B_j]^2\rangle} . \quad (10.154)$$

La dernière égalité vient du fait que, si $j \neq j'$, $\langle A_j B_{j'} \rangle = \langle A_j \rangle \langle B_{j'} \rangle$ et que la valeur moyenne d'une fonction linéaire de B et/ou b^\dagger est nulle. Alors, (10.153) donne :

$$Q = e^{\frac{1}{2}\langle[\sum_j A_j + B_j]^2\rangle} e^{+\frac{1}{2}\sum_j [A_j, B_j]} . \quad (10.155)$$

$[A_j, B_j]$ est un scalaire et peut être reporté dans l'argument de la première exponentielle ; ce dernier devient :

$$\begin{aligned} \frac{1}{2}\sum_j \langle(A_j + B_j)^2\rangle + [A_j, B_j] &= \frac{1}{2}\sum_j \langle A_j^2 + B_j^2 + A_j B_j + B_j A_j + A_j B_j - B_j A_j \rangle \\ &= \frac{1}{2}\sum_j \langle A_j^2 + B_j^2 + 2A_j B_j \rangle . \end{aligned} \quad (10.156)$$

Au total, Q défini en (10.144) est égal à :

$$Q = e^{\frac{1}{2}\sum_j \langle A_j^2 + B_j^2 + 2A_j B_j \rangle} , \quad (10.157)$$

de sorte que $Q \stackrel{\text{def}}{=} \langle e^{i\vec{q}\cdot\vec{u}(\vec{0})} e^{-i\vec{q}\cdot\vec{u}(\vec{R})} \rangle$ est donné par :

$$\begin{aligned} Q &= \exp \left\langle -\frac{1}{2}[\vec{q}\cdot\vec{u}(\vec{0})]^2 - \frac{1}{2}[\vec{q}\cdot\vec{u}(\vec{R})]^2 + [\vec{q}\cdot\vec{u}(\vec{0})][\vec{q}\cdot\vec{u}(\vec{R})] \right\rangle \\ &= \exp \left\langle -[\vec{q}\cdot\vec{u}(\vec{0})]^2 + [\vec{q}\cdot\vec{u}(\vec{0})][\vec{q}\cdot\vec{u}(\vec{R})] \right\rangle \equiv e^{-2W} \exp \left\langle [\vec{q}\cdot\vec{u}(\vec{0})][\vec{q}\cdot\vec{u}(\vec{R})] \right\rangle . \end{aligned} \quad (10.158)$$

Dans la dernière expression, on a introduit la quantité $2W$:

$$2W = \langle [\vec{q}\cdot\vec{u}(\vec{0})]^2 \rangle . \quad (10.159)$$

e^{-2W} s'appelle le facteur de Debye - Waller. En définitive, la quantité centrale contenant la diffusion élastique prend la forme (voir (10.142)) :

$$\langle |S|^2 \rangle = N e^{-2W} \sum_{\vec{R}} e^{-i\vec{q}\cdot\vec{R}} \exp \left\langle [\vec{q}\cdot\vec{u}(\vec{0})][\vec{q}\cdot\vec{u}(\vec{R})] \right\rangle \equiv N e^{-2W} \sum_{\vec{R}} e^{-i\vec{q}\cdot\vec{R}} e^{\Phi(\vec{R})} . \quad (10.160)$$

Ce résultat montre deux effets des vibrations de réseau sur les spectres de diffraction. D'une part, l'apparition du facteur de Debye - Waller révèle une atténuation des intensités de diffraction, d'autant plus grande que la moyenne $\langle [\vec{q}\cdot\vec{u}(\vec{0})]^2 \rangle$ est élevée, c'est-à-dire que la température est haute.

D'autre part, la somme de réseau ne conduit plus à une superposition de fonctions de Dirac, en vertu de la dépendance spatiale de la moyenne $e^{\Phi(\vec{R})}$ définie en (10.160). Si cette fonction – qui est constante et égale à 1 pour le réseau immobile – vaut 1 pour la plupart des valeurs de \vec{R} , on retrouve en gros la situation du réseau fixe, les angles spécifiques de diffraction étant légèrement entachés d'un "flou" qui se traduit par un petit élargissement angulaire des taches de diffraction. À l'inverse, si $e^{\Phi(\vec{R})}$ varie significativement à l'échelle

⁴⁷La relation (10.152) est connue sous le nom d'identité de Bloch, voir par exemple [37]. On peut y voir un avatar élaboré du fait que la transformée de Fourier d'une gaussienne est une gaussienne (voir [8], complément L_V).

du réseau, l'incertitude $\delta\theta$ devient de l'ordre de 2π ; dit autrement, la diffusion coopérative perd tout caractère d'hypersélectivité⁴⁸. Le point crucial est donc l'allure de $\Phi(\vec{R})$.

À l'infini, cette fonction vaut 1 : la moyenne du produit devient le produit des moyennes pour deux points très éloignés l'un de l'autre, chaque moyenne vaut 0, l'exponentielle est donc égale à 1 :

$$e^{\Phi(\infty)} = 1 . \quad (10.161)$$

À l'opposé, en $\vec{R} = \vec{0}$, c'est la valeur quadratique moyenne des déplacements qui apparaît :

$$e^{\Phi(\vec{0})} = \exp \left\langle [\vec{q} \cdot \vec{u}(\vec{0})]^2 \right\rangle . \quad (10.162)$$

La question se résume donc à savoir sur quelle distance Φ passe d'une valeur à l'autre. On observe que ceci se produit toujours sur quelques distances atomiques à peine, de sorte que, en première approximation, Φ est presque constante : pour "presque toutes" les valeurs de \vec{R} , $e^{\Phi(\vec{R})} \simeq e^{\Phi(\infty)}$. La somme dans (10.160) se décompose comme suit :

$$\sum_{\vec{R}} e^{-i\vec{q} \cdot \vec{R}} e^{\Phi(\vec{R})} = \sum_{\vec{R}} e^{-i\vec{q} \cdot \vec{R}} e^{\Phi(\infty)} - \sum_{\vec{R}} e^{-i\vec{q} \cdot \vec{R}} [e^{\Phi(\infty)} - e^{\Phi(\vec{R})}] ; \quad (10.163)$$

la première somme est d'ordre N pour un réseau fini, et donne un produit de trois peignes de Dirac si $N \rightarrow +\infty$, reproduisant la condition de von Laue (ou de Bragg) ; au contraire, la deuxième somme est d'ordre quelques unités, et est négligeable devant la première. Il en résulte que l'incertitude angulaire reste très petite et, dans les cas usuels, quasi-inobservable : en pareil cas, l'effet des vibrations se borne à une simple réduction d'intensité sans élargissement perceptible.

Cette conclusion doit cependant être un peu tempérée eu égard à des effets de dimensionnalité : si elle est correcte pour des cristaux à trois dimensions, elle peut être invalidée à $d = 2$ (ou *a fortiori* pour un réseau unidimensionnel), pour les raisons suivantes.

L'équation (10.162) implique la valeur moyenne du carré d'une quantité qui est à nouveau une combinaison linéaire d'opérateurs de création et d'annihilation. Pour un seul oscillateur de fréquence ω , on a :

$$u = \sqrt{\frac{\hbar}{2M\omega}} (b + b^\dagger) . \quad (10.164)$$

La moyenne de u^2 est donc :

$$\langle u^2 \rangle = \frac{\hbar}{2M\omega} \frac{1}{Z} \sum_{n=0}^{+\infty} e^{-\beta[n+\frac{1}{2}]\hbar\omega} \langle b^2 + b^{\dagger 2} + bb^\dagger + b^\dagger b \rangle . \quad (10.165)$$

Un calcul simple donne le résultat⁴⁹ :

$$\langle u^2 \rangle = \frac{\hbar}{2M\omega} \coth \frac{\beta\hbar\omega}{2} . \quad (10.166)$$

Pour les combinaisons linéaires que constituent les $\vec{q} \cdot \vec{u}(\vec{0})$, on va donc voir apparaître des arguments en exponentielle qui sont des sommes analogues sur l'ensemble des modes :

$$q^2 \sum_{\vec{k}\lambda} \frac{\hbar}{2M\omega_\lambda(\vec{k})} \coth \frac{\beta\hbar\omega_\lambda(\vec{k})}{2} . \quad (10.167)$$

Fondamentalement, pour chaque branche, la somme fait apparaître une intégrale du genre :

$$\int \frac{k^{d-1} dk}{\omega_\lambda(\vec{k})} \coth \frac{\beta\hbar\omega_\lambda(\vec{k})}{2} , \quad (10.168)$$

⁴⁸On raconte que von Laue, le pionnier de la diffraction X, ne croyait pas trop à la possibilité d'observer la diffusion sélective, en raison des effets de vibration de réseau. Si l'expérience s'est révélée positive, c'est *justement* parce que, en pratique, la fonction $\Phi(\vec{R})$ est presque partout constante.

⁴⁹À la limite classique, on retrouve bien le théorème d'équipartition : $\frac{1}{2}M\langle u^2 \rangle = \frac{1}{2}k_B T$.

où d est la dimension spatiale. Cette intégrale présente une singularité en $\vec{k} = 0$ en dimension inférieure à 3. En effet, près de l'origine, l'intégrand est comme $k^{d-1}\omega_\lambda(\vec{k})^{-2} \sim k^{d-3}$; l'intégrale ne converge donc que si $d-3 > -1$ soit $d > 2$. À $d = 2$, la divergence est "douce" (logarithmique), à $d = 1$, elle est relativement dure. On sait évidemment traiter ces divergences⁵⁰ en introduisant des coupures appropriées faisant intervenir (par exemple) la taille de l'échantillon. Quoiqu'il en soit, elles constituent un signal d'alarme montrant que, dans ces systèmes de basse dimensionnalité, les corrections dues aux vibrations jouent un rôle important et donnent, ici, une véritable largeur angulaire aux taches de diffraction⁵¹.

⁵⁰dites infra-rouges car survenant à basse fréquence.

⁵¹Finalement, ce sont les phonons acoustiques qui donnent une amplification spectaculaire de l'effet vibrationnel. Ceci n'est pas surprenant : la diffusion sélective étant un phénomène coopératif impliquant l'ordre à longue distance, les phonons de grande longueur d'onde sont les plus capables de modifier sérieusement le comportement par rapport au réseau fixe.

Chapitre 11

Notions de transport dans les solides

11.1 Généralités

Dans un milieu, quel qu'il soit, l'existence d'un transport (de matière, de charge, de chaleur, *etc.*) résulte d'un déséquilibre conduisant à l'apparition de courants. Il peut s'agir de courants transitoires, quand on part d'une situation hors d'équilibre et que le système, abandonné à lui-même, relaxe vers l'équilibre thermodynamique. Par exemple, une impulsion thermique ponctuelle en temps et en espace provoque un gradient de température qui induit un courant de chaleur transitoire, disparaissant lorsque l'équilibre thermodynamique est retrouvé. Il peut aussi s'agir de courants permanents, provoqués par un déséquilibre maintenu de l'extérieur : quand on soumet un métal – volume équipotentiel à l'équilibre – à une différence de potentiel imposée de l'extérieur¹, il y apparaît un courant électrique permanent, donné par la loi d'Ohm.

D'une façon générale, la description fine d'une situation hors d'équilibre est d'une grande difficulté et passe par la manipulation d'équations contenant à la fois la dynamique (mouvement en terme de forces) et la relaxation. Les deux aspects doivent être simultanément présents : en l'absence de relaxation, le système serait purement mécanique et ne saurait revenir à un état stable, indépendant des conditions initiales. Dans un cadre classique, cette complexité est bien représentée par l'équation de Boltzmann, qui est l'équation d'évolution temporelle de la densité $\rho(\{\vec{r}\}, \{\vec{p}\}, t)$, définie dans l'espace des phases, en présence de forces données et compte tenu des collisions. La description précise de ces dernières repose crucialement sur l'hypothèse du chaos moléculaire et c'est cette hypothèse qui, de fait, introduit la flèche du temps : au sortir d'une collision, deux particules sont fortement corrélées par les lois de conservation (énergie et impulsion) ; en revanche, lorsqu'elles entrent à nouveau en collision plus tard, il s'est passé tant d'autres collisions avec d'autres particules que toute mémoire de la corrélation initiale s'est perdue. L'équation de Boltzmann est d'une grande complexité (équation aux dérivées partielles, intégrale et non-linéaire) et relève d'un certain nombre d'approximations dont les plus simples débordent déjà largement le cadre de ce cours. Enfin, dernière difficulté de taille : s'agissant de la conductivité d'un métal par exemple, il faut bien évidemment se placer dans le cadre de la Mécanique Quantique !

On s'en tiendra ici à une description élémentaire, limitée à la description de la conductivité électrique, où les phénomènes dissipatifs sont introduits "à la main", sur la base d'arguments plausibles. Ce choix pragmatique permet d'incorporer un aspect essentiel, déjà rencontré dans le modèle simpliste de Drude : si les porteurs de charge avaient un mouvement purement mécanique, il ne saurait y avoir de loi d'Ohm – donc de conductivité – puisque, en pareil cas, ces porteurs sont uniformément accélérés par un champ extérieur. Les collisions, de quelque nature qu'elles soient², jouent donc un rôle fondamental : leur effet antagoniste vient limiter l'accélération due au champ externe, pour produire finalement une vitesse constante, donc un courant, donc une conductivité finie, σ . En définitive, la description qui suit précise d'abord le mouvement d'un électron

¹se traduisant en pratique par des conditions aux limites.

²Il peut s'agir de collisions entre les porteurs de charge (les électrons usuellement) et des défauts du cristal (lacunes, impuretés, *etc.*), toujours présents en pratique. Conceptuellement, on peut imaginer un cristal absolument parfait ; mais même dans ce cas, la possibilité de dégradation existe encore, par collision entre les électrons et les phonons.

entre deux collisions, puis ajuste le cadre théorique pour prendre compte *a minima* la dissipation provoquée par les collisions. La modestie de cet objectif ne doit pas faire penser qu'il s'agit d'une entreprise un peu vaine : les résultats obtenus dans ce cadre étroit, complétés par des arguments physiques qualitatifs, fournissent de précieux renseignements et mettent en évidence des phénomènes intéressants et aisément observables.

Il reste que la description quantique (purement mécanique) du mouvement d'un électron est encore fort complexe. On se bornera à présenter le modèle dit semi-classique, dont la justification précise est longue et difficile, mais dont la validité est surprenante. C'est ceci qui en fait la valeur.

L'exposé suivant procède donc en deux étapes principales :

1. dans un premier temps (section 11.2), on construit une description purement mécanique en termes de forces, incorporant l'essentiel des caractéristiques du mouvement quantique d'une particule dans un potentiel périodique.
2. dans un deuxième temps (section 11.4), la dissipation est introduite "à la main" par des arguments physiques plausibles.

11.2 Le modèle semi-classique

Commençons par réinterpréter le résultat obtenu antérieurement donnant la vitesse d'un électron décrit par l'état de Bloch $\psi_{\vec{k}_0 n}$, associé à la valeur propre $\varepsilon_n(\vec{k}_0)$ du Hamiltonien ; on a vu que la moyenne (quantique) de la vitesse de l'électron dans cet état est donnée par :

$$\vec{v}_n(\vec{k}_0) \equiv \langle \psi_{n \vec{k}_0} | \frac{\vec{p}}{m} | \psi_{n \vec{k}_0} \rangle = \hbar^{-1} \left(\frac{\partial \varepsilon_n}{\partial \vec{k}} \right)_{\vec{k}_0}, \quad (11.1)$$

qui est finalement l'expression de la vitesse de groupe associée à la loi de dispersion $\varepsilon_n(\vec{k})$. En effet, soit un paquet d'ondes formé sur des états de Bloch tous relatifs à une bande d'indice n donné :

$$\Psi(\vec{r}, t) = \sum_{\vec{k}} A(\vec{k}) \psi_{\vec{k}}(\vec{r}) e^{-i\omega(\vec{k})t} \quad (\varepsilon = \hbar\omega), \quad (11.2)$$

où la somme est à prendre au sens large (sommation ou intégration). L'amplitude $A(\vec{k})$ est caractérisée par une largeur en \vec{k} , Δk . Autrement dit, en désignant par \vec{k}_0 le vecteur où A est maximum :

$$A(\vec{k}) \simeq 0 \text{ si } \|\vec{k} - \vec{k}_0\| > \Delta k. \quad (11.3)$$

Si Δk est très petit par rapport à la taille de la 1^{ère} zone de Brillouin³, la vitesse de groupe de ce paquet d'ondes n'est autre que la vitesse donnée par (11.1) – résultat d'ailleurs quasi-évident : avec un tout petit Δk , l'état $\Psi(\vec{r}, t)$ est très voisin de l'état de Bloch (stationnaire) caractérisé par \vec{k}_0 . Bien évidemment, un tel paquet d'ondes a une largeur spatiale Δr donnée en ordre de grandeur par $\frac{1}{\Delta k} \gg a$: il s'étend donc sur un très grand nombre de mailles primitives de \mathcal{B} .

En définitive, raisonner – comme on le fera dans la suite – avec le seul ingrédient $\varepsilon_n(\vec{k})$, c'est finalement traiter implicitement des paquets d'ondes ayant les propriétés ci-dessus. D'un autre côté, on sait bien – c'est le théorème d'Ehrenfest – qu'un paquet d'ondes soumis à un champ extérieur dont les caractéristiques varient très lentement à l'échelle de ce paquet d'ondes a un mouvement quasi-classique : en effet, le centre de ce paquet suit en gros la trajectoire classique. Dans toute la suite, on supposera que les inégalités suivantes sont satisfaites :

$$a \ll \frac{1}{\Delta k} \ll l, \quad (11.4)$$

où l désigne l'échelle de variation spatiale typique du champ extérieur appliqué. Comme a est une longueur atomique et que, au contraire, l est une longueur macroscopique, on conçoit que les inégalités (11.4) sont

³Si a est l'ordre de grandeur de la dimension linéaire de la maille primitive de \mathcal{B} , cette condition se traduit par $\Delta k \ll a^{-1}$.

relativement faciles à satisfaire⁴. Ceci étant posé, on pourra convenir dans la suite de parler d'un électron "de vitesse et de position déterminées" pour dire que les incertitudes sur \vec{p} et sur \vec{r} – tout en respectant évidemment les relations de Heisenberg – sont l'une très petite par rapport à $\frac{\hbar}{a}$, l'autre très petite par rapport à la longueur macroscopique l .

Le modèle de transport semi-classique repose donc entièrement sur le seul ingrédient $\varepsilon_n(\vec{k})$: il présuppose la seule connaissance préalable de ces énergies de bande (il ne dit rien sur la façon de les obtenir). Il donne des prescriptions permettant de trouver les propriétés de transport (réponse du système à des champs (gradients électrostatiques), ou à des gradients de température). Inversement, de la mesure des propriétés de transport, il permet d'obtenir des renseignements sur la nature des énergies de bande. Ses prescriptions se résument en l'énoncé de trois affirmations :

1. l'index de bande n ne change pas (est une constante du mouvement). Ceci élimine la possibilité de transitions inter-bandes et suppose donc les champs externes petits par rapport aux champs microscopiques, en tout cas suffisamment faibles pour ne pas donner à un électron assez d'énergie pour franchir un *gap*. Avec les champs usuels, cette hypothèse est facilement satisfaite⁵.
2. en présence d'un champ électromagnétique ($\vec{\mathcal{E}}(\vec{r}, t)$, $\vec{\mathcal{B}}(\vec{r}, t)$), les équations (du mouvement) donnant la variation de la position \vec{r} et de la pseudo-impulsion $\hbar\vec{k}$ sont les suivantes⁶ :

$$\dot{\vec{r}} \equiv \vec{v}_n(\vec{k}) = \frac{1}{\hbar} \frac{\partial \varepsilon_n}{\partial \vec{k}}, \quad (11.5)$$

$$\hbar\dot{\vec{k}} = e \left[\vec{\mathcal{E}}(\vec{r}, t) + \vec{v}_n(\vec{k}) \times \vec{\mathcal{B}}(\vec{r}, t) \right]. \quad (11.6)$$

La première affirmation est visiblement nécessaire pour que ces équations soient utilisables (il faut bien savoir ce qu'est la vitesse \vec{v}_n).

3. le remplissage des bandes s'effectue conformément au principe de Pauli, en suivant la distribution de Fermi - Dirac :

$$f(\varepsilon_n(\vec{k})) = \frac{1}{e^{\beta[\varepsilon_n(\vec{k}) - \mu]} + 1} \quad (\beta = (k_B T)^{-1}). \quad (11.7)$$

Bien évidemment, comme les états de l'électron sont définis à une translation près dans $\vec{\mathcal{B}}$, les deux états \vec{k} et $\vec{k} + \vec{K}$ sont identiques : il ne peut exister deux électrons ayant le même indice de bande, n , le même spin et ne différant que par un vecteur \vec{K} . Ceci étant, pour un nombre donné d'électrons (par unité de volume, par exemple), les états sont remplis, à température nulle, jusqu'à un certain niveau d'énergie ε_F .

Avant de poursuivre quelques commentaires s'imposent. L'équation (11.5) traduit la propriété d'un paquet d'ondes ayant une faible dispersion en \vec{k} , et était prévisible ; on peut la considérer comme une équation de Hamilton, $\varepsilon_n(\vec{k})$ jouant le rôle d'un Hamiltonien. L'équation (11.6), en revanche, peut conduire à une méprise : en dépit des apparences, elle n'est pas la version quantique d'une équation de Hamilton, car le premier membre n'est pas l'impulsion de l'électron (ou, plus précisément, sa valeur moyenne dans un paquet d'ondes peu dispersé). Outre les forces extérieures, l'électron est en outre soumis au potentiel du réseau, complètement pris en compte (à noyaux fixés) par l'intervention de $\varepsilon_n(\vec{k})$. D'ailleurs, comme \vec{k} est un bon nombre quantique, c'est une constante du mouvement en l'absence de forces extérieures, *i. e.* en l'absence de forces non incluses dans le Hamiltonien conduisant aux lois de dispersion exprimées à l'aide de \vec{k} . Dit autrement, \vec{k} étant fixé, l'énergie est aussi fixée, ce qui est bien le moins pour un état propre de H ; seule une force extérieure peut changer l'énergie d'un électron.

⁴On sait bien que les paquets d'ondes s'élargissent au cours du temps. La largeur Δr peut donc se rapprocher de l . On admettra que tout intervalle de temps considéré est borné de sorte que les inégalités (11.4) sont toujours satisfaites – gardant à l'esprit que, de toute façon, le modèle ici développé ne décrit dans le détail que le mouvement entre deux collisions.

⁵On se souvient des ordres de grandeurs trouvés lors de l'étude de l'effet Stark pour un atome : un champ de 10^4 V/cm est ridiculement faible par rapport à un champ intra-atomique.

⁶Avec l'hypothèse $\Delta k \ll a^{-1}$, on confond \vec{k} et sa valeur moyenne, ce qui met finalement le paquet d'ondes (11.2) au second plan. L'équation (11.5) doit être considérée comme la *définition* de $\dot{\vec{r}}$, étant entendu que \vec{r} désigne en fait la *moyenne* (quantique) de l'observable position dans un paquet d'ondes ayant les propriétés ci-dessus.

La limite de validité des équations semi-classiques est difficile à cerner en détail. On se bornera à énoncer quelques faits d'évidence. Comme ce modèle interdit les transitions inter-bandes, toute énergie liée aux champs externes doit être très petite par rapport aux énergies de *gap*. Ces dernières sont d'autant plus grandes que le potentiel de réseau est grand ; autrement dit, une première condition de validité est que celui-ci doit être assez intense. Par ailleurs, avec les modules des champs \mathcal{E} et \mathcal{B} , on peut former deux énergies, $|e|\mathcal{E}a$ et $\hbar(|e|\mathcal{B}/m) \equiv \hbar\omega_c$; les deux conditions suivantes doivent donc être *a priori* vérifiées :

$$|e|\mathcal{E}a \ll E_{\text{gap}} \quad \hbar\omega_c \ll E_{\text{gap}} . \quad (11.8)$$

Ces inégalités laissent de côté une énergie pourtant importante, ε_F , qui doit apparaître dans les conditions de validité. Une analyse plus subtile montre que les bonnes conditions sont plus précisément :

$$|e|\mathcal{E}a \ll \frac{E_{\text{gap}}^2}{\varepsilon_F} \quad (11.9)$$

et :

$$\hbar\omega_c \ll \frac{E_{\text{gap}}^2}{\varepsilon_F} . \quad (11.10)$$

La condition (11.9) est usuellement satisfaite ; en prenant des valeurs typiques $E_{\text{gap}} \sim 0.1$ eV et $\varepsilon_F \sim 1$ eV, le rapport de droite vaut $\sim 10^{-2}$ eV. Le premier membre vaut $\sim 10^{-10} \mathcal{E}$ eV et ne devient comparable au second que pour $\mathcal{E} \sim 10^8$ V/m, une valeur gigantesque⁷. Au contraire, la condition (11.10) est violée pour des champs magnétiques que l'on peut réaliser facilement (~ 1 T)⁸.

Comme en général $\frac{E_{\text{gap}}}{\varepsilon_F}$ est petit devant 1, les conditions (11.9) et (11.10) sont plus drastiques que (11.8). Ceci ne doit pas surprendre : finalement, cette dernière condition ne s'appuie que sur la nécessité de la pertinence d'une seule bande (1^{ère} affirmation), qui ne suffit pas à justifier les équations du mouvement (11.5) et (11.6).

Enfin, avec des champs externes variables dans le temps à la pulsation ω et dans l'espace à la longueur d'onde λ , il existe des contraintes du genre :

$$\hbar\omega \ll E_{\text{gap}} \quad \lambda \gg a . \quad (11.11)$$

Il convient d'avoir toujours à l'esprit que, à température finie, pour des raisons physiques évidentes, seuls les électrons situés au voisinage de l'énergie de Fermi dans un intervalle d'ordre $k_B T$ sont concernés par le transport et y participent effectivement.

11.3 Conséquences du modèle semi-classique

Une première conséquence des hypothèses définissant le modèle semi-classique est l'inertie absolue de toute bande complète. Séparée en énergie de toutes les autres et pleine par hypothèse, il ne peut y avoir de mouvement électronique : dans cette bande, tous les états sont remplis, de sorte que tout "courant" entre états apparaissant dans un sens est forcément compensé par un courant identique mais dirigé en sens contraire – dont le mécanisme sera précisé ci-dessous⁹. Inversement, la conduction apparaît due entièrement aux bandes incomplètes, là où les électrons peuvent trouver des places libres, moyennant un peu d'énergie supplémentaire apportée par les champs externes.

Dans chaque bande, le nombre d'états est égal à deux fois le nombre de mailles primitives¹⁰ de l'échantillon sur lequel se referment les conditions cycliques. En effet, pour un échantillon macroscopique, soit N_i le nombre

⁷Bien sûr, avec des champs électriques très intenses, on peut provoquer un phénomène de claquage (*electric breakdown*). Le champ est alors assez fort pour provoquer des passages d'une bande à l'autre ; l'énergie d'un électron peut alors augmenter vertigineusement, d'où le claquage conduisant à un plasma.

⁸et on parle alors de claquage magnétique (*magnetic breakdown*).

⁹et repose sur le fait que \vec{k} est défini à un vecteur du réseau réciproque près.

¹⁰égal au nombre de nœuds du réseau direct.

points suivant la direction \vec{a}_i ; le nombre de mailles primitives est $N = N_1 N_2 N_3$. Par ailleurs, la condition cyclique parallèlement à la direction \vec{a}_i est :

$$e^{i\vec{k}\cdot N_i \vec{a}_i} = 1 \iff \vec{k}\cdot\vec{a}_i = \text{entier} \times \frac{2\pi}{N_i a_i} \iff k_i = n_i \frac{2\pi}{N_i a_i} \quad (n_i = 1, 2, \dots, N_i) . \quad (11.12)$$

Fixer \vec{k} (donc l'état de Bloch, pour un indice de bande n donné), c'est se donner les trois composantes, qui peuvent être choisies de $N_1 N_2 N_3 = N$ façons, qui est aussi le nombre de mailles primitives dans l'échantillon.

Il ne peut donc exister de bandes vides et pleines que pour les solides ayant un nombre pair d'électrons par maille élémentaire. La réciproque est fautive : si différentes bandes se recouvrent en énergie, un solide à deux électrons par maille peut être pour cette raison un conducteur (c'est d'ailleurs souvent le cas). Par exemple, pour un cristal carré à deux dimensions contenant deux électrons par maille, la sphère (cercle) de Fermi déborde de BZ1, de sorte que deux bandes se trouvent partiellement remplies et le solide est conducteur.

11.3.1 Mouvement dans un champ électrique constant dans le temps

En présence d'un champ électrique constant, les équations semi-classiques deviennent :

$$\dot{\vec{r}} \equiv \vec{v}_n(\vec{k}) = \frac{1}{\hbar} \frac{\partial \varepsilon_n}{\partial \vec{k}} , \quad (11.13)$$

$$\hbar \dot{\vec{k}} = e \vec{\mathcal{E}}(\vec{r}) . \quad (11.14)$$

Il en résulte que :

$$\vec{k}(t) - \vec{k}(0) = \frac{e}{\hbar} \vec{\mathcal{E}}(\vec{r}) t . \quad (11.15)$$

L'écart entre le vecteur d'onde à l'instant t et sa valeur de départ augmente donc linéairement avec le temps ; compte tenu des prescriptions antérieures permettant de considérer le modèle semi-classique comme un prérequis du transport, cette variation dans le temps n'a de sens que si t reste petit devant le temps τ séparant deux collisions successives (ou en tout cas n'est pas plus grand que τ). Ceci permet de se faire une idée numérique de l'augmentation Δk de \vec{k} sous l'effet du champ. En prenant $\mathcal{E} = 10$ V/cm et $\tau = 10^{-14}$ s, on trouve $\Delta k \sim e \frac{\mathcal{E}\tau}{\hbar} \sim 10^2$ cm⁻¹, alors que la dimension linéaire de la zone est $\sim 10^8$ cm⁻¹. Cette variation – la seule qui ait un sens avec la condition $t < \tau \sim 10^{-13}$ s – est donc microscopique comparée à la taille de la zone. D'ailleurs, l'augmentation d'énergie correspondante, $\frac{\hbar^2 (\Delta k)^2}{2m}$ est infime est vaut $\sim \frac{\hbar^2}{2ma_0^2} (a_0 \Delta k)^2 \sim 13.6 \times (0.5 \cdot 10^{-6})^2$ eV $\sim 10^{-12}$ eV.

À ce stade, il faut à nouveau se souvenir que $\hbar \vec{k}$ n'est pas l'impulsion de l'électron et que, par voie de conséquence, $\frac{\hbar \vec{k}}{m}$ n'est pas la vitesse de ce dernier, donnée au contraire par (11.13). Ici, quand t augmente, \vec{k} augmente : la variation de la vitesse avec le temps est donc dans le même sens que sa variation vis-à-vis de \vec{k} . Soit maintenant une bande de dispersion $\varepsilon_n(\vec{k})$ dont la variation, le long d'une direction de $\vec{\mathcal{B}}$ a l'allure d'un cosinus inversé, minimum en centre de zone, maximum en bord de zone. Quand on part du centre, ε_n augmente, sa dérivée est positive, puis s'annule au bord de zone ; comme \vec{k} et t varient dans le même sens, on peut dire que l'accélération est positive jusqu'au milieu de zone, négative ensuite. En quelque sorte, \vec{k} augmentant uniformément (tout comme le temps), la vitesse augmente (l'accélération est donc positive), puis diminue (l'accélération devient négative) ; en particulier, près du bord de zone, l'accélération est dirigée en sens inverse de la force, exactement comme si l'électron avait une masse négative (cette analyse est une autre justification de la relation exprimant la masse effective de l'électron comme l'inverse de la dérivée seconde de ε_n , au facteur \hbar^2 près). L'explication de ce comportement paradoxal est toujours la même : la force totale agissant sur l'électron ne se réduit pas à celle du champ extérieur appliqué mais contient également le potentiel périodique de réseau, incorporé dans la loi de dispersion $\varepsilon_n(\vec{k})$. En pratique, comme l'augmentation de \vec{k} est limité par le fait que l'on doit raisonner à une échelle de temps petite par rapport au temps de collision, on se borne à retenir qu'un électron partant près du centre de zone est accéléré par le champ, cependant qu'un autre partant tout près du bord de zone a, de fait, une masse négative, est envoyé par le champ vers d'autres états de masse négative et, au total, est freiné par l'introduction de ce dernier (l'énergie de l'électron commence par croître jusqu'au bord de zone puis, se retrouvant avec l'impulsion $-\frac{\vec{K}}{2}$, l'électron voit son énergie diminuer).

On peut d'ailleurs à ce propos justifier la notion de réflexion de Bragg. Quand \vec{k} augmente le long d'une direction de $\vec{\mathcal{B}}$, son extrémité se rapproche d'un plan de Bragg $\mathcal{P}_{\vec{K}}$. Quand \vec{k} est précisément égal à $\frac{\vec{K}}{2}$ ($\vec{k} \perp \mathcal{P}_{\vec{K}}$), comme $\vec{k} = +\frac{\vec{K}}{2}$ est défini à un vecteur de $\vec{\mathcal{B}}$ près, $\vec{k} - \vec{K} = -\frac{\vec{K}}{2}$ représente exactement le même état que \vec{k} . L'image que l'on peut se forger est bien celle d'une réflexion (comme sur un miroir), la vitesse changeant de signe¹¹.

Le courant induit par le champ extérieur se calcule à partir de la vitesse des électrons, chaque état \vec{k} pouvant abriter deux électrons de spins opposés. On a donc :

$$\vec{j} = 2 \int \frac{d^3k}{(2\pi)^3} e f(\varepsilon(\vec{k})) \vec{v}(\vec{k}) \quad (11.16)$$

soit, à température nulle, f devenant une marche :

$$\vec{j} = 2 \int_{\text{états occupés}} \frac{d^3k}{(2\pi)^3} e \vec{v}(\vec{k}) . \quad (11.17)$$

Cette intégrale est nulle si on somme sur toute une bande :

$$2 \int_{\text{bande complète}} \frac{d^3k}{(2\pi)^3} e \vec{v}(\vec{k}) = 0 , \quad (11.18)$$

– puisque les deux états \vec{k} et $-\vec{k}$ apparaissent à égalité et que \vec{v} est impair en tant que gradient d'une fonction paire – d'où, pour bande partiellement :

$$\int_{\text{états occupés}} \frac{d^3k}{(2\pi)^3} e \vec{v}(\vec{k}) + \int_{\text{états vides}} \frac{d^3k}{(2\pi)^3} e \vec{v}(\vec{k}) = 0 , \quad (11.19)$$

Pour une bande incomplètement remplie, on peut donc tout autant écrire (11.17) sous la forme :

$$\vec{j} = 2 \int_{\text{états vides}} \frac{d^3k}{(2\pi)^3} (-e) \vec{v}(\vec{k}) . \quad (11.20)$$

Ce résultat d'apparence banale est important ; il montre que le courant est le même que si les états occupés étaient vides et si tous les états vides étaient occupés par des particules de charge opposée à l'électron. Ces particules sont appelées *trous*. Ainsi, selon la commodité, on pourra raisonner soit avec les électrons d'une bande, soit avec leurs "antiparticules".

11.3.2 Mouvement dans un champ magnétique constant

En présence d'un champ magnétique constant¹² $\vec{\mathcal{B}}$ parallèle à Oz , les équations semi-classiques se réduisent à :

$$\dot{\vec{r}} \equiv \vec{v}_n(\vec{k}) = \frac{1}{\hbar} \frac{\partial \varepsilon_n}{\partial \vec{k}} , \quad (11.21)$$

$$\hbar \dot{\vec{k}} = e \vec{v}_n(\vec{k}) \times \vec{\mathcal{B}}(\vec{r}) \quad (11.22)$$

La seconde équation montre que $\dot{\vec{k}}$ est perpendiculaire à $\vec{\mathcal{B}}$, autrement dit, la composante de \vec{k} le long du champ est une constante du mouvement. Techniquement, le produit scalaire $\vec{k} \cdot \vec{\mathcal{B}}$ a une dérivée nulle :

$$\frac{d}{dt} \vec{k} \cdot \vec{\mathcal{B}} = 0 . \quad (11.23)$$

¹¹Il s'agit bien seulement d'une image, puisque précisément $\hbar \vec{k}$, qui change de signe, n'est pas la vraie impulsion de l'électron, donc $\frac{\hbar \vec{k}}{m}$ n'est pas la vitesse. Celle-ci est toujours $\vec{v}_n(\vec{k})$ et s'annule continûment en t_0 , si t_0 désigne l'instant où $\vec{k}(t)$ arrive en $\frac{\vec{K}}{2}$. Il reste que comme $\varepsilon(\vec{k})$ (et sa dérivée) sont des fonctions périodiques en \vec{k} , la (vraie) vitesse oscille : là où la concavité de ε est dirigée vers le bas, l'accélération est en sens contraire de la force (oscillations de Bloch)

¹²Télescopage de notations ! Ne pas confondre le module du champ magnétique, \mathcal{B} , avec le réseau de Bravais !!!

L'énergie moyenne est aussi une constante du mouvement. En effet, la variation de \vec{k} avec t entraîne *a priori* celle de $\varepsilon_n(\vec{k})$:

$$\frac{d}{dt} \varepsilon_n(\vec{k}) = (\vec{\nabla}_{\vec{k}} \varepsilon_n) \frac{d\vec{k}}{dt}, \quad (11.24)$$

mais compte tenu de (11.21) et de (11.22), il vient :

$$\frac{d}{dt} \varepsilon_n(\vec{k}) = e\hbar\vec{v}_n \cdot [\vec{v}_n \times \vec{B}] = 0. \quad (11.25)$$

Tout comme en Dynamique classique, la force magnétique ne travaille pas.

Le vecteur \vec{k} est cependant variable dans le temps ; on peut définir la trajectoire dans l'espace \vec{k} comme la ligne décrite par l'extrémité de ce vecteur au cours du temps. L'existence des deux constantes du mouvement k_z et ε_n entraîne que ces trajectoires sont les intersections entre des plans perpendiculaires à \vec{B} et les surfaces d'énergie constante. Le sens de rotation sur ces orbites est fixé par \vec{k} (qui donne la vitesse sur la trajectoire dans l'espace \vec{k}) et donc, d'après (11.22), dépend de \vec{v} , lui-même déterminé par l'orientation de $\vec{\nabla}_{\vec{k}} \varepsilon_n$ (qui détermine la normale à la surface d'énergie constante). Pour l'électron (de charge négative !), la rotation de \vec{k} est dans le sens positif autour du champ \vec{B} si les états extérieurs à la surface ont une énergie inférieure à celle de ceux situés à l'intérieur.

Bien évidemment, ces orbites ne sont pas nécessairement fermées, d'autant plus que la topologie des surfaces (de Fermi) peut être très complexe. Par exemple, dans un schéma en zones étendues, la surface de Fermi peut ressembler à des boules vaguement sphériques (centrées sur les nœuds de \vec{B}) reliées par des tubes perpendiculaires aux plans de Bragg¹³ ; en pareil cas, suivant l'orientation du champ, les orbites peuvent être ouvertes ou fermées. Ceci produit des phénomènes facilement observables (magnétorésistance).

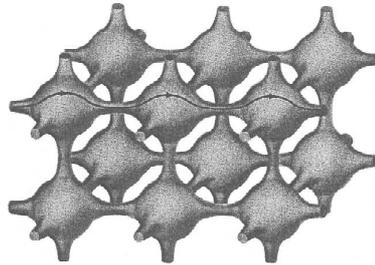


Figure 11.1: Orbites ouvertes pour un cubique simple convenablement orienté par rapport au champ magnétique (figure extraite de [38]).

Il est également possible de trouver la trajectoire dans l'espace physique. Dans un premier temps, on peut obtenir facilement la variation de $\vec{r}_\perp = \vec{r} - \frac{\vec{B}\vec{r}}{B}$, qui donne la projection du mouvement dans un plan perpendiculaire au champ. En effet, en posant $\vec{r} = \vec{r}_\perp + \vec{r}_\parallel$, le deuxième membre de (11.22) devient :

$$\hbar\dot{\vec{k}} = e\dot{\vec{r}}_\perp \times \vec{B}, \quad (11.26)$$

et une simple intégration en temps de (11.22) donne :

$$\frac{\hbar}{e} [\vec{k}(t) - \vec{k}(0)] = [\vec{r}_\perp(t) - \vec{r}_\perp(0)] \times \vec{B}. \quad (11.27)$$

La solution de cette équation est¹⁴ :

$$\vec{r}_\perp(t) - \vec{r}_\perp(0) = \frac{\hbar}{eB^2} [\vec{k}(t) - \vec{k}(0)] \times \vec{B}. \quad (11.29)$$

¹³Voir les livres pour des dessins plus explicites, et notamment [38] dont est extraite la fig. 11.1.

¹⁴Soit l'équation :

$$\vec{A} \times \vec{B} = \vec{C}. \quad (11.28)$$

En posant $\vec{A} = \vec{A}_\perp + \vec{A}_\parallel$, (11.28) devient $\vec{A}_\perp \times \vec{B} = \vec{C}$. En construisant la figure illustrant cette relation vectorielle, on voit immédiatement que \vec{A}_\perp , \vec{B} et \vec{C} forment un trièdre trirectangle et que \vec{A}_\perp est égal à $\frac{\vec{B} \times \vec{C}}{\|\vec{B}\|^2}$.

Il en résulte que la projection des orbites “spatiales” sur un plan perpendiculaire au champ s’obtient en tournant de $\frac{\pi}{2}$ autour du champ les orbites pour \vec{k} et en les multipliant par le facteur dimensionnel $\frac{\hbar}{eB^2}$. Dans le cas des électrons libres, les surfaces d’énergie constante sont des sphères dont l’intersection avec les plans perpendiculaires au champ sont des cercles ; en tournant ces cercles de $\frac{\pi}{2}$, on obtient encore des cercles, en conformité avec le résultat classique. En présence du potentiel périodique, les orbites peuvent être très compliquées, en résultat de la topologie complexe des surfaces d’énergie constante.

L’obtention de la variation de la troisième coordonnée est moins simple, puisqu’il faut résoudre :

$$z(t) - z(0) = \int_0^t v_z(\vec{k}(t')) dt' , \quad (11.30)$$

avec :

$$v_z(\vec{k}(t)) = \hbar^{-1} \frac{\partial \varepsilon_n}{\partial k_z}(\vec{k}(t)) . \quad (11.31)$$

Si l’électron était libre ($\varepsilon(\vec{k}) = \frac{\hbar^2 k^2}{2m}$), le second membre de (11.31) serait :

$$v_z(\vec{k}(t)) = \frac{\hbar k_z(t)}{m} \quad (11.32)$$

et comme k_z est, en toute hypothèse, une constante du mouvement, on trouverait immédiatement :

$$z(t) - z(0) = \frac{\hbar k_z}{m} t , \quad (11.33)$$

reproduisant le mouvement uniforme classique le long du champ. En présence du potentiel périodique, la symétrie sphérique est perdue et $\frac{\partial \varepsilon_n}{\partial k_z}$ contient en général les trois composantes k_x , k_y et k_z .

11.4 Conductivité statique statique d’un métal

Le modèle classique de Drude met bien en évidence un fait évident physiquement : si les particules responsables du transport de courant électrique n’étaient pas freinées, il ne pourrait y avoir de loi d’Ohm, puisque sous l’effet du champ électrique, leur vitesse augmenterait linéairement en temps. L’existence d’un courant constant dans un conducteur en régime permanent soumis à un champ électrique constant exige que la vitesse des porteurs – en tout cas celle qui produit le courant – est elle-même constante.

Le modèle semi-classique exposé ci-dessus prétend seulement donner la dynamique des électrons dans un cristal entre deux collisions, quelle que soit la nature de ces dernières (défauts structuraux et/ou chimiques, phonons, *etc.*). Il a permis d’obtenir des équations dynamiques où $\varepsilon(\vec{k})$ joue le rôle d’un Hamiltonien au sens classique et où $\hbar \vec{k}$ a toutes les apparences d’un moment conjugué (impulsion). Ces analogies permettent de construire une théorie simple du transport électrique, en associant les équations dynamiques semi-classiques précédentes à des termes représentant des collisions au sens large et selon une méthode universelle, suivie notamment en Mécanique Statistique classique et adaptée au cadre quantique remodelé par le modèle semi-classique.

La description mécanique précise des collisions¹⁵ est impossible en pratique, compte tenu du fait qu’elle exige la connaissance de l’état initial d’un système contenant un nombre gigantesque de degrés de liberté. Cette impossibilité, bien connue dans le domaine classique, vaut tout autant en Physique Quantique. D’ailleurs, quand bien même elle serait envisageable, elle ne serait d’aucun intérêt en fournissant une quantité d’information inutilisable et pléthorique – alors que l’expérience révèle qu’il suffit de connaître un très petit nombre de paramètres pour décrire un système macroscopique. Il convient donc de s’y prendre autrement, en recourant à une description statistique reposant sur des hypothèses plausibles physiquement incorporant de façon effective les degrés de

¹⁵ Leur nature mérite d’être précisée physiquement. Il peut s’agir de diffusion sur des impuretés inévitables, d’une part, de la mise en vibration des ions du réseau d’autre part. Ce dernier mécanisme est toujours présent et donne lieu à l’émission ou la destruction de phonons.

liberté laissés en coulisse. La quantité centrale est alors la fonction de répartition donnant, ρ , à un certain instant t la (densité de) probabilité de trouver un système avec des valeurs données pour les grandeurs dynamiques précisant son état. Ici, chaque électron "semi-classique" est défini par les deux vecteurs \vec{r} et \vec{k} ; la fonction de répartition est donc notée plus précisément $\rho(\vec{r}, \vec{k}, t)$. Le nombre dN d'électrons ayant à l'instant t la position \vec{r} à d^3r près et le vecteur d'onde \vec{k} à d^3k près est :

$$dN = \rho(\vec{r}, \vec{k}, t) 2 \frac{d^3r d^3k}{(2\pi)^3} , \quad (11.34)$$

où le facteur 2 prend en compte le spin. À l'équilibre thermodynamique, pour les fermions que sont les électrons, ρ n'est autre que la distribution de Fermi - Dirac donnée en (11.7).

En l'absence de collisions, $\rho(\vec{r}, \vec{k}, t)$ obéit à une équation purement mécanique (l'un des avatars de l'équation de Liouville) que l'on écrit simplement sur la base d'un argument de conservation. En effet, l'intégrale de ρ est une constante dans le temps, égale au nombre total d'électrons :

$$\int_{\mathbb{R}^3} d^3r \int_{\mathbb{R}^3} \frac{d^3k}{4\pi^3} \rho(\vec{r}, \vec{k}, t) = N . \quad (11.35)$$

Dès lors, le mouvement des N électrons (dans l'espace des phases !) est analogue à celui d'un fluide sans perte pour lequel il existe une équation locale de conservation :

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{\mathcal{J}} = 0 . \quad (11.36)$$

$\vec{\mathcal{J}}$ désigne le courant dans l'espace des phases ; c'est le produit de la densité ρ par la vitesse de déplacement d'un point de l'espace des phases ; un tel point est défini par les 6 composantes des vecteurs \vec{r} et \vec{k} , c'est donc un vecteur à six composantes – les composantes de la vitesse correspondante sont donc les six composantes de $\dot{\vec{r}}$ et de $\dot{\vec{k}}$. Quant à la divergence, c'est le gradient à six composantes, réunion de $\vec{\nabla}_{\vec{r}}$ et de $\vec{\nabla}_{\vec{k}}$. Dès lors, (11.36) s'écrit :

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}_{\vec{r}}(\rho \dot{\vec{r}}) + \vec{\nabla}_{\vec{k}}(\rho \dot{\vec{k}}) = 0 . \quad (11.37)$$

Les deux gradients donnent un premier terme en ρ :

$$\rho (\vec{\nabla}_{\vec{r}} \dot{\vec{r}} + \vec{\nabla}_{\vec{k}} \dot{\vec{k}}) . \quad (11.38)$$

S'il existait un authentique Hamiltonien fixant la dynamique par les équations canoniques correspondantes, la parenthèse serait globalement nulle : le premier terme serait de la forme $\frac{\partial}{\partial q} \frac{\partial H}{\partial p}$ et le second $-\frac{\partial}{\partial p} - \frac{\partial H}{\partial q}$; ici les choses se présentent différemment. En fait, le premier terme de (11.38) est nul (ε ne dépend pas de \vec{r}) :

$$\vec{\nabla}_{\vec{r}} \dot{\vec{r}} \equiv \vec{\nabla}_{\vec{r}} \vec{v} = \hbar^{-1} \vec{\nabla}_{\vec{r}} \varepsilon(\vec{k}) \equiv 0 . \quad (11.39)$$

Par ailleurs, le second terme de (11.38) est :

$$e\hbar^{-1} \vec{\nabla}_{\vec{k}} (\vec{\mathcal{E}} + \vec{v} \times \vec{\mathcal{B}}) = e\hbar^{-1} \vec{\nabla}_{\vec{k}} (\vec{\nabla}_{\vec{k}} \varepsilon \times \vec{\mathcal{B}}) . \quad (11.40)$$

C'est un produit mixte impliquant deux fois le même vecteur ($\vec{\nabla}_{\vec{k}}$), qui est donc nul. Cette fois, la parenthèse dans (11.38) est nulle parce que chacun de ses termes l'est.

En définitive, l'équation décrivant l'écoulement du fluide décrit par les variables dynamiques *ad hoc* \vec{v} et \vec{k} obéissant aux équations du mouvement semi-classiques (11.5) et (11.6) est, en l'absence de collisions :

$$\frac{\partial \rho}{\partial t} + \dot{\vec{r}} \cdot \vec{\nabla}_{\vec{r}} \rho + \dot{\vec{k}} \cdot \vec{\nabla}_{\vec{k}} \rho = 0 . \quad (11.41)$$

Les collisions modifient position et "impulsion" et induisent des pertes au sens où un électron qui devrait se trouver en (\vec{r}, \vec{k}) à l'instant t s'il suivait les équations strictement dynamiques, ne s'y retrouve pas (mais,

bien évidemment, se trouve ailleurs), et inversement : on retrouve en (\vec{r}, \vec{k}) des électrons qui, sans collisions, seraient ailleurs. Il en résulte que, en présence des collisions, le second membre de (11.37) n'est pas nul, et est précisément égal à la variation temporelle de ρ sous l'effet des collisions ; on écrit alors :

$$\frac{\partial \rho}{\partial t} + \dot{\vec{r}} \cdot \vec{\nabla}_{\vec{r}} \rho + \dot{\vec{k}} \cdot \vec{\nabla}_{\vec{k}} \rho = \left(\frac{\partial \rho}{\partial t} \right)_{\text{coll}} . \quad (11.42)$$

Ceci n'avance guère, tant que l'on ne connaît pas le second membre. Une hypothèse raisonnable consiste à dire que cette variation sous l'effet des collisions est d'autant plus grande que ρ diffère beaucoup de la distribution d'équilibre de Fermi - Dirac, f ; on est donc tenté d'écrire :

$$\left(\frac{\partial \rho}{\partial t} \right)_{\text{coll}} \propto (\rho - f) . \quad (11.43)$$

La constante de proportionnalité est l'inverse d'un temps ; enfin, pour qu'il s'agisse d'une relaxation, il convient d'introduire le bon signe. En réfléchissant, on voit qu'une forme acceptable est :

$$\left(\frac{\partial \rho}{\partial t} \right)_{\text{coll}} = -\frac{1}{\tau} (\rho - f) , \quad (11.44)$$

où τ , positif, est appelé temps de relaxation. *A priori*, c'est une fonction de \vec{r} (si le milieu est inhomogène) et de \vec{k} (dépendance en énergie du temps de relaxation). En définitive, dans cette description, dite du temps de relaxation, l'équation d'évolution de ρ est la suivante :

$$\frac{\partial \rho}{\partial t} + \dot{\vec{r}} \cdot \vec{\nabla}_{\vec{r}} \rho + \dot{\vec{k}} \cdot \vec{\nabla}_{\vec{k}} \rho = -\frac{1}{\tau} (\rho - f) . \quad (11.45)$$

C'est cette équation que l'on va maintenant utiliser pour obtenir la conductivité électrique statique d'un métal.

Pour simplifier, on suppose que le solide est à une température donnée homogène (pas de gradient de température) et que le champ électrique appliqué est lui aussi homogène (pas de gradient de champ électrique). Enfin, on s'intéresse uniquement au régime permanent, après extinction de tous les transitoires. Ainsi, dans un conducteur homogène (ρ indépendant de \vec{r}), le régime permanent est décrit par (11.45) avec $\frac{\partial \rho}{\partial t} = 0$ et $\vec{\nabla}_{\vec{r}} \rho = 0$:

$$\dot{\vec{k}} \cdot \vec{\nabla}_{\vec{k}} \rho = -\frac{1}{\tau} (\rho - f) \iff \frac{e}{\hbar} \vec{\mathcal{E}} \cdot \vec{\nabla}_{\vec{k}} \rho = -\frac{1}{\tau} (\rho - f) , \quad (11.46)$$

où (11.6) a été utilisé pour écrire la deuxième forme. Il en résulte :

$$\rho = f - \frac{e\tau}{\hbar} \vec{\mathcal{E}} \cdot \vec{\nabla}_{\vec{k}} \rho . \quad (11.47)$$

Ceci montre que ρ diffère de la distribution de Fermi - Dirac par un terme au moins d'ordre 1 en champ. À l'ordre le plus bas, il vient donc :

$$\rho \simeq f - \frac{e\tau}{\hbar} \vec{\mathcal{E}} \cdot \vec{\nabla}_{\vec{k}} f . \quad (11.48)$$

Afin de rendre les résultats plus transparents, on utilise :

$$\vec{\nabla}_{\vec{k}} f = \left(\frac{\partial f}{\partial \varepsilon} \right) \vec{\nabla}_{\vec{k}} \varepsilon = \hbar \vec{v} \frac{\partial f}{\partial \varepsilon} , \quad (11.49)$$

de sorte que (11.48) devient :

$$\rho = f - e\tau \vec{\mathcal{E}} \cdot \vec{v} \frac{\partial f}{\partial \varepsilon} . \quad (11.50)$$

Il est bien entendu que, partout, ε désigne la fonction $\varepsilon(\vec{k})$. Maintenant, il suffit de revenir à l'expression du courant de charge par unité de volume¹⁶ :

$$\vec{j} = e \int \frac{d^3 k}{4\pi^3} \rho \vec{v} \quad (11.51)$$

¹⁶ f est ici indépendant de \vec{r} ; l'intégration sur \vec{r} donne un simple facteur extensif, le volume de l'échantillon.

pour obtenir :

$$\vec{j} = e \int \frac{d^3k}{4\pi^3} \left(f - e\tau \vec{\mathcal{E}} \cdot \vec{v} \frac{\partial f}{\partial \varepsilon} \right) \vec{v} . \quad (11.52)$$

L'intégration avec f donne évidemment zéro (pas de courant à l'équilibre en l'absence de champ !). En explicitant les produits scalaires à l'aide des coordonnées cartésiennes, il vient :

$$j_\alpha = -e^2 \sum_{\beta=x,y,z} \int \frac{d^3k}{4\pi^3} \tau v_\alpha \mathcal{E}_\beta v_\beta \frac{\partial f}{\partial \varepsilon} \equiv \sum_{\beta} \sigma_{\alpha\beta} \mathcal{E}_\beta , \quad (11.53)$$

où le tenseur de conductivité $\sigma_{\alpha\beta}$ se trouve défini par la dernière égalité, généralisant la relation usuelle et incorporant une éventuelle anisotropie de conductivité¹⁷ :

$$\vec{j} = \hat{\sigma} \vec{\mathcal{E}} \quad \hat{\sigma} = -e^2 \int \frac{d^3k}{4\pi^3} \tau(\varepsilon(\vec{k})) \vec{v} \otimes \vec{v} \frac{\partial f}{\partial \varepsilon} . \quad (11.54)$$

Dans un cristal cubique, prenant x , y et z le long des arêtes du cube, les trois éléments diagonaux $\sigma_{\alpha\alpha}$ sont égaux entre eux : le tenseur de conductivité est donc un simple scalaire, traduisant l'isotropie (réalisée aussi bien sûr dans le cas libre).

Il est possible de faire apparaître le tenseur de masse effective introduit plus haut. On note d'abord que, aux températures usuelles, le rapport $\frac{k_{\text{F}} T}{\varepsilon_{\text{F}}}$ est très petit devant 1, de sorte que, même à ces températures, la fonction f est très voisine d'une marche. Il en résulte que $\frac{\partial f}{\partial \varepsilon} \simeq -\delta(\varepsilon - \varepsilon_{\text{F}})$, ce qui autorise à sortir $\tau(\varepsilon)$ de l'intégrale et donne :

$$\hat{\sigma} \simeq -e^2 \tau(\varepsilon_{\text{F}}) \int \frac{d^3k}{4\pi^3} \vec{v} \vec{v} \frac{\partial f}{\partial \varepsilon} . \quad (11.55)$$

Par ailleurs, $\vec{v} \frac{\partial f}{\partial \varepsilon} = \hbar^{-1} \vec{\nabla}_{\vec{k}} f(\varepsilon(\vec{k}))$. Ainsi, l'élément $\alpha\beta$ de $\hat{\sigma}$ est :

$$\sigma_{\alpha\beta} = -\frac{e^2}{\hbar} \tau(\varepsilon_{\text{F}}) \int \frac{d^3k}{4\pi^3} v_\alpha \frac{\partial f(\varepsilon(\vec{k}))}{\partial k_\beta} \quad (11.56)$$

et une intégration par parties donne :

$$\sigma_{\alpha\beta} = \frac{e^2}{\hbar} \tau(\varepsilon_{\text{F}}) \int \frac{d^3k}{4\pi^3} f(\varepsilon(\vec{k})) \frac{\partial v_\alpha}{\partial k_\beta} . \quad (11.57)$$

Le tenseur de masse effective d'éléments $m_{\alpha\beta}$ est défini comme :

$$(m^{-1})_{\alpha\beta} = \hbar^{-2} \frac{\partial^2 \varepsilon}{\partial k_\alpha \partial k_\beta} = \hbar^{-1} \frac{\partial v_\alpha}{\partial k_\beta} = \frac{\partial^2 \varepsilon}{\partial k_\beta \partial k_\alpha} = \hbar^{-1} \frac{\partial v_\beta}{\partial k_\alpha} . \quad (11.58)$$

Il en résulte :

$$\sigma_{\alpha\beta} = e^2 \tau(\varepsilon_{\text{F}}) \int \frac{d^3k}{4\pi^3} f(\varepsilon(\vec{k})) (m^{-1})_{\alpha\beta} . \quad (11.59)$$

σ , tout comme m , est un tenseur symétrique. À très basse température, on peut remplacer f par la fonction de Heaviside $\Theta(\varepsilon_{\text{F}} - \varepsilon)$; alors la conductivité a pour éléments :

$$\sigma_{\alpha\beta} = e^2 \tau(\varepsilon_{\text{F}}) \int_{\text{états occupés}} \frac{d^3k}{4\pi^3} (m^{-1})_{\alpha\beta} . \quad (11.60)$$

Comme $(m^{-1})_{\alpha\beta}$ est la dérivée d'une fonction périodique, sa moyenne sur une maille primitive est nulle ; (11.60) s'écrit donc tout autant :

$$\sigma_{\alpha\beta} = e^2 \tau(\varepsilon_{\text{F}}) \int_{\text{états vides}} \frac{d^3k}{4\pi^3} (-m^{-1})_{\alpha\beta} . \quad (11.61)$$

¹⁷ σ a le bon signe puisque $\frac{\partial f}{\partial \varepsilon}$ est négatif.

Dans cette optique, ce sont les trous positifs qui conduisent le courant et ont une masse effective changée de signe par rapport à celle des électrons. Par ailleurs, si $(m^{-1})_{\alpha\beta}$ est scalaire et indépendant de \vec{k} :

$$(m^{-1})_{\alpha\beta} = \frac{1}{m_{\text{eff}}} \delta_{\alpha\beta} , \quad (11.62)$$

alors :

$$\sigma_{\alpha\beta} = \sigma \delta_{\alpha\beta} , \quad \sigma = e^2 \tau (\varepsilon_F) \frac{1}{m_{\text{eff}}} \int_{\text{états occupés}} \frac{d^3 k}{4\pi^3} = \frac{n e^2 \tau}{m_{\text{eff}}} , \quad (11.63)$$

où n est le nombre d'électrons par unité de volume : ceci n'est autre que le résultat de Drude, corrigé de la masse effective.

Bibliographie

- [1] A. Messiah, *Mécanique Quantique*, tome 1 (Dunod, Paris 1965)
- [2] H. V. McIntosh, Amer. J. Phys., **27**, 620, 1959
- [3] M. Bander et C. Itzykson, Rev. Mod. Phys., **38**, 330, 1966
- [4] L. Landau et I. Lifchitz, *Mécanique Quantique* (Mir, Moscou, 1967)
- [5] L. I. Schiff, *Quantum Mechanics* (McGraw - Hill, New York, 1968)
- [6] Cl. Cohen-Tannoudji, *Cours du Collège de France, 1973-1974*
- [7] E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1970)
- [8] Cl. Cohen-Tannoudji, B. Diu et F. Laloë, *Mécanique Quantique* (Hermann, Paris, 1973)
- [9] A. Sokal et J. Bricmont, *Impostures intellectuelles* (Odile Jacob, Paris, 1997)
- [10] *Encyclopedia Universalis*, **16**, 300 (1992).
- [11] A. Messiah, *Mécanique Quantique*, tome 2 (Dunod, Paris 1965)
- [12] A. S. Davydov, *Quantum Mechanics* (Pergamon Press, 1965)
- [13] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group*, sect. 2. 2. 1 (Addison - Wesley, New York, 1997)
- [14] H. A. Bethe et E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Springer, Berlin, 1957)
- [15] J. C. Slater, *Quantum Theory of Atomic Structure, I*, Appendice 19 (McGraw - Hill, New York, 1960)
- [16] <http://www.britannica.com/bcom/eb/article/5/0,5716,59835+1+58374,00.html>
- [17] B. Misra & E. C. G. Sudarshan, *The Zeno's paradox in quantum theory*, J. Math. Phys., **18**, 756 (1977)
- [18] A. Peres, *Zeno paradox in quantum theory*, Am. J. Phys., **48**, 931 (1980)
- [19] R. J. Cook, Phys. Scr., **T21**, 49, (1988) (cité en [20])
- [20] W. M. Itano, D. J. Heinzen, J. J. Bollinger & D. J. Wineland, *Quantum Zeno effect*, Phys. Rev. A, **41**, 2295 (1990)
- [21] Claude Cohen-Tannoudji, *Cours du Collège de France*, (1989/1990)
- [22] M. B. Plenio & P. L. Knight, *The quantum-jump approach to dissipative dynamics in quantum optics*, Rev. Mod. Phys., **70**, 101 (1998)
- [23] H. G. Dehmelt, Bull. Am. Phys. Soc., **20**, 60 (1975)

- [24] R. J. Cook & H. J. Kimble, *Possibility of Direct Observation of Quantum Jumps*, Phys. Rev. Let., **54**, 1023 (1985)
- [25] W. Nagourney, J. Sandberg & H. Dehmelt, *Shelved Optical Electron Amplifier: Observation of Quantum Jumps*, Phys. Rev. Let., **56**, 2797 (1986)
- [26] Th. Sauter, W. Neuhauser, R. Blatt & P. E. Toschek *Observation of Quantum Jumps*, Phys. Rev. Let., **57**, 1696 (1986)
- [27] J. Dalibard, Y. Castin & K. Mølmer, *Wave-Function Approach to Dissipative Processes in Quantum Optics*, Phys. Rev. Let., **68**, 580 (1992)
- [28] <http://www.cs.mcgill.ca/~crepeau/CRYPTO/Biblio-QC.html>
- [29] A. K. Ekert, *Quantum Cryptography Based on Bell's Theorem*, Phys. Rev. Let., **68**, 661 (1991)
- [30] A. K. Ekert, *La Mécanique Quantique au secours des agents secrets*, La Recherche, **22**, 790 (1991)
- [31] M. Jammer, *Le paradoxe d'Einstein - Podolsky - Rosen*, La Recherche, **111**, 510 (1980)
- [32] I. N. Levine, *Quantum Chemistry*, vol. I (Allyn & Bacon, Boston, 1970)
- [33] O. Burrau, *Kgl. Danske Vid. Selsk. Matt.-Fys. Medd.*, **7**, 14 (1927)
- [34] J. C. Slater, *Quantum Theory of Molecules and Solids*, vol. I, (McGraw-Hill, New York, 1963)
- [35] M. Gell-Mann & F. Low, *Bound States in Quantum Field Theory*, Phys. Rev., **84**, 350 (1951)
- [36] A.L.Fetter & J.D Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill Book Company, New York, 1971)
- [37] W. Marshall et S. W. Lovesey, *Theory of Thermal Neutron Scattering* (Holt, Rinehart et Winston, Oxford, 1971)
- [38] N. W. Ashcroft et N. D. Mermin, *Solid State Physics* (Clarendon Press, Oxford, New York 1976)